The Number of Random 2-SAT Solutions Is Asymptotically Log-Normal

Arnab Chatterjee
Department of Computer Science, TU Dortmund, Germany

Amin Coja-Oghlan
Department of Computer Science, TU Dortmund, Germany

Noela Müller
Department of Mathematics and Computer Science, Eindhoven University of Technology, Netherlands

Connor Riddlesden
Department of Mathematics and Computer Science, Eindhoven University of Technology, Netherlands

Maurice Rolvien
Department of Computer Science, TU Dortmund, Germany

Pavel Zakharov
Department of Computer Science, TU Dortmund, Germany

Haodong Zhu
Department of Mathematics and Computer Science, Eindhoven University of Technology, Netherlands

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.

2012 ACM Subject Classification Theory of computation → Randomness, geometry and discrete structures

Keywords and phrases satisfiability problem, 2-SAT, random satisfiability, central limit theorem

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2024.39

Category RANDOM

Related Version *Full Version*: https://arxiv.org/abs/2405.03302

Funding *Amin Coja-Oghlan*: DFG CO 646/3, DFG CO 646/5 and DFG CO 646/6.
Noela Müller: NWO Gravitation grant NETWORKS-024.002.003.
Pavel Zakharov: DFG CO 646/6
Haodong Zhu: European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 945045 and NWO Gravitation project NETWORKS-024.002.003

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, 24]. 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 [33]. 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, 15, 28]. 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, 26, 40, 41]). 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, 20].

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 [19, 31], 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 [45]. Nonetheless, Monasson and Zecchina [36] 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 $d > 2$ [19, 31]. 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

$$Z(\mathbf{\Phi}) = \exp(n\phi(d) + o(n)) \qquad \text{w.h.p.}\,, \tag{1}$$

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

However, (1) fails to identify the limiting distribution of $Z(\mathbf{\Phi})$. To be precise, since (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 [34]. 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

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

The order $\Theta(\sqrt{n})$ of fluctuations confirmed by (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, 20, 40]. 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 (2) establishes that the number of solutions fails to superconcentrate (for more details see [20]).

1.2 The main result

The formula for the standard deviation $\eta(d)$ from (2) comes in terms of a fixed point equation on a space of probability measures. Thus, let $\mathcal{P}(\mathbb{R}^2)$ 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

$$\texttt{logBP}^{\otimes}_{d,t} : \mathcal{P}\left(\mathbb{R}^2\right) \to \mathcal{P}\left(\mathbb{R}^2\right), \qquad \rho \mapsto \hat{\rho} = \texttt{logBP}^{\otimes}_{d,t}(\rho), \tag{3}$$

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

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

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

$$\rho_{d,t} = \texttt{logBP}^{\otimes}_{d,t}(\rho_{d,t}) \qquad \textit{and} \qquad \int_{\mathbb{R}^2} \|\xi\|_2^2 \mathrm{d}\rho_{d,t}(\xi) < \infty. \tag{5}$$

[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, 20].

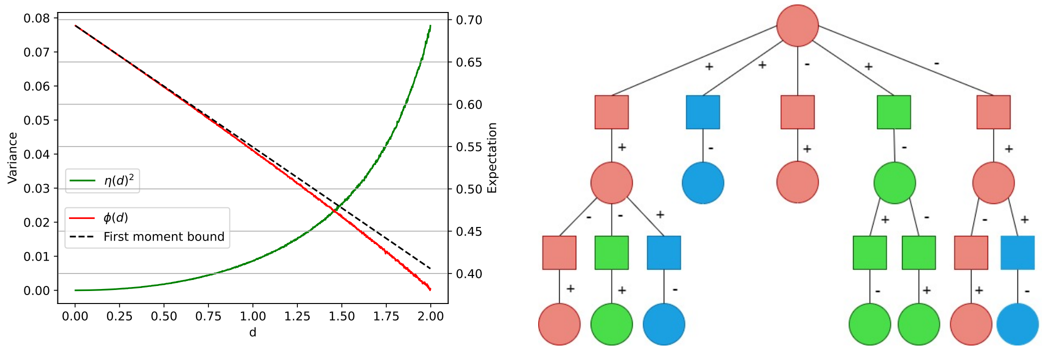

Figure 1 *Left:* Numerical approximations to the function $\phi(d)$ from (1) (red) and the variance $\eta(d)^2$ from (7) (green). The black dashed line is the first moment bound $d \mapsto \log(2) + \frac{d}{2}\log(3/4)$. *Right:* An illustration of the tree $\boldsymbol{T}^{\otimes}$ from Section 2.6.

Furthermore,

$$\lim_{n\to\infty} \frac{\log Z(\mathbf{\Phi}) - \mathbb{E}[\log Z(\mathbf{\Phi}) \mid Z(\mathbf{\Phi}) > 0]}{\sqrt{m}} = \mathbf{\Gamma}_{\eta(d)} \qquad \textit{in distribution, where} \tag{6}$$

$$\eta(d)^2 = \int_0^1 \mathcal{B}^{\otimes}_{d,t}(\rho_{d,t})\mathrm{d}t - \mathcal{B}^{\otimes}_{d,0}(\rho_{d,0}) \in (0,\infty). \tag{7}$$

The conditioning on $\log Z(\mathbf{\Phi}) > 0$ is necessary in (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 (5) ensures that the integral (7) is well-defined. Finally, (6) implies (2).

How can the formula (7) be evaluated? Because the proof of the uniqueness of the stochastic fixed point $\rho_{d,t}$ from (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* [34, Chapter 14]. Since $\mathcal{B}^{\otimes}_{d,t}(\rho_{d,t})$ 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 is to get a handle on the variance of $\log Z(\mathbf{\Phi})$ given satisfiability. The key idea, inspired by spin glass theory [18] 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 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).

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 [23, 42]. 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

$$\frac{1}{n}\log \mathbb{E}[Z(\mathbf{\Phi})] \sim \log 2 + \frac{d}{2}\log(3/4). \tag{8}$$

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) for any $0 < d < 2$. Consequently, (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 [36] 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 $\mathrm{sign}_\Phi(x, a) = \mathrm{sign}(x, a) \in \{\pm 1\}$ 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, $\mu_{x\to a}(\pm 1), \mu_{a\to x}(\pm 1) \ge 0$ 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

$$\mathrm{BP} : (\mu_{x\to a}, \mu_{a\to x})_{a, x\in\partial a} \mapsto (\hat\mu_{x\to a}, \hat\mu_{a\to x})_{a, x\in\partial a} = \mathrm{BP}((\mu_{x\to a}, \mu_{a\to x})_{a, x\in\partial a}). \tag{9}$$

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

$$\hat\mu_{a\to x}(\mathrm{sign}(x,a)) = \frac{1}{1 + \mu_{y\to a}(\mathrm{sign}(y,a))}, \quad \hat\mu_{a\to x}(-\mathrm{sign}(x,a)) = \frac{\mu_{y\to a}(\mathrm{sign}(y,a))}{1 + \mu_{y\to a}(\mathrm{sign}(y,a))}. \tag{10}$$

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

$$\hat\mu_{x\to a}(s) = \frac{\prod_{b\in\partial x\setminus\{a\}} \mu_{b\to x}(s)}{\prod_{b\in\partial x\setminus\{a\}} \mu_{b\to x}(1) + \prod_{b\in\partial x\setminus\{a\}} \mu_{b\to x}(-1)} \qquad (s \in \{\pm 1\}) \ ; \tag{11}$$

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

$$\mu_x(s) = \frac{\prod_{b\in\partial x} \mu_{b\to x}(s)}{\prod_{b\in\partial x} \mu_{b\to x}(1) + \prod_{b\in\partial x} \mu_{b\to x}(-1)} \qquad (s \in \{\pm 1\}). \tag{12}$$

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 (12) and plug them into a generic formula called the *Bethe free*

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

entropy, which yields the BP "approximation" of $\log Z(\Phi)$; an excellent exposition can be found in [34]. 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 $\mathbf{\Phi}$ contains cycles in abundance. Nonetheless, (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 also harnesses Gibbs uniqueness, let us elaborate. To mimic the local structure of $\mathbf{\Phi}$ consider a multitype Galton-Watson tree $\boldsymbol{T}$ whose types are *variable nodes* and *clause nodes* of four sub-types (s, s') with $s, s' \in \{\pm 1\}$. 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 $\ell \geq 0$ 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 $\mathbf{\Phi}$ 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 $\mathbf{\Phi}$ 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 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** ([5, Proposition 2.2]). *We have*

$$\lim_{\ell \to \infty} \mathbb{E}\left[\max_{\tau \in S(\boldsymbol{T}^{(2\ell)})} \left|\mathbb{P}\left[\boldsymbol{\sigma}_o = 1 \mid \boldsymbol{T}^{(2\ell)}, \boldsymbol{\sigma}_{\partial^{2\ell} o} = \tau_{\partial^{2\ell} o}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_o = 1 \mid \boldsymbol{T}^{(2\ell)}\right]\right|\right] = 0. \tag{13}$$

2.3 Approaching the variance

The proof of the formula (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 $\log Z(\mathbf{\Phi})$ by a long shot. Actually, it is anything but clear how even the *order* of the standard deviation of $\log Z(\mathbf{\Phi})$ could be derived along these lines. One specific problem is that the rate of convergence of (13) diminishes as d approaches the satisfiability threshold.

To tackle this challenge we devise a combinatorial interpretation of $\log^2 Z(\mathbf{\Phi})$. A key idea, which we borrow from spin glass theory [18], is to set up a family of correlated random formulas. Specifically, given integers $M, M' \geq 0$ we construct a correlated pair $(\mathbf{\Phi}_1(M, M'), \mathbf{\Phi}_2(M, M'))$ of formulas on the variable set $V_n = \{x_1, \ldots, x_n\}$ 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

$$\begin{aligned} \mathbf{\Phi}_1(M, M') &= \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_M \wedge \boldsymbol{a}'_1 \wedge \cdots \wedge \boldsymbol{a}'_{M'}, \\ \mathbf{\Phi}_2(M, M') &= \boldsymbol{a}_1 \wedge \cdots \wedge \boldsymbol{a}_M \wedge \boldsymbol{a}''_1 \wedge \cdots \wedge \boldsymbol{a}''_{M'}. \end{aligned} \tag{14}$$

Thus, the two formulas share clauses $\boldsymbol{a}_1, \dots, \boldsymbol{a}_M$. 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

$$\begin{aligned}\log Z(\boldsymbol{\Phi}_1(m,0)) \cdot \log Z(\boldsymbol{\Phi}_2(m,0)) - \log Z(\boldsymbol{\Phi}_1(0,m)) \cdot \log Z(\boldsymbol{\Phi}_2(0,m)) \qquad (15)\\ = \sum_{M=1}^{m} \log Z(\boldsymbol{\Phi}_1(M, m-M)) \cdot \log Z(\boldsymbol{\Phi}_2(M, m-M))\\ - \log Z(\boldsymbol{\Phi}_1(M-1, m-M+1)) \cdot \log Z(\boldsymbol{\Phi}_2(M-1, m-M+1)).\end{aligned}$$

If we *could* take the expectation on the l.h.s. of (15), 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 (15), 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 $\mathcal{L}$ 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.

Algorithm 1 Pessimistic Unit Clause Propagation ("`PUC`").

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

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

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

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

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

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

Clearly, trouble brews if `PUC` ends up placing both a literal l and its negation $\neg l$ into the set $\mathcal{L}$. 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 $\mathcal{C}$ 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 $\texttt{PUC}(\Phi, \mathcal{L} = \{l\})$. Let $\mathcal{C}(\Phi, \{l\})$ be the set of conflict clauses returned by PUC. Obtain the *pruned formula* $\hat{\Phi}$ from Φ by removing all clauses in $\mathcal{C}(\Phi) = \bigcup_l \mathcal{C}(\Phi, \{l\})$. Then it is easy to verify the following.

▶ **Fact 3.** *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 4.** *With probability $1 - o(n^{-1/2})$ we have $|\log Z(\hat{\mathbf{\Phi}}) - \log Z(\mathbf{\Phi})| \leq n^{1/3}$.*

2.5 Variance redux

The error bound from Proposition 4 is tight enough so that towards the proof of Theorem 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 (15), we obtain the following expression. Recalling (14), we write $\hat{\mathbf{\Phi}}_h(M, M') = \mathbf{\Phi}_h\widehat{(M, M')}$ for the formula obtained by pruning $\mathbf{\Phi}_h(M, M')$.

▶ **Lemma 5.** *Let*

$$\Delta(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M, m-M))}{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M, m-M))}{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M))}\right)\right], \tag{16}$$

$$\Delta'(M) = \mathbb{E}\left[\log\left(\frac{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M+1))}{Z(\hat{\mathbf{\Phi}}_1(M-1, m-M))}\right) \cdot \log\left(\frac{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M+1))}{Z(\hat{\mathbf{\Phi}}_2(M-1, m-M))}\right)\right]. \tag{17}$$

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

Lemma 5 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 (16)–(17). Specifically, the quotients in (16)–(17) 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 6.** *Let $1 \leq M \leq m$. 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 6 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 6.

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 $(\pm 1, \pm 1)$ 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 $s, s' \in \{\pm 1\}$.

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 $s, s' \in \{\pm 1\}$.
- An *h-distinct variable* begets $\mathrm{Po}(d/4)$ h-distinct clauses of type (s, s') for any $s, s' \in \{\pm 1\}$ $(h = 1, 2)$.
- 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 7.** *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

$$\boldsymbol{\mu}^{(2\ell)} = \left(\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_1^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}\right], \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}_2^{(2\ell)},o} = 1 \mid \boldsymbol{T}^{\otimes}\right]\right) \in (0,1)^2. \tag{18}$$

Since BP is exact on trees, we could calculate these marginals by iterating (9)–(11) 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 (18) 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 (3). To be precise, that operator expresses the updates of the log-likelihood ratios of the BP messages from (10)–(11). 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 $\rho \in \mathcal{P}(\mathbb{R}^2)$ let $\mathfrak{t}(\rho)$ be the pushforward probability measure on $(0,1)^2$.[3]

▶ **Proposition 8.** *Let $\rho^{(0)}_{d,t} \in \mathcal{P}(\mathbb{R}^2)$ 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 $\mathfrak{t}(\rho^{(\ell)}_{d,t})$.*

[3] That is, for a measurable $\mathfrak{A} \subseteq (0,1)^2$ we have $\mathfrak{t}(\rho)(\mathfrak{A}) = \rho(\mathfrak{t}^{-1}(\mathfrak{A}))$.

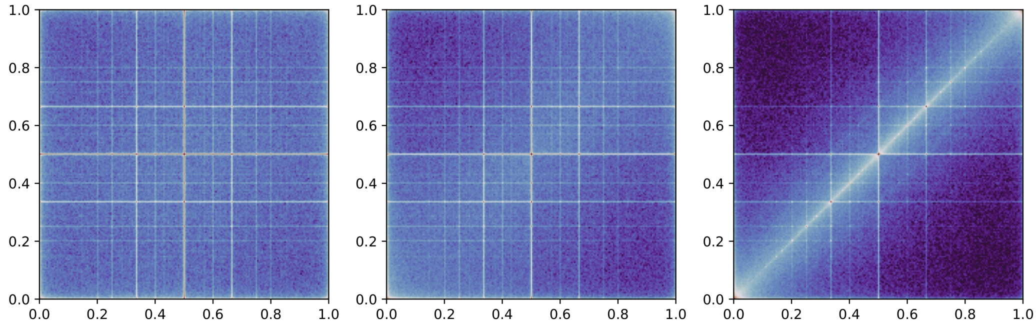

Figure 2 The distributions $\mathfrak{t}(\rho_{d,t})$ for $d = 1.9$ and $t = 0.1, 0.5, 0.9$.

We employ the contraction method to show that the sequence $(\rho^{(\ell)}_{d,t})_{\ell \geq 1}$ of measures converges.

▶ **Proposition 9.** *There exists a unique $\rho_{d,t} \in \mathcal{P}(\mathbb{R}^2)$ that satisfies (5) and $\lim_{\ell \to \infty} \rho^{(\ell)}_{d,t} = \rho_{d,t}$ weakly.*

Furthermore, the Gibbs uniqueness property (13) extends to $\boldsymbol{T}_1$ and $\boldsymbol{T}_2$.

▶ **Corollary 10.** *For all $t \in [0,1]$ and $h = 1, 2$ we have*

$$\mathbb{E}\left[\max_{\tau \in S(\boldsymbol{T}^{(2\ell)}_h)} \left|\mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}^{(2\ell)}_h, o} = 1 \mid \boldsymbol{T}^{\otimes}, \boldsymbol{\sigma}_{\boldsymbol{T}^{(2\ell)}_h, \partial^{2\ell} o} = \tau_{\partial^{2\ell} o}\right] - \mathbb{P}\left[\boldsymbol{\sigma}_{\boldsymbol{T}^{(2\ell)}_h, o} = 1 \mid \boldsymbol{T}^{\otimes}\right]\right|\right] \to 0, \tag{19}$$

as $\ell \to +\infty$.

Combining Propositions 8 and 9 and Corollary 10, 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

$$\pi_{\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M')} = \frac{1}{n}\sum_{i=1}^{n} \delta_{(\mathbb{P}[\boldsymbol{\sigma}_{\hat{\boldsymbol{\Phi}}_1(M,M'), x_i} = 1 | \hat{\boldsymbol{\Phi}}_1(M,M')], \mathbb{P}[\boldsymbol{\sigma}_{\hat{\boldsymbol{\Phi}}_2(M,M'), x_i} = 1 | \hat{\boldsymbol{\Phi}}_2(M,M')])} \in \mathcal{P}([0,1]^2)$$

be the empiricial distribution of the joint marginals of $\hat{\boldsymbol{\Phi}}_1(M, M')$ and $\hat{\boldsymbol{\Phi}}_2(M, M')$, which we need to know to evaluate the expressions from Proposition 6. Furthermore, denote by $W_1(\,\cdot\,,\,\cdot\,)$ the Wasserstein L^1-distance of two probability measures on $[0,1]^2$.

▶ **Corollary 11.** *For any $t \in [0,1]$ and any $M \sim tnd/2$, $M' \sim (1-t)dn/2$ we have*

$$\mathbb{E}\left[W_1\left(\pi_{\hat{\boldsymbol{\Phi}}_1(M,M'), \hat{\boldsymbol{\Phi}}_2(M,M')}, \mathfrak{t}(\rho_{d,t})\right)\right] = o(1).$$

Finally, combining Proposition 6 with Corollary 11, we obtain the variance of $\log Z(\hat{\boldsymbol{\Phi}})$.

▶ **Corollary 12.** *With $\eta(d)^2$ from (7) we have $\eta(d) > 0$ and $\mathrm{Var} \log Z(\hat{\boldsymbol{\Phi}}) \sim m\eta_d^2$.*

Because the proof of Proposition 9 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. 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 $(\mathfrak{F}_{n,M})_{0\leq M\leq m_n}$ 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

$$\boldsymbol{Z}_{n,M} = m^{-1/2}\mathbb{E}\left[\log Z(\hat{\boldsymbol{\Phi}}) \mid \mathfrak{F}_{n,M}\right] \tag{20}$$

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

▶ **Proposition 13.** *For all $0 < d < 2$ the martingale (20) satisfies*

$$\lim_{n\to\infty}\mathbb{E}\left[\max_{1\leq M\leq m}|\boldsymbol{X}_{n,M}|\right] = 0 \qquad \textit{and} \qquad \lim_{n\to\infty}\mathbb{E}\left|\eta(d)^2 - \sum_{M=1}^{m}\boldsymbol{X}_{n,M}^2\right| = 0. \tag{21}$$

Thanks to pruning, the first condition from (21) is easily checked. Furthermore, the steps that we pursued towards the proof of Corollary 12, i.e., the variance calculation, also imply the second condition without further ado. Finally, as (21) demonstrates that the marginal differences are small and that the variance process converges to a deterministic limit, Theorem 1 follows the general martingale central limit theorem from [27].

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 [19, 31]. Subsequent successes include the 1-in-k-SAT threshold [3] and the k-XORSAT threshold [26, 40]. Furthermore, Friedgut [29] 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 $k \geq 3$ [35]. A line of rigorous work [6, 8, 22] culminated in the verification of this physics prediction for large k [24].

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 [14] investigated the scaling window around the satisfiability threshold, a point on which a recent contribution by Dovgal, de Panafieu and Ravelomanana elaborates [25]. Abbe and Montanari [2] made the first substantial step towards the study of the number of satisfying assignments that $\frac{1}{n}\log Z(\boldsymbol{\Phi})$ 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 [37] obtain a "law-of-large-numbers" estimate of the number of assignments that violate all but $o(n)$ clauses for $d < 1.16$. Finally, the aforementioned article of Achlioptas et al. [5] verifies the prediction from [36] 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 [39]. This bound matches the predictions of the cavity method [34]. 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 [37, 38, 44].

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].[4] 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, 26, 40]. 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, 20, 41]. The same is true of random k-SAT instances with regular literal degrees [23]. 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 [43] 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, 23, 20, 41], 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 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 [18] 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 [18]. 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 [32] and Glasgow, Kwan, Sah, Sawhney [30] 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.

References

1. E. Abbe, S. Li, and A. Sly. Proof of the contiguity conjecture and lognormal limit for the symmetric perceptron. In *Proc. 62nd FOCS*, pages 327–338, 2022.
2. E. Abbe and A. Montanari. On the concentration of the number of solutions of random satisfiability formulas. *RSA*, 45:362–382, 2014.

[4] Formally, by "symmetry" we mean that the empirical distribution of the marginals of random solutions converges to an atom; cf. [21].

3 D. Achlioptas, A. Chtcherba, G. Istrate, and C. Moore. The phase transition in 1-in-k sat and nae 3-sat. In *Proc. 12th SODA*, pages 721–722, 2001.
4 D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th FOCS*, pages 793–802, 2008.
5 D. Achlioptas, A. Coja-Oghlan, M. Hahn-Klimroth, J. Lee, N. Müller, M. Penschuck, and G. Zhou. The number of satisfying assignments of random 2-sat formulas. *Random Structures and Algorithms*, 58:609–647, 2021.
6 D. Achlioptas and C. Moore. Random k-sat: two moments suffice to cross a sharp threshold. *SIAM Journal on Computing*, 36:740–762, 2006.
7 D. Achlioptas, A. Naor, and Y. Peres. Rigorous location of phase transitions in hard optimization problems. *Nature*, 435:759–764, 2005.
8 D. Achlioptas and Y. Peres. The threshold for random k-sat is $2^k \ln 2 - O(k)$. *Journal of the AMS*, 17:947–973, 2004.
9 D. Aldous and J. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In H. Kesten, editor, *Probability on Discrete Structures*. Springer, 2004.
10 B. Aspvall, M. F. Plass, and R. E. Tarjan. A linear-time algorithm for testing the truth of certain quantified boolean formulas. *Information Processing Letters*, 8:121–123, 1979.
11 P. Ayre, A. Coja-Oghlan, P. Gao, and N. Müller. The satisfiability threshold for random linear equations. *Combinatorica*, 40:179–235, 2020.
12 V. Bapst, A. Coja-Oghlan, and C. Efthymiou. Planting colourings silently. *Combinatorics, Probability and Computing*, 26:338–366, 2017.
13 V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, and D. Vilenchik. The condensation phase transition in random graph coloring. *Communications in Mathematical Physics*, 341:543–606, 2016.
14 B. Bollobás, C. Borgs, J. Chayes, J. Kim, and D. Wilson. The scaling window of the 2-sat transition. *RSA*, 18:201–256, 2001.
15 G. Bresler and B. Huang. The algorithmic phase transition of random k-sat for low degree polynomials. In *Proc. 62nd FOCS*, pages 298–309, 2021.
16 S. Cao. Central limit theorems for combinatorial optimization problems on sparse erdős-rényi graphs. *Annals of Applied Probability*, 31:1687–1723, 2021.
17 P. Cheeseman, B. Kanefsky, and W. Taylor. Where the *really* hard problems are. In *Proc. IJCAI*, pages 331–337, 1991.
18 W.-K. Chen, P. Dey, and D. Panchenko. Fluctuations of the free energy in the mixed p-spin models with external field. *Probability Theory and Related Fields*, 168:41–53, 2017.
19 V. Chvátal and B. Reed. Mick gets some (the odds are on his side). In *Proc. 33th FOCS*, pages 620–627, 1992.
20 A. Coja-Oghlan, T. Kapetanopoulos, and N. Müller. The replica symmetric phase of random constraint satisfaction problems. *Combinatorics, Probability and Computing*, 29:346–422, 2020.
21 A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborová. Information-theoretic thresholds from the cavity method. *Advances in Mathematics*, 333:694–795, 2018.
22 A. Coja-Oghlan and K. Panagiotou. The asymptotic k-sat threshold. *Advances in Mathematics*, 288:985–1068, 2016.
23 A. Coja-Oghlan and N. Wormald. The number of satisfying assignments of random regular k-sat formulas. *Combinatorics, Probability and Computing*, 27:496–530, 2018.
24 J. Ding, A. Sly, and N. Sun. Proof of the satisfiability conjecture for large k. *Annals of Mathematics*, 196:1–388, 2022.
25 S. Dovgal, É. de Panafieu, and V. Ravelomanana. Exact enumeration of satisfiable 2-sat formulae. *arXiv:2108.08067*, 2021. arXiv:2108.08067.
26 O. Dubois and J. Mandler. The 3-xorsat threshold. In *Proc. 43rd FOCS*, pages 769–778, 2002.
27 G. Eagleson. Martingale convergence to mixtures of infinitely divisible laws. *Annals of Probability*, 3:557–562, 1975.

28 C. Efthymiou. On sampling symmetric gibbs distributions on sparse random graphs and hypergraphs. In *Proc. 49th ICALP*, page 57, 2022.

29 E. Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the AMS*, 12:1017–1054, 1999.

30 M. Glasgow, M. Kwan, A. Sah, and M. Sawhney. A central limit theorem for the matching number of a sparse random graph. *arXiv:2402.05851*, 2024.

31 A. Goerdt. A threshold for unsatisfiability. *Journal of Computer and System Sciences*, 53:469–486, 1996.

32 E. Kreačič. *Some problems related to the Karp-Sipser algorithm on random graphs*. PhD thesis, University of Oxford, 2017.

33 F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proceedings of the National Academy of Sciences*, 104:10318–10323, 2007.

34 M. Mézard and A. Montanari. *Information, physics and computation*. Oxford University Press, 2009.

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

36 R. Monasson and R. Zecchina. The entropy of the k-satisfiability problem. *Physical Review Letters*, 76:3881, 1996.

37 A. Montanari and D. Shah. Counting good truth assignments of random k-sat formulae. In *Proc. 18th SODA*, pages 1255–1264, 2007.

38 D. Panchenko. On the replica symmetric solution of the K-sat model. *Electronic Journal of Probability*, 19:67, 2014.

39 D. Panchenko and M. Talagrand. Bounds for diluted mean-fields spin glass models. *Probability Theory and Related Fields*, 130:319–336, 2004.

40 B. Pittel and G. Sorkin. The satisfiability threshold for k-xorsat. *Combinatorics, Probability and Computing*, 25:236–268, 2016.

41 F. Rassmann. On the number of solutions in random graph k-colouring. *Combinatorics, Probability and Computing*, 28:130–158, 2019.

42 R. Robinson and N. Wormald. Almost all regular graphs are hamiltonian. *Random Structures and Algorithms*, 5:363–374, 1994.

43 A. Sly, N. Sun, and Y. Zhang. The number of solutions for random regular nae-sat. *Probability Theory and Related Fields*, 182:1–109, 2022.

44 M. Talagrand. The high temperature case for the random K-sat problem. *Probability Theory and Related Fields*, 119:187–212, 2001.

45 L. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8:410–421, 1979.