Counting good truth assignments of random k-SAT formulae

Andrea Montanari* and Devavrat Shah†

July 17, 2006

Abstract

We present a deterministic approximation algorithm to compute *logarithm* of the number of 'good' truth assignments for a random k-satisfiability (k-SAT) formula in polynomial time (by 'good' we mean that violate a small fraction of clauses). The relative error is bounded above by an arbitrarily small constant ϵ with high probability[1] as long as the clause density (ratio of clauses to variables) $\alpha < \alpha_{\rm u}(k) = 2k^{-1}\log k(1+o(1))$. The algorithm is based on computation of marginal distribution via belief propagation and use of an interpolation procedure. This scheme substitutes the traditional one based on approximation of marginal probabilities via MCMC, in conjunction with self-reduction, which is not easy to extend to the present problem.

We derive $2k^{-1}\log k(1+o(1))$ as threshold for uniqueness of the Gibbs distribution on satisfying assignment of random infinite tree k-SAT formulae to establish our results, which is of interest in its own right.

1 Introduction

Setup and Problem Statement. Given N boolean variables $x_i, 1 \le i \le N$, an M clause k-satisfiability (k-SAT) formula has the form $F = \wedge_{j=1}^{M} C_j$, where $C_j = \vee_{\ell=1}^{k} z_{j_\ell}$ with literal z_{j_ℓ} being either x_i for $\bar{x}_i$ for some $1 \le i \le N$. An assignment $\underline{x} \in \{0,1\}^N$ of variables $x_i, 1 \le i \le N$ satisfies clauses C_j if at least of one the k literals of C_j evaluates to be true. We will denote true by "1" and false by "0". For given F, $E(\underline{x})$ denote the number of unsatisfied clauses of F under assignment $\underline{x}$. Given $\beta \in \mathbb{R}_+$ (called *inverse temperature* in statistical physics), define *partition function* as

$$Z_N(\beta, F) \equiv \sum_{\underline{x}\in\{0,1\}^N} e^{-\beta E(\underline{x})}\,. \tag{1}$$

Notice that $Z_N(\beta, F)$ weighs in favor of "good" assignments, i.e. assignments that satisfy more clauses. As $\beta \to \infty$, $Z_N(\beta, F)$ becomes the number of assignments that satisfy (all clauses of) F. The partition function naturally arises as normalizing constant in the following probability measure on $\{0,1\}^N$, often denoted as *Boltzmann* distribution [1] related to F: for $\underline{x} \in \{0,1\}^N$,

$$\mu_{\beta,F}(\underline{x}) = \frac{1}{Z_N(\beta,F)}\prod_{j=1}^{M}\psi_j(\underline{x}) \;=\; \frac{e^{-\beta E(\underline{x})}}{Z_N(\beta,F)}\,, \quad \text{where} \quad \psi_j(\underline{x}) = \begin{cases} 1 & \text{if } \underline{x} \text{ satisfy clause } C_j,\\ e^{-\beta} & \text{otherwise.}\end{cases} \tag{2}$$

*Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, Paris. Research is partially supported by European Union under the ip EVERGROW. Email: montanar@lpt.ens.fr

†LIDS, MIT. Research is partially supported by NSF CAREER. Email: devavrat@mit.edu.
Keywords: Random k-SAT, Correlation Decay, Uniqueness, Gibbs Distribution

[1]In this paper, by term "with high probability" (whp) we mean with probability $1 - o_N(1)$.

We shall write $\mu(\,\cdot\,) = \mu_{\beta,F}(\,\cdot\,)$ whenever it will not be necessary to specify the formula and inverse temperature. We further denote by $\langle\,\cdot\,\rangle = \langle\,\cdot\,\rangle_{\beta,F}$ expectations with respect to the measure μ.

In this paper, we are interested in *random k-SAT* formulas. These are generated by selecting M clauses independently and uniformly at random from all possible $2^k\binom{N}{k}$ k-clauses. Specifically, let M scale linearly in N, i.e. $M = \alpha N$ for $\alpha \in \mathbb{R}_+$.

The main motivation in this paper is to describe an efficient algorithm to compute a good approximation of $Z_N(\beta, F)$ for such random formulas. An important open conjecture is to show, that for any $\alpha, \beta \in \mathbb{R}_+$, under the probability distribution induced by random k-SAT formula, the limit $\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta, F)$ exists with probability 1. The analysis of our algorithm implies such a result for all finite β, and α smaller than a critical value.

Related Previous Work. The well-known threshold conjecture for random k-SAT states that for all $k \geq 2$, there exists $\alpha_{\rm c}(k)$ such that for $\alpha < \alpha_{\rm c}(k)$ (resp. $\alpha > \alpha_{\rm c}(k)$) the randomly generated formula is satisfiable (resp. not satisfiable) with probability 1 as $N \to \infty$. There has been a lot of interesting work on this topic, and a convergence of methods from different communities [2, 3, 4]. Due to space limitation, we will recall only some of the key relevant results.

Friedgut [14] established existence of a sharp threshold. More precisely, he proved that there exists $\alpha_{\rm c}(k, N)$ such that the satisfiability probability tends to 1 (to 0) if $\alpha < \alpha_{\rm c}(k, N)(1 - \eta)$ (respectively $\alpha > \alpha_{\rm c}(k, N)(1 + \eta)$). While it is expected that $\lim_{N\to\infty} \alpha_c(k, N)$ exists, it has still remained elusive. Recently, Achlioptas and Peres [6] established that $\alpha_{\rm c}(k, N) = 2^k \ln k(1 + o_k(1))$ thus implying that $\alpha_{\rm c}(k, N)$ can be taken N independent to first order for large k.

The existence of $\lim_{N\to\infty} \lim_{\beta\to\infty} \frac{1}{N} \log Z_N(\beta, F)$ with probability 1, for all $\alpha \in \mathbb{R}_+$ and k naturally establishes the threshold conjecture. More generally, the log-partition function at $\beta = \infty$ provides detailed information about the satisfying assignments (computing it exactly is of course #-P complete). In [7] a formula for the limit log-partition function was derived through the non-rigorous replica method from statistical physics. The existence of the $N \to \infty$ limit was proved by Franz, Leone and Toninelli [8, 9] for even k and all values of α. These authors also provided an upper bound on $\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta, F)$. However evaluating the bound requires solving an *a priori* complex optimization problem, and a matching lower bound wasn't proved there. Talagrand [5] established the existence of the limit and its value for very small value of β (depending on k).

Overview of Results. In this paper, we essentially prove that the Boltzmann distribution (2) is a *pure state* [1] by establishing appropriate worst-case *correlation decay* for tree formulae. The approach of Talagrand [5] also crucially relied of proving correlation decay, albeit with different means. This resulted in a limitation to small values of β and thus leaving out interesting regime of large β.

An analogy can be drawn with the Markov Chain Monte Carlo (MCMC) approach to the approximate computation of partition functions (see, for example, work by Jerrum and Sinclair [10]). In that case, the crucial step consists in proving an appropriate mixing condition ('temporal' correlation decay) for some Markov Chain. The same role is played here by 'spatial' correlation decay with respect to the measure (2).

In this paper, we establish correlation decay for random k-SAT formula for a range of α and all β. This allows to estabilish that deterministic Belief Propagation algorithm provides a good approximation of the marginals with respect to the distribution (2), cf. Section 3. In the usual MCMC approach, marginals are used to approximate the partition function by recursively fixing the variables and exploiting self-reducibility. This cannnot be done in the present case because the reduced SAT formulae are not random anymore. Instead, we use *interpolation* in β, to obtain $\log Z_N(\beta, F)$ approximately (Theorem 1). The analysis of the approximation scheme implies the existence of the limit $\lim_{N\to\infty} \frac{1}{N} \log Z_N(\beta, F)$ (Theorem 3). We hope that our novel approach for counting will find applications in other hard combinatorial problems. Similar schemes were recently discussed by Weitz [11], and Bandyopadhyay and Gamarnik [12] for counting independent sets approximately via deterministic algorithms.

Finally, we show that the computation of the partition function leads to an estimate of the number of truth assignments that violate at most $N\varepsilon$ clauses, for small ε (Theorem 4). As a byproduct, we obtain an asymptotically (in k) threshold for uniqueness Gibbs measure on infinite k-SAT tree formula (Theorem 2).

Organization. Section 2 presents preliminaries and statements of the main results. The Section 3 describes the approximate counting algorithm and the proof of key Lemmas related to the correlation decay (or uniqueness) of Gibbs distribution on random *tree* k-SAT. The Section 4 completes the proofs of all main results stated in Section 2. We present direction for future work in Section 5.

2 Preliminaries and Main Results

Given α and k, define $\alpha_*(k)$ to be the smallest positive root of the equation $\kappa(\alpha) = 1$, where

$$\kappa(\alpha) \equiv k(k-1)\alpha \left(1 - \frac{1}{4}\, e^{-k\alpha/2}\right) \left(1 - \frac{1}{2}\, e^{-k\alpha/2}\right)^{k-2} . \tag{3}$$

For $k = 2, 3, 4, 6$, the $\alpha_*(k)$ is approximately $0.58216, 0.293, 0.217, 0.16670$. Asymptotically, $\alpha_*(k) = 2k^{-1}\log k \left(1 + O\left(\frac{\log\log k}{\log k}\right)\right)$. Now, we state the main result of this paper about approximating logarithm of partition function.

Theorem 1. *Given $\varepsilon > 0$ and $\alpha < \alpha_*(k)$, there exists $\delta' > 0$ and a polynomial (in N, independent of ε) time algorithm that computes a number $\Phi(\beta, F)$ (the input being $\beta \in \mathbb{R}$ and a satisfiability formula F) such that the following is true. If $\beta \in [0, N^{\delta'}]$ and F is random k-SAT formula with N variables and $M = N\alpha$ clauses, then, with high probability,*

$$\Phi(\beta, F)\,(1-\varepsilon) \le \log Z_N(\beta, F) \le \Phi(\beta, F)\,(1+\varepsilon). \tag{4}$$

The proof of Theorem 1 requires us to prove uniqueness of Gibbs measure for the model (2) on infinite tree random k-SAT formulae. To state this result, we first need some definitions. An appropriate model for tree random k-SAT, $\mathsf{T}_*(r)$ is described as follows: For $r = 0$, it is the graph containing a unique variable node. For any $r \ge 1$, start by a single variable node (the root) and add $l \stackrel{\mathrm{d}}{=} \mathsf{Poisson}(k\alpha)$ clauses, each one including the root, and $k-1$ new variables (first generation variables). For each one of the l clauses, the corresponding literals are non-negated or negated indipendently with equal probability. If $r \ge 2$, generate an independent copy of $\mathsf{T}_*(r-1)$ for each variable node in the first generation and attach it to them. By construction, for any $r' < r$ the first r' generations of a tree from $\mathsf{T}_*(r)$ are distributed according to the model $\mathsf{T}_*(r')$. As a consequence, the infinite tree distribution $\mathsf{T}_*(\infty)$ is also well defined. In what follows, we denote the root of $\mathsf{T}_*(\cdot)$ as 0. Let μ denote the Gibbs distribution on random formula on $\mathsf{T}_*(r)$ (cf. (2)) and $\mu_{0|r}(x_0|\underline{x}_r)$ be the conditional distribution of root variable conditional to the assignment of r-th generation nodes of $\mathsf{T}_*(r)$ according to $\underline{x}_r$. The key property for most of the results of this paper is that of correlation decay with respect to random tree formulas $\mathsf{T}_*(\cdot)$.

Definition 1. *Given $\alpha, \beta \in \mathbb{R}_+$ and $k \ge 2$, the Gibbs distribution defined by (2) on the random tree $\mathsf{T}_*(\cdot)$ is unique with exponential correlation decay if there exists positive constants $A, \gamma > 0$, such that*

$$\mathbb{E}\left[\sup_{\underline{x}_r, \underline{z}_r} \left|\left|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\right|\right|_{\mathrm{TV}}\right] \le A\, e^{-\gamma r}\,, \tag{5}$$

for any $r \ge 0$. The uniqueness threshold $\alpha_{\mathrm{u}}(k)$ is the supremum value of α such that the above condition is verified for any $\beta \in [0, \infty]$.

The property defined here is a lot stronger than the usual notion of correlation decay, which only requires $\left|\left|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\right|\right|_{\mathrm{TV}} \to 0$ as $r \to \infty$ almost surely. Let $\alpha'_{\mathrm{u}}(k)$ denote the threshold for this weaker property. To the best of our knowledge, nothing has been known about the precise values of $\alpha_{\mathrm{u}}(k), \alpha'_{\mathrm{u}}(k)$ or the relation between them other than trivial lower bound from percolation threshold of $\Omega(k^{-2})$. We establish the precise asymptotic behavior of $\alpha_{\mathrm{u}}(k)$ and show that $\alpha_{\mathrm{u}}(k) = \alpha'_{\mathrm{u}}(k)(1 + o_k(1))$ as stated below.

Theorem 2. *For the Gibbs distribution (2) defined on $\mathsf{T}_*(\cdot)$ as above,*

$$\alpha_{\mathrm{u}}(k) = \frac{2\log k}{k}\left\{1 + O\left(\frac{\log\log k}{\log k}\right)\right\}, \qquad \alpha'_{\mathrm{u}}(k) = \frac{2\log k}{k}\left\{1 + O\left(\frac{\log\log k}{\log k}\right)\right\}. \tag{6}$$

Though algorithmically we obtain approximation of $\log Z_N(\beta, F)$, it is possible to establish the convergence of $\frac{1}{N}\log Z_N(\beta, F)$ with probability 1. Before stating this result, we need some definitions. In what follows, define function $f : \mathbb{R}^{k-1} \to \mathbb{R}$ as

$$f(x_1, \ldots, x_{k-1}) = -\frac{1}{2}\log\left\{1 - \frac{1-e^{-\beta}}{2^{k-1}}\prod_{i=1}^{k-1}(1-\tanh x_i)\right\}. \tag{7}$$

Let $\mathcal{D}$ denote the space of probability distributions on the real line $\mathbb{R}$. Define functions $S, S_1, S_2 : \mathcal{D} \to \mathcal{D}$ as follows: Given $\mu \in \mathcal{D}$, define random variable $u = f(h_1, \ldots, h_{k-1})$ where $h_1, \ldots, h_{k-1}$ are i.i.d. with distribution μ. Define distribution of u as $S_1(\mu)$. Given a distribution $\nu \in \mathcal{D}$, let random variable $h_0 = \sum_{a=1}^{\ell^+} u_a - \sum_{b=1}^{\ell^-} u_b$, where ℓ^+, ℓ^- are independent Poisson random variables with mean $k\alpha/2$ and u_a, u_b be i.i.d. with distribution ν. Let distribution of h_0 be denoted by $S_2(\nu)$. Define $S \equiv S_1 \circ S_2$. Now, we state the result.

Theorem 3. *Given k, let $\alpha < \alpha_*(k)$ and $\beta \in [0, \infty)$. Then, the function $S : \mathcal{D} \to \mathcal{D}$ as defined above has unique fixed point, say μ^*. Let $\nu^* = S_2(\mu^*)$. Then,*

$$\frac{1}{N}\log Z(\beta, F_N) \overset{\text{a.s.}}{\to} \phi(\beta), \tag{8}$$

$$\begin{aligned} \textit{where } \phi(\beta) \quad = \quad & -k\alpha\mathbb{E}\log[1+\tanh h \tanh u] + \alpha\mathbb{E}\log\left\{1 - \frac{1}{2^k}(1-e^{-\beta})\prod_{i=1}^{k}(1-\tanh h_i)\right\} + \\ & + \mathbb{E}\log\left\{\prod_{i=1}^{\ell_+}(1+\tanh u_i^+)\prod_{i=1}^{\ell_-}(1-\tanh u_i^-) + \prod_{i=1}^{\ell_+}(1-\tanh u_i^+)\prod_{i=1}^{\ell_-}(1+\tanh u_i^-)\right\}, \end{aligned} \tag{9}$$

where $u, u_i^{\pm}$ are i.i.d. with distribution μ^*, h, h_j are i.i.d. with distribution ν^* and $\ell_\pm$ are Poisson of mean $k\alpha/2$.

Finally, define $\Xi(\zeta, F)$ to be the number of assigments that violate at most ζ clauses. The next result formalizes the relation between the approximation of $Z_N(\beta, F)$ and counting the number of truth assignments that violate a small fraction of clauses.

Theorem 4. *For any $k \geq 2$, $\varepsilon > 0$, and $\alpha < \alpha_*(k)$ there exists $A, C > 0$, $a > 0$ such that the following is true. If F is a random k-SAT $M = N\alpha$ clauses over N variables, and $\beta = A\log 1/\varepsilon$, then*

$$|\log \Xi(N\varepsilon, F) - \Phi(\beta, F)| \leq NC\varepsilon^a, \tag{10}$$

with high probability, where $\Phi(\beta, F)$ as defined in Theorem 1.

3 Algorithm and Key Lemmas

3.1 Algorithm

We first define a factor graph G_F for a given formula F: each variable is represented by (circle) variable node and each clause by a (square) clause node with an edge between a variable and a clause node only if corresponding variable belongs to the clause. The edge is solid if variable is non-negated and dashed if variable is negated. The Belief Propagation (BP) algorithm is a heuristic (exact for tree factor graphs) to estimate the marginal distribution of node variables for any factor graph. Specifically, we will use BP to approximately compute marginals of the distribution (2).

We will quickly recall BP for our specific setup. We refer reader to see [15, 16] for further details on the algorithm. BP is a message passing algorithm in which at each iteration messages are sent from variable nodes to neighboring clause nodes and vice versa. The messages at iteration $t+1$ are functions of messages received at iteration t. To describe the message update equations, we need some notation. Let

∂a denote the set of all variables that belong to clause a. If variable x_i is involved in clause a as literal z (either $z = x_i$ or $z = \bar{x}_i$), then define $\partial_+ i(a)$ as the set of all clauses (minus a) in which x_i appears as z. Similarly, $\partial_- i(a)$ denotes the set of all clauses in which x_i appears as $\bar{z}$. Let $\{h^{(t)}_{i\to a}\}$, $\{u^{(t)}_{a\to i}\}$ denote the messages (ideally they are half log-likelihood ratios) that are passed along the ddirected edges $i \to a$ and $a \to i$ respectively at time t, then the precise update equations are

$$h^{(t+1)}_{i\to a} = \sum_{b\in\partial_+ i(a)} u^{(t)}_{b\to i} - \sum_{b\in\partial_- i(a)} u^{(t)}_{b\to i}\,, \qquad u^{(t)}_{a\to i} = f(\{h^{(t)}_{j\to a}; j \in \partial a\backslash i\})\,, \tag{11}$$

where the function $f(\,\cdot\,)$ has been defined in Eq. (7). We shall assume [2] that the update equations are initialized by $h^{(0)}_{i\to a} = 0$ and algorithm stops at iteration $t_{\max}$ which is equal to the diameter of G_F. Let $(h_{i\to a}, u_{a\to i})$ be messages passed in the last iteration of BP. Using these messages, an estimate of the probability that a clause is satisfied can be obtained as follows. Let $E_a(\underline{x}_{\partial a})$ be the indicator function for the a-th clause not being satisfied. As mentioned above, $h_{i\to a}$ is thought of as half log-likelihood ratio for i satisfying a and i not satisfying a, in the absence of clause a itself. A little algebra then shows that the BP estimate for the expectation of $E_a(\underline{x}_{\partial a})$ is

$$\langle E_a(\underline{x}_{\partial a})\rangle_{\rm BP} = \frac{\sum_{\underline{x}_{\partial a}} E_a(\underline{x}_{\partial a})\ \exp\{-\beta E_a(\underline{x}_{\partial a}) + h_{i\to a}\sigma_{ai}(x_i)\}}{\sum_{\underline{x}_{\partial a}} \exp\{-\beta E_a(\underline{x}_{\partial a}) + h_{i\to a}\sigma_{ai}(x_i)\}}\,, \tag{12}$$

where $\sigma_{ai}(x) = +1$ if setting $x_i = x$ satisfies clause a, and $= -1$ otherwise. We further introduce the number of clauses violated by $\underline{x}$, $E(\underline{x}) = \sum_a E_a(\underline{x}_{\partial a})$, and its BP estimate $\langle E(\underline{x})\rangle_{\rm BP} = \sum_a \langle E_a(\underline{x})\rangle_{\rm BP}$.

Given $\beta > 0$, we let $\beta_i = i\beta/N^2$, for $i = 0, \ldots, n \equiv N^2$. Then,

$$\log Z(\beta, F) \;\;=\;\; \log Z(0, F) + \sum_{i=0}^{n-1} \log\frac{Z(\beta_{i+1}, F)}{Z(\beta_i, F)} = N\log 2 + \sum_{i=0}^{n-1} \log\langle e^{-\Delta E(\underline{x})}\rangle_i\,, \tag{13}$$

where $\Delta \equiv \beta_{i+1} - \beta_i$, and $\langle\,\cdot\,\rangle_i$ is a shorthand for expectation with respect to the measure $\mu_{\beta_i,F}(\,\cdot\,)$. The above expression is difficult to evaluate. However, due to Δ being small the $\langle -\Delta E(\underline{x})\rangle$ is a good estimate of $\log\langle e^{-\Delta E(\underline{x})}\rangle_i$. Hence, define the algorithm estimate as

$$\Phi(\beta, F) = N\log 2 - \sum_{i=1}^{n-1} \Delta\,\langle E(\underline{x})\rangle_{{\rm BP},i}\,, \tag{14}$$

where the subscript in $\langle\,\cdot\,\rangle_{{\rm BP},i}$ emphasizes that the BP computation must be performed at inverse temperature β_i.

3.2 Key Lemmas

Before presenting useful Lemmas, let us mention a few facts. Given factor graph G_F and variable node i, $1 \le i \le N$, let $\mathsf{B}_i(r)$ denote subgraph induced by the set of all variable that are within shortest path distance r of node i (distance between two variables sharing a clause is unit). Analogously, for a clause node a, $\mathsf{B}_a(r)$ is the union of $\mathsf{B}_i(r)$ with i running over the variables involved in a. Let A be subset of variable nodes. Then, let $\underline{x}_A$ denote an assignment to the corresponding variables. Given two such subsets $A, B \subseteq [N]$ and assignments $\underline{x}_A$, $\underline{x}_B$, let $\mu_{A|B}(\underline{x}_A|\underline{x}_B)$ be the conditional probability under the distribution (2) of the variables in A, given assignment $\underline{x}_B$ on B. The following is a well-known result about BP algorithm (see [17]).

Lemma 1. *Given a clause a and r, let $\mathsf{B}_a(r+1)$ be a tree. Let $U = \mathsf{B}_a(r)$ and $V = [N]\backslash U$. Then*

$$|\langle E_a(\underline{x}_{\partial a})\rangle - \langle E_a(\underline{x}_{\partial a})\rangle_{\rm BP}| \le \sup_{\underline{y},\underline{z}} \Big|\Big|\mu_{\partial a|V}(\,\cdot\,|\underline{y}_V) - \mu_{\partial a|V}(\,\cdot\,|\underline{z}_V)\Big|\Big|_{\rm TV}\,, \tag{15}$$

$$0 \le \langle E_a(\underline{x}_{\partial a})\rangle,\ \langle E_a(\underline{x}_{\partial a})\rangle_{\rm BP} \le \max_{\underline{z}_V}\left\{\sum_{\underline{x}_{\partial a}} E_a(\underline{x})\mu_{\partial a|V}(\underline{x}_{\partial a}|\underline{z}_V)\right\} \tag{16}$$

[2] In fact an arbitrary initial condition and a smaller number of iterations wouldn't change our main results.

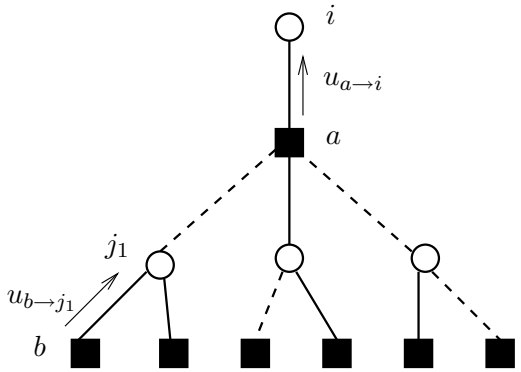

Figure 1: Pictorial representation of the recursion (21) on the factor graph G_F: filled squares represent function nodes and empty circles variable nodes. Dashed edges correspond to negations.

Next, we present a known result about locally tree-like structure of random k-SAT formula (an analogous result concerns the local structure of sparse random graphs).

Lemma 2. *Consider $k \geq 2$, $\alpha \in [0, \infty)$ and a random k-SAT formula F with clause density α. For $r \geq 0$, let $\mathsf{B}_i(r)$ be the ball of radius r centered at a uniformly random variable node i. Let $S(r)$ be an r-generation tree with distribution same as $\mathsf{T}_*(r)$ (with the same values of k and α). Then, there exists A, ρ (dependent on α, k) such that*

$$|| \, \mathbb{P}\{\mathsf{B}_i(r) \in \cdot \, \} - \mathbb{P}\{S(r) \in \cdot \, \} \, ||_{\mathrm{TV}} \leq \frac{A \, e^{\rho r}}{N} \, . \tag{17}$$

Lemma 3. *Let $\alpha_*(k)$ be the smallest positive root of the equation $\kappa(\alpha) = 1$, with $\kappa(\alpha)$ defined as in Eq. (3). Then $\alpha_*(k) \leq \alpha_{\mathrm{u}}(k)$.*

Proof: Given an r-generations tree formula F, consider an edge $i \to a$ directed toward the root and the subtree rooted at i and *not* containing a. Denote by $\mu_{i\to a}(\, \cdot \,)$ the marginal distribution of x_i with respect to the model associated to this subtree, and let $h_{i\to a} \in [-\infty, \infty]$ be the corresponding log-likelihood ratio

$$h_{i\to a} \equiv \frac{1}{2} \log \left\{ \frac{\mu_{i\to a}(x_i \text{ satisfies } a)}{\mu_{i\to a}(x_i \text{ doesn't satisfy } a)} \right\} . \tag{18}$$

Analogously, given an edge $a \to i$, we consider the subtree rooted at i and containing only a among the clauses involving i. We denote by $\mu_{a\to i}(\, \cdot \,)$ the corresponding marginal distribution at i, and let

$$u_{a\to i} \equiv \frac{1}{2} \log \left\{ \frac{\mu_{a\to i}(x_i \text{ satisfies } a)}{\mu_{a\to i}(x_i \text{ doesn't satisfy } a)} \right\} . \tag{19}$$

It is easy to show that these log-likelihoods satisfy the recursions[3]

$$h_{j\to a} = \sum_{b \in \partial_+ j(a)} u_{b\to j} - \sum_{b \in \partial_- j(a)} u_{b\to j} \, , \qquad u_{a\to i} = f(\{h_{j\to a}; j \in \partial a \backslash i\}) \, , \tag{20}$$

with the function $f(\, \cdot \,)$ being defined as in Eq. (7). For the calculations below, it is convernient to eliminate the $h_{i\to a}$ variables, to get

$$u_{a\to i} = f \left(\sum_{b \in \partial_+ j_1(a)} u_{b\to j_1} - \sum_{b \in \partial_- j_1(a)} u_{b\to j_1}; \; \cdots \; ; \sum_{b \in \partial_+ j_{k-1}(a)} u_{b\to j_{k-1}} - \sum_{b \in \partial_- j_{k-1}(a)} u_{b\to j_{k-1}} \right) , \tag{21}$$

[3]The reader will notice that these coincide with the BP update equations, cf. Eq. (11), which are known to be exact on trees.

where we denoted by $j_1, \dots, j_{k-1}$ the indices of variables involved in clause a (other than i). A pictorial representation of this recursion is provided in Fig. 1.

Notice that the above recursions hold irrespective whether one considers the unconditional measure $\mu(\,\cdot\,)$, or the conditional one $\mu(\,\cdot\,|\underline{x}_r)$. What changes in the two cases are the initial condition for the recursion, i.e. the value of $h_{i\to a}$ associated with the variables i at the r-th generation. For the unconditioned measure ('free boudary condition'), the appropriate initialization is $h_{i\to a} = 0$. If one conditions to $\underline{x}_r$, $h_{i\to a} = +\infty$, or $= -\infty$ depending (respectively) whether x_i satisfy clause a or not.

In the rest of the proof, we shall think always to the conditioned measure $\mu(\,\cdot\,|\underline{x}_r)$. As a consequence, the log-likelihoods are, implicity, functions of $\underline{x}_r$: $u_{a\to i} = u_{a\to i}(\underline{x}_r)$ (indeed of the restriction of $\underline{x}_r$ to the subtree rooted at i, and only containing a). We then let

$$\overline{u}_{a\to i} = \max_{\underline{x}_r} u_{a\to i}(\underline{x}_r)\,, \qquad \underline{u}_{a\to i} = \min_{\underline{x}_r} u_{a\to i}(\underline{x}_r)\,. \tag{22}$$

In the case $\beta = \infty$, the maximum (minimum) is taken over all boundary conditions $\underline{x}_r$, such that the sub-formula rooted at i admits at least one solutions, under the condition $\underline{x}_r$ (there is always at least one such boundaries). We further let $\Delta_{a\to i} = \overline{u}_{a\to i} - \underline{u}_{a\to i} \geq 0$.

Consider a random tree distributed as $\mathsf{T}_*(r)$, conditioned to the root having degree 1, i.e. to the root variable being involved in a unique clause, to be denoted by a. Let $\Delta^{(r)} = \Delta_{a\to i}$ be the corresponding log-likelihoods interval. We will show that $\mathbb{E}\tanh\Delta^{(r)} \leq e^{-\gamma r}$ for some positive constant γ. Before proving this claim, let us show that it indeed implies the thesis. Denoting by $\partial_+ 0$ the set of clauses in which the root is involved as the direct literal, and by $\partial_- 0$ the set in which it is involved as negated, we have

$$\big|\big|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\big|\big|_{\rm TV} = \frac{1}{2}\big|\tanh h_0(\underline{x}_r) - \tanh h_0(\underline{z}_r)\big|\,, \tag{23}$$

$$h_0(\underline{x}_r) \equiv \sum_{a\in\partial_+ 0} u_{a\to 0}(\underline{x}_r) - \sum_{a\in\partial_- 0} u_{a\to 0}(\underline{x}_r)\,. \tag{24}$$

Since $x \mapsto \tanh(x)$ is monotonically increasing in x, we have

$$\big|\big|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\big|\big|_{\rm TV} \leq \frac{1}{2}\big\{\tanh\overline{h}_0 - \tanh\underline{h}_0\big\}\,, \tag{25}$$

$$\overline{h}_0 \equiv \sum_{a\in\partial_+ 0} \overline{u}_{a\to 0} - \sum_{a\in\partial_- 0} \underline{u}_{a\to 0}\,, \qquad \underline{h}_0 \equiv \sum_{a\in\partial_+ 0} \underline{u}_{a\to 0} - \sum_{a\in\partial_- 0} \overline{u}_{a\to 0}\,. \tag{26}$$

Using the elementary properties $\tanh x - \tanh y \leq 2\tanh(x-y)$ for any $x \geq y$, and $\tanh(x+y) \leq \tanh x + \tanh y$ for $x, y \geq 0$, we get

$$\big|\big|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\big|\big|_{\rm TV} \leq \tanh\left\{\sum_{a\in\partial 0}\Delta_{a\to 0}\right\} \leq \sum_{a\in\partial 0}\tanh\Delta_{a\to 0}\,. \tag{27}$$

We can take the maximum over boundary condition and the expectation with respect to the tree ensemble. Recalling that $|\partial 0|$ is a Poisson random variable of mean $k\alpha$, we get

$$\mathbb{E}\max_{\underline{x},\underline{z}}\big|\big|\mu_{0|r}(\,\cdot\,|\underline{x}_r) - \mu_{0|r}(\,\cdot\,|\underline{z}_r)\big|\big|_{\rm TV} \leq k\alpha\,\mathbb{E}\tanh\Delta^{(r)}\,, \tag{28}$$

which implies the thesis upon taking $A = k\alpha$.

We are now left with the task of proving $\mathbb{E}\tanh\Delta^{(r)} \leq e^{-\gamma r}$. It is easy to realize that $f(x_1, \dots, x_{k-1})$ is monotonically decreasing in each of its arguments. Therefore Eq. (21) yields the following recursion for upper/lower bounds

$$\overline{u}_{a\to i} = f\left(\sum_{b\in\partial_+ j_1(a)}\underline{u}_{b\to j_1} - \sum_{b\in\partial_- j_1(a)}\overline{u}_{b\to j_1};\ \dots\ ;\ \sum_{b\in\partial_+ j_{k-1}(a)}\underline{u}_{b\to j_{k-1}} - \sum_{b\in\partial_- j_{k-1}(a)}\overline{u}_{b\to j_{k-1}}\right), \tag{29}$$

together with the equation obtained by interchanging $\underline{u}_{\cdot\cdot}$ and $\overline{u}_{\cdot\cdot}$. By taking the difference of these two equations, we get

$$\Delta_{a\to i} = f(\underline{h}_1;\ \dots\ ;\underline{h}_{k-1}) - f(\overline{h}_1;\ \dots\ ;\overline{h}_{k-1})\,, \tag{30}$$

where we defined $\underline{h}_i = \sum_{b\in\partial_+ j_i(a)} \underline{u}_{b\to j_i} - \sum_{b\in\partial_- j_i(a)} \overline{u}_{b\to j_i}$ and $\overline{h}_i = \sum_{b\in\partial_+ j_i(a)} \overline{u}_{b\to j_i} - \sum_{b\in\partial_- j_i(a)} \underline{u}_{b\to j_i}$ (obviously $\underline{h}_i \geq \overline{h}_i$).

Suppose now n out of the $k-1$ variables $x_{j_1},\dots,x_{j_{k-1}}$ are pure literals, let's say variables $x_{j_1},\dots,x_{j_n}$. This means that $\partial_- j_1(a),\dots\partial_- j_n(a) = \emptyset$, and therefore, since the loglikelihoods $u_{b\to j}$ are non-negative (because f is non-negative), $\underline{h}_1,\dots\underline{h}_n \geq 0$. It is an easy exercise of analysis to show that, if $x_1,\dots,x_n \geq 0$,

$$0 \leq -\frac{\partial f}{\partial x_i}(x_1,\dots,x_{k-1}) \leq \frac{1}{2^n}\,. \tag{31}$$

Therefore, by the Mean Value Theorem

$$\Delta_{a\to i} \leq \frac{1}{2^n}\sum_{l=1}^{k-1}(\overline{h}_l - \underline{h}_l) = \frac{1}{2^n}\sum_{l=1}^{k-1}\sum_{b\in\partial j_l} \Delta_{b\to j_l}\,, \tag{32}$$

Next we take the hyperbolic tangent of both sides, and use again $\tanh(x+y) \leq \tanh x + \tanh y$, for $x, y \geq 0$ to get

$$\tanh \Delta_{a\to i} \leq \frac{1}{2^n}\sum_{l=1}^{k-1}\sum_{b\in\partial j_l} \tanh \Delta_{b\to j_l}\,. \tag{33}$$

Finally we take expectation of this inequality. In order to do this, we recall that n is just the number of pure literals among $x_{j_1},\dots x_{j_{k-1}}$. In our notations this can be written as $n = \sum_{l=1}^{k-1} \mathbb{I}(|\partial_- j_l(a)| = 0)$. We further assume that i is the root of a tree from $\mathsf{T}_*(r+1)$, $r \geq 0$ and therefore $\Delta_{a\to i}$ is distributed as $\Delta^{(r)}$. Furthermore the differences $\Delta_{b\to j_l}$ will be distributed as $\Delta^{(r+1)}$. We thus obtain

$$\mathbb{E}\tanh\Delta^{(r+1)} \leq \mathbb{E}\left\{\prod_{l=1}^{k-1}\frac{1}{2^{\mathbb{I}(|\partial_- j_l(a)|=0)}}\sum_{l=1}^{k-1}\sum_{b\in\partial j_l(a)}\tanh\Delta^{(r)}\right\} = \tag{34}$$

$$= (k-1)\,\mathbb{E}\left\{\frac{1}{2^{\mathbb{I}(|\partial_- j|=0)}}|\partial j|\right\}\left\{\mathbb{E}\,2^{-\mathbb{I}(|\partial_- j|=0)}\right\}^{k-2}\mathbb{E}\tanh\Delta^{(r)}\,. \tag{35}$$

The expectations over $|\partial_+ j|$, $|\partial_- j|$ are easily evaluated by recalling that these are inpependent Poisson random variables of mean $k\alpha/2$. One finally obtains $\mathbb{E}\tanh\Delta^{(r+1)} \leq \kappa(\alpha)\mathbb{E}\tanh\Delta^{(r)}$. The thesis follows (with $\gamma = -\log\kappa(\alpha)$) by noticing that $\mathbb{E}\tanh\Delta^{(0)} \leq 1$, and recalling that $\kappa(\alpha) < 1$ for $\alpha < \alpha_*(k)$. □

Next, we state result about the error in expectation w.r.t. to BP estimate in a clauses being satisfied or not. To obtain bound in the error of BP estimate of $\langle E_a(\underline{x})\rangle$, we need to study the error in estimation of the joint distribution of k variables in a clause. For this, we first choose a clause at random and treat all of its k variables as root of k independent rooted random trees (of suitable depth r) as before. Note that, this asymptotically does not bias the distribution of the original random formula as this process tilt the original distribution by at most $O(1/N)$.

To this end, let $\underline{x}_r$ be an assigment for the r-th generation variables. We shall denote by $\langle\,\cdot\,\rangle^{(r)}$ the expectation with respect to the graphical model (2) associated to a formula constructed as follows. First we generate a uniformly random clause over variables $x_1,\dots,x_k$. Then we sample k independent trees according to $\mathsf{T}_*(r)$ and root them at $x_1,\dots,x_k$. We let $\langle\,\cdot\,\rangle^{(r)}_{\underline{x}_r}$ be the corresponding conditional expectation, given the assignment to the r-th generation.

Lemma 4. *Let $k \geq 2$, $\alpha < \alpha_*(k)$ and $\beta \in [0,\infty]$. Then there exist two positive constants A, γ, such that*

$$\mathbb{E}\max_{\underline{x}_r,\underline{z}_r}\left|\langle E_a(\underline{x})\rangle^{(r)}_{\underline{x}_r} - \langle E_a(\underline{x})\rangle^{(r)}_{\underline{z}_r}\right| \leq A\,e^{-\gamma r}\,. \tag{36}$$

Proof: Denote by $\underline{x}_{\partial a} = \{x_1, \dots, x_k\}$ the zeroth generation variables, by $T_1, \dots, T_k$ the tree factor graphs drawn from $\mathsf{T}_*(r)$ and rooted, respectively, at variable nodes $1, \dots, k$. We then denote by $\mu_i(x_i|\underline{x}_r)$, $i \in \{1, \dots, k\}$ the conditional distribution for variable x_i with respect to the model associated with the tree T_i. We also let $h_i(\underline{x}_r)$ be the associated log-likelihoods (defined analogously to Eq. (18)), and $\underline{h}_i = \max_{\underline{x}_r} h_i(\underline{x}_r)$ ($\overline{h}_i = \min_{\underline{x}_r} h_i(\underline{x}_r)$) be their maximum (minimum) values with respect to the boundary condition.

It is not hard to show that $\langle E_a(\underline{x})\rangle^{(r)}_{\underline{x}_r} = g(h_1(\underline{x}_r), \dots, h_k(\underline{x}_r))$ where the function $g : \mathbb{R}^k \to \mathbb{R}$ is defined as follows

$$g(x_1, \dots, x_k) \equiv \frac{e^{-\beta} \prod_{i=1}^k \frac{1}{2}(1 - \tanh x_i)}{1 - (1 - e^{-\beta}) \prod_{i=1}^k \frac{1}{2}(1 - \tanh x_i)} \,. \tag{37}$$

Since $g(x_1, \dots, x_k)$ is monotonically decreasing in each of its arguments, we have

$$\mathbb{E} \max_{\underline{x}_r, \underline{z}_r} \left| \langle E_a(\underline{x})\rangle^{(r)}_{\underline{x}_r} - \langle E_a(\underline{x})\rangle^{(r)}_{\underline{z}_r} \right| \le \mathbb{E} \left\{ g(\underline{h}_1, \dots, \underline{h}_k) - g(\overline{h}_1, \dots, \overline{h}_k) \right\} , \tag{38}$$

where the couples $(\underline{h}_1, \overline{h}_1), \dots, (\underline{h}_k, \overline{h}_k)$ are i.i.d.'s and distributed as $(\underline{h}_0, \overline{h}_0)$ in the proof of Lemma 3, cf. Eq. (26). In particular, proceeding as in that proof, we deduce that $\mathbb{E} \tanh(\overline{h}_i - \underline{h}_i) \le A\, e^{-\gamma r}$. We are left with the task of proving that this implies an analogous bound on the right hand side of Eq. (38).

To this end, we first consider a single variable function $\tilde{g} : \mathbb{R} \to \mathbb{R}$ with $0 \le \tilde{g}(x) \le 1$ and $-1 \le \tilde{g}'(x) \le 0$. Then

$$\begin{aligned} \mathbb{E}\{\tilde{g}(\underline{h}_1) - \tilde{g}(\overline{h}_1)\} &\le \mathbb{P}\{\overline{h}_1 - \underline{h}_1 \ge \Delta\} + \mathbb{E}\{(\overline{h}_1 - \underline{h}_1)\, \mathbb{I}(\overline{h}_1 - \underline{h}_1 < \Delta)\} \le \\ &\le \frac{1}{\tanh \Delta} \mathbb{E} \tanh(\overline{h}_1 - \underline{h}_1) + \frac{\Delta}{\tanh \Delta} \mathbb{E}\{\tanh(\overline{h}_1 - \underline{h}_1)\, \mathbb{I}(\overline{h}_1 - \underline{h}_1 < \Delta)\} \le \\ &\le \frac{1 + \Delta}{\tanh \Delta} \mathbb{E} \tanh(\overline{h}_1 - \underline{h}_1) \,. \end{aligned} \tag{39}$$

The proof is completed by writing $\mathbb{E}\{g(\underline{h}_1, \dots, \underline{h}_k) - g(\overline{h}_1, \dots, \overline{h}_k)\} = \sum_{i=0}^k \mathbb{E}\{\tilde{g}_i(\underline{h}_i) - \tilde{g}_i(\overline{h}_i)\}$ where $\tilde{g}_i(x) \equiv g(\overline{h}_1 \dots \overline{h}_{i-1}, x, \underline{h}_{i+1}, \dots, \underline{h}_k)$ and noticing that $-1 \le \frac{\partial g}{\partial x_i} \le 0$ (the last statement is proved in the appendix) □

Finally, a result that puts together the above observations to derive the net error in BP estimation.

Lemma 5. *Let $k \ge 2$, $\alpha < \alpha_*(k)$ and $\beta \in [0, \infty]$. Then there exists two positive constants C and $\delta < 1$ such that for any N,*

$$\mathbb{E} \left| \langle E(\underline{x})\rangle - \langle E(\underline{x})\rangle_{\rm BP} \right| \le C N^{\delta} \,. \tag{40}$$

Proof: By linearity of expectation and using Lemma 1, we get

$$\mathbb{E} \left| \langle E(\underline{x})\rangle - \langle E(\underline{x})\rangle_{\rm BP} \right| \le M \mathbb{E} \left| \langle E_a(\underline{x})\rangle - \langle E_a(\underline{x})\rangle_{\rm BP} \right| \le M \mathbb{E} \left\{ \max_{\underline{x}, \underline{z}} \left| \langle E_a(\underline{x})\rangle^{(r)}_{\underline{x}_r} - \langle E_a(\underline{x})\rangle^{(r)}_{\underline{z}_r} \right| \right\} . \tag{41}$$

We would like to apply Lemma 4, but the expectation in the last expression is taken with respect to the formula F drawn from the random k-SAT ensemble, instead of the tree model $\widehat{\mathsf{T}}_*(r)$. However, the quantity in curly brackets depends only of the radius r neighborhood $\mathsf{B}_a(r)$ of vertex a in G_F. Furthermore is non negative and upper bounded by 1. We can therefore apply Lemma 2 and 4 to upper bound the last expression by (here $\mathbb{E}_{\widehat{\mathsf{T}}}$ denotes expectation with respect to the tree ensemble):

$$\begin{aligned} M \left|\right| \mathbb{P}\{\mathsf{B}_i(r) \in \cdot\,\} - \mathbb{P}\{S(r) \in \cdot\,\} \left|\right|_{\rm TV} + M \mathbb{E}_{\widehat{\mathsf{T}}} \left\{ \max_{\underline{x}, \underline{z}} \left| \langle E_a(\underline{x})\rangle^{(r)}_{\underline{x}_r} - \langle E_a(\underline{x})\rangle^{(r)}_{\underline{z}_r} \right| \right\} \le \\ \le A\alpha\, e^{\rho r} + N A' \alpha\, e^{-\gamma r} \end{aligned} \tag{42}$$

The proof is completed by setting $r = \frac{1}{\rho + \gamma} \log N$, which yields Eq. (40) with $\delta = \frac{\rho}{\gamma + \rho}$. □

4 Proofs of Theorems

4.1 Proof of Theorem 1

Clearly, the running time of algorithm described in Section 3 is $O(N^4)$ as total number of BP runs are $O(N^2)$ and each BP run takes $O(N)$ iterations or $O(N^2)$ serial operations. Now, we'll prove Eq. (4).

Using te existing lower bounds on $\alpha_{\rm c}(k,N)$ (see [6] and references therein), it is not hard to show that $\alpha_*(k) \leq \alpha_{\rm c}(k,N)(1-\eta)$ for some $\eta > 0$ all $k \geq 2$ and N large enough. By definition, for $\alpha < \alpha_{\rm c}(k,N)(1-\eta)$, $\beta \in [0,\infty]$ there exists a constant $C(\alpha) > 0$ such that $\log Z(\beta,F) \geq C(\alpha)N\log 2$ whp. This follows from the following two facts for appropriate $C(\alpha)$: (1) at least $C(\alpha)N$ variables do not appear in any clause whp and (2) at least one solution is satisfying assignment whp as $\alpha < \alpha_{\rm c}(k,N)(1-\eta)$. Thus, there are at least $2^{C(\alpha)N}$ satisfying assignment, whence $Z_N(\beta,F) \geq 2^{C(\alpha)N}$. Given this, it is sufficient to show that $|\log Z(\beta,F) - \Phi(\beta,F)| \leq N\varepsilon$ w.h.p. for any $\varepsilon > 0$ and N large enough.

Now, Eqs. (13) and (14) imply that

$$\begin{aligned} |\log Z(\beta,F) - \Phi(\beta,F)| &\leq \sum_{i=0}^{n-1} \left| \log\langle e^{-\Delta E(\underline{x})}\rangle_i + \Delta \langle E(\underline{x})\rangle_{{\rm BP},i} \right| \\ &\leq \sum_{i=0}^{n-1} \left| \log\langle e^{-\Delta E(\underline{x})}\rangle_i + \Delta \langle E(\underline{x})\rangle_i \right| + \sum_{i=0}^{n-1} \Delta \left| \langle E(\underline{x})\rangle_i - \langle E(\underline{x})\rangle_{{\rm BP},i} \right| . \quad (43) \end{aligned}$$

Consider the first term in (43): for any non-negative random variable X, $\log\langle e^{-X}\rangle \leq \langle e^{-X}\rangle - 1 \leq \langle 1 - X + X^2\rangle - 1 \leq -\langle X\rangle + \langle X^2\rangle$. As a consequence, we obtain

$$\sum_{i=0}^{n-1} \left| \log\langle e^{-\Delta E(\underline{x})}\rangle_i + \Delta \langle E(\underline{x})\rangle_i \right| \leq \sum_{i=0}^{n-1} \Delta^2 \langle E(\underline{x})^2\rangle_i \leq \beta\Delta \sup_i \langle E(\underline{x})^2\rangle_i \leq N^{2\delta'}\alpha^2 , \quad (44)$$

where we used $\beta \leq N^{\delta'}$, $\Delta = \beta/N^2 \leq N^{\delta'-2}$ and $0 \leq E(\underline{x}) \leq N\alpha$. If we choose $\delta' < 1/2$, this contribution is smaller than $N\varepsilon/2$ for all N large enough.

Now, the second term in Eq. (43): the bound (40) holds for any β in the compact region $[0,\infty]$. Furhter, the left hand side is uniformly bounded (in terms of N) and continuous in β. Hence, there exists a C so that the bound (40) holds uniformly for $\beta \in [0,\infty]$. This will imply that

$$\sum_{i=0}^{n-1} \Delta\, \mathbb{E} \left| \langle E(\underline{x})\rangle_i - \langle E(\underline{x})\rangle_{{\rm BP},i} \right| \leq \beta C N^{\delta} \leq C N^{\delta+\delta'} \quad (45)$$

Choosing $\delta' \in (0, 1-\delta)$ and Markov inequality will imply that the second term is also bounded above by $N\varepsilon/2$ whp. This completes the proof of Theorem 1. □

4.2 Proofs of Theorems 2, 3, and 4

Due to shortage of space, they are moved to Appendix A.

5 Discussion and Future Work

We presented a novel deterministic algorithm for approximately counting good truth assignments of random k-SAT formula with high probability. The algorithm is built upon the well-known Belief Propagation heuristic and an interpolation method for the log-partition function. In the process of establishing the correctness of the algorithm, we obtained the threshold for uniqueness of Gibbs distribution for random k-SAT formula as $2k^{-1}\log k(1+o_k(1))$. This result if of interest in its own right.

We believe that our result can be extended to a reasonable class of non-random k-SAT formula. We also believe that the approximation guarantees of Theorem 1 should hold for any $\beta \in [0,\infty]$.

References

[1] H. O. Georgii, "Gibbs Measures and Phase Transitions". Berlin, Walter de Gruyter and Co., 1988

[2] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, "Determining computational complexity from characteristic 'phase transitions' ", Nature 300 (1999), 133–137

[3] M. Mézard, G. Parisi, and R. Zecchina, "Analytic and Algorithmic Solution of Random Satisfiability Problems", Science 297 (2002), 812–815

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

[5] M. Talagrand, " The high temperature case of the K-sat problem", Probability Theory and Related Fields 119, 2001, 187-212.

[6] D. Achlioptas and Y. Peres, "The threshold for random k-SAT is $2k\log 2 - O(k)$", Journal of the AMS, 17 (2004), 947-973.

[7] R. Monasson and R. Zecchina, "Entropy of the K-Satisfiability Problem", Phys. Rev. Lett. 76 (1996), 38813885

[8] S. Franz and M. Leone, "Replica bounds for optimization problems and diluted spin systems", Journal of Statistical Physics, 111 (2003), 535.

[9] S. Franz, M. Leone and F. L. Toninelli, "Replica bounds for diluted non-Poissonian spin systems", Journal of Physics, A 36 (2003) 10967 .

[10] M. Jerrum and A. Sinclair, "Polynomial-time Approximation Algorithms for the Ising Model", SIAM Journal on Computing 22 (1993), pp. 1087-1116.

[11] D. Weitz, "Counting independent sets up to the tree threshold", In Proceedings of STOC, 2006.

[12] A. Bandyopadhyay and D. Gamarnik, "Counting without sampling. New algorithms for enumeration problems using statistical physics", In Proceedings of SODA, 2006.

[13] A. Montanari and D. Shah, "k-SAT: Counting Satisfying Assignment and Threshold for Correlation Decay", Longer version, in preparation.

[14] E. Friedgut, "Sharp Thresholds of Graph Proprties, and the k-sat Problem", Journal of American Mathematical Society, 12 (1999), no. 4, 1017–1054.

[15] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", San Francisco, CA: Morgan Kaufmann, 1988.

[16] M. Wainwright and M. Jordan, "Graphical models, exponential families, and variational inference," *Tech. Report*, Dept. of Stat.,University of Cal., Berkeley, 2003.

[17] S. Tatikonda and M. Jordan, "Loopy Belief Propagation and Gibbs Measure," Berkeley Working Paper, 2002.

A Proof Sketches: Theorems 2, 3 and 4

Due to space limitations, we only provide sketch of proofs for Theorems 2, 3 and 4.

Proof sketch for Theorem 2. By using the definition $\kappa(\alpha_*) = 1$ (with $\kappa(\alpha)$ being defined as in Eq. (3)), it is easy to show that $\alpha_*(k) = 2k^{-1}\log k\{1 + O(\log\log k/\log k)\}$. To complete the proof, we need a (asymptotically in k) matching upper bound. In order to obtain such an upper bound, we consider the case $\beta = \infty$, i.e. only satisfying assignments have positive weight. Consider a tree formula which is distributed as $\mathsf{T}_*(r)$. Let P_r be the probability that there exists two boundary conditions $\underline{x}_r^{(0)}$, $\underline{x}_r^{(1)}$, such that the root takes values, respectively, 0 or 1 in all the satisfying assignments with the respective boundary conditions. Clearly for the Gibbs measure to be unique (or have correlation decay) in the sense of Definition 1 (but also in the weaker sense correspondint to the threshold $\alpha'_{\rm u}(k)$), it must be that $P_r \to 0$ as $r \to \infty$. Hence, if we establish that for $\alpha > 2k^{-1}\log k\{1 + O(\log\log k/\log k)\}$, there

exists such boundary conditions with positive probability, then the proof will be complete. Next, we do that.

For this, consider a tree from $\mathsf{T}_*(r)$ with the root having degree 1. Given such a tree, let ρ_r be the probability that there exists a boundary condition $\underline{x}_r$, such that the root variable is the only variable that satisfies the only clause in which it belongs (recall that the root variable has degree 1) for all possible satisfying assignments with the given boundary condition. If $P_r \to 0$, then $\rho_r \to 0$. To prove this claim, assume by contraddiction that ρ_r remains bounded away from zero (say $\rho_r \geq \underline{\rho} > 0$) and consider an tree from $\mathsf{T}_*(r)$ (without conditioning). With finite probability the root belongs to two clauses in which it appears, respectively, directed and negated. With probability at least $\underline{\rho}^2 > 0$, for each of the corresponding subtrees there exists a boundary condition that fixes the root variable to be (respectively) directed or negated. By extending arbitrarily this boundary conditions to the full tree, we obtain the desired $\underline{x}_r^{(1)}$, $\underline{x}_r^{(0)}$.

It turns out that ρ_r can be determined recursively. Set $\rho_0 = 1$ and $\rho_{r+1} = \{1 - \exp(-k\alpha\rho_r/2)\}^{k-1}$. Recursively, $\rho_r \to 0$ as $r \to \infty$ only if $\alpha < \alpha^*(k)$, where $\alpha^*(k)$ for the above recursion (with little bit of algebra) evaluates to $\alpha^*(k) = 2k^{-1}\log k\{1 + O(\log\log k/\log k)\}$. This completes the proof sketch of Theorem 2.

Proof sketch for Theorem 3. First notice that, if F and F' differ in a single clause, then $|\log Z(\beta, F) - \log Z(\beta, F')| \leq 2\beta$. Hence, by application of Azuma-Hoeffding's inequality, it follows that $|\log Z - \mathbb{E}\log Z| \leq N\delta$ with probability at least $1 - e^{-NC_\beta\delta^2}$, for some $C_\beta > 0$ for any $\beta \in [0, \infty)$. Given this, to obtain the almost sure convergence as in (8), it is sufficient to prove that $\lim_{N\to\infty} N^{-1}\mathbb{E}\,\Phi(\beta, F) = \phi(\beta)$, in light of Theorem 1 and Borel-Cantelli's Lemma.

To do so, first we need to establish that

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\langle E(\underline{x})\rangle_{\mathrm{BP},\beta} = \alpha\mathbb{E}g(h_1, \dots, h_k)\,, \tag{46}$$

where g is defined as in Eq. (37); the random variables $h_1, \dots, h_k$ are i.i.d. with distribution ν^* that is fixed point of operator S as defined in the statement of Theorem 3. We claimed that the fixed point is unique for S. To justify this claim, first note that the image of S is contained in the space of distributions supported on $[0, \beta/2]$, call it $\mathcal{D}_\beta$, which is a compact space with respect to the weak topology. Being continuous on $\mathcal{D}_\beta$, S admits at least one fixed point in it. Moreover, the contraction condition implied by the correlation decay (proved as a part of Theorem 2) implies the attractiveness as well as the uniqueness of the fixed point of S.

Once we establish existence of the unique fixed point, the (46) follows from Lemma 2 and correlation decay established in Theorem 2. Now, by integrating Eq. (46) over β and observing that $\beta_{i+1} - \beta_i = \beta/N^2$ (hence integration error is negligible at scale $1/N$) one gets

$$\lim_{N\to\infty} N^{-1}\mathbb{E}\,\Phi(\beta, F) = \log 2 - \alpha\int_0^\beta \mathbb{E}_{\beta'} g(h_1, \dots, h_k)\,\mathrm{d}\beta'\,, \tag{47}$$

where a subscript has been added in $\mathbb{E}_{\beta'}$ to stress that the fixed point distribution has to be taken at inverse temperature β'. The proof of Theorem 3 is completed by showing that the integral on the right hand side of the last equation is given by $\phi(\beta)$ as in Eq. (9). In fact, by taking the derivative of this expression wrt β, one gets a contribution coming from the explicit β dependence, which evaluates to $-\alpha\mathbb{E}g(h_1, \dots, h_k)$, and one from the β dependence of the fixed poit distribution, that can be shown to vanish.

Proof Sketch of Theorem 4. For the ease of notation, let $Z(\beta) \equiv Z_N(\beta, F)$, $\Xi(\zeta) \equiv \Xi(\zeta, F)$ and $U(\beta) \equiv \langle E(\underline{x})\rangle_{\beta,F}$. Because of Theorem 1, it is sufficient to prove that $|\log\Xi(N\epsilon) - \log Z(\beta)| \leq N\epsilon^a$ whp. This follows from two inequalities.

First inequality. For any $\zeta \geq 0$,

$$Z(\beta) \;=\; \sum_{\underline{x}:E(\underline{x})\geq\zeta} e^{-\beta E(\underline{x})} + \sum_{\underline{x}:E(\underline{x})<\zeta} e^{-\beta E(\underline{x})} \geq e^{-\beta\zeta}\Xi(\zeta). \tag{48}$$

Second inequality. For any $\zeta \geq 0$ and using the first equality in (48), we obtain

$$Z(\beta) \quad \leq \quad \sum_{\underline{x}: E(\underline{x}) \geq \zeta} e^{-\beta E(\underline{x})} + \Xi(\zeta).$$

Equivalently, $Z(\beta)\mu(E(\underline{x}) < \zeta) \leq \Xi(\zeta)$. Now, take $\zeta = 2U(\beta)$ then, we get using Markov's inequality

$$\mu\{E(\underline{x}) < 2U(\beta)\} \quad \geq \quad 1 - \frac{U(\beta)}{2U(\beta)} \quad = \quad \frac{1}{2}. \tag{49}$$

¿From (48) and (49), we obtain

$$\log Z(\beta) - \log 2 \quad \leq \quad \log \Xi(2U(\beta)) \quad \leq \quad \log Z(\beta) + 2\beta U(\beta). \tag{50}$$

The next sep consists in controlling $U(\beta)$ at large β. Arguing analogously to the proof of Theorem 1 one can show that there exist constants C_1, C_2, C_3, $a > 0$ such that, for any $\beta \in [0, \infty]$, $NC_1 e^{-2\beta} \leq U(\beta) \leq NC_2 e^{-b\beta} + C_3 N^{\delta}$ whp.

Fix β_1 in such a way that $2C_1 e^{-2\beta_1} = \varepsilon$. Then $2U(\beta_1) \geq N\varepsilon$ whp. By the upper bound in Eq. (50) and monotonicit of $\Xi(\zeta)$, we get

$$\log \Xi(N\varepsilon) \leq \log Z(\beta_1) + 2\beta_1 U(\beta_1) \leq \log Z(\beta_1) + 2\beta_1 NC_2 e^{-b\beta_1} + 2\beta_1 C_3 N^{\delta}\,. \tag{51}$$

Using the definition of β_1, which gives $\beta_1 = \frac{1}{2}\log\frac{2C_1}{\varepsilon}$, we get that there exists $C, a > 0$ such that

$$\log \Xi(N\varepsilon) \leq \log Z(\beta_1) + NC\varepsilon^a\,. \tag{52}$$

with high probability.

The lower bound on $\log \Xi(N\varepsilon)$ is proved analogously by taking β_2 such that $2C_2 e^{-b\beta} + 2C_3 N^{-1+\delta} = \varepsilon$ thus getting $\log \Xi(N\varepsilon) \geq \log Z(\beta_2) - NC\varepsilon^a$ whp. One concludes by bounding the difference of the two partition functions: $|\log Z(\beta_2) - \log Z(\beta_1)| \leq U(\beta_2)|\beta_1 - \beta_2| \leq NC\varepsilon^a$ whp. □