Proof of the Satisfiability Conjecture for Large k

[Extended Abstract]*

Jian Ding
Department of Statistics
University of Chicago
Chicago, Illinois
jianding@galton.uchicago.edu

Allan Sly
Department of Statistics
Univ. of California–Berkeley
Berkeley, California
sly@stat.berkeley.edu

Nike Sun
Microsoft Research and
Dept. of Mathematics, MIT
Cambridge, Massachusetts
nisun@microsoft.com

ABSTRACT

We establish the satisfiability threshold for random k-SAT for all $k \geqslant k_0$. That is, there exists a limiting density $\alpha_{\rm s}(k)$ such that a random k-SAT formula of clause density α is with high probability satisfiable for $\alpha < \alpha_{\rm s}$, and unsatisfiable for $\alpha > \alpha_{\rm s}$. The satisfiability threshold $\alpha_{\rm s}(k)$ is given explicitly by the one-step replica symmetry breaking (1RSB) prediction from statistical physics. We believe that our methods may apply to a range of random constraint satisfaction problems in the 1RSB class.

Categories and Subject Descriptors

G.2 [**Discrete Mathematics**]: Graph theory

General Terms

Theory

Keywords

Constraint satisfaction problem, random k-SAT, satisfiability threshold, condensation, replica symmetry breaking, belief propagation, survey propagation

1. INTRODUCTION

Random k-SAT is a natural and well-studied model of a random *constraint satisfaction problem* (CSP). Advances in the understanding of random CSPs have been contributed by researchers from several different communities, including computer science, probability, combinatorics, and statistical physics. Much of this work concerned a sharp satisfiability transition that was conjectured to occur as the clause density α is increased past some critical threshold $\alpha_{\rm s}$. In this work we prove the conjecture for large k.

*The full version of this paper is available as an online preprint at `http://arxiv.org/abs/1411.0650`.

STOC '15, June 14-17, 2015, Portland, Oregon, USA.

ACM 978-1-4503-3536-2/15/06 ...$15.00.
DOI: http://dx.doi.org/10.1145/2746539.2746619.

An instantiation of random k-SAT on n variables at clause density α is defined by a k-CNF formula $\phi : \{+,-\}^n \to \{+,-\}$ (with $+ \equiv$ TRUE, $- \equiv$ FALSE), consisting of M clauses where M is a Poisson($n\alpha$) random variable, and each clause is the disjunction of k literals chosen independently and uniformly at random from $\{+x_1, -x_1, \dots, +x_n, -x_n\}$. We say ϕ is SAT if $\phi^{-1}(+) \neq \varnothing$. An example of a 3-CNF formula with $n = 5$ and $M = 2$ is given by

$$\phi(\underline{x}) = (+x_1 \text{ OR } -x_3 \text{ OR } +x_4) \text{ AND } (-x_1 \text{ OR } -x_2 \text{ OR } +x_5).$$

Write $\mathbb{P} \equiv \mathbb{P}_n^\alpha$ for probability under the above model. It is widely conjectured that for each fixed $k \geqslant 2$, random k-SAT has a *sharp satisfiability threshold*: that is, there exists a positive constant $\alpha_{\rm s}$ — depending on k but not on n — such that for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n^{\alpha_{\rm s}+\epsilon}(\text{SAT}) = 1 = \lim_{n\to\infty} \mathbb{P}_n^{\alpha_{\rm s}+\epsilon}(\text{UNSAT}). \tag{1}$$

This is known for $k = 2$, with $\alpha_{\rm s} = 1$ [11, 19, 25]. For $k \geqslant 3$, however, even existence of $\alpha_{\rm s}$ has remained a long-standing open question. A breakthrough by Friedgut establishes, for all k, existence of a *sharp threshold sequence* $\alpha_{\rm s}(n)$ [22]. The sequence may not converge as $n \to \infty$, but is known to be within ϵ_k of $2^k \ln 2 - (1 + \ln 2)/2$ for n large, where $\epsilon_k \to 0$ in the limit $k \to \infty$ [26, 16]. By heuristic methods, physicists conjecture an explicit value $\alpha_\star$ for $\alpha_{\rm s}$, the *one-step replica symmetry breaking* (1RSB) *threshold*, which is expected to be correct for all $k \geqslant 3$ [29, 33]. The main result of this work resolves the k-SAT threshold conjecture for large k, proving that $\alpha_{\rm s}$ exists and matches the predicted value $\alpha_\star$:

THEOREM 1. *There exists an absolute constant k_0 such that for all $k \geqslant k_0$, the satisfiability threshold $\alpha_{\rm s}$ for random k-SAT exists, with explicit value given by $\alpha_\star$.*

The explicit characterization $\alpha_\star$ is as follows. We assume throughout the paper, even when not explicitly stated, that $k \geqslant k_0$ and $2^k \ln 2 - 2 \equiv \alpha_{\rm lbd} \leqslant \alpha \leqslant \alpha_{\rm ubd} \equiv 2^k \ln 2$. Sample d^+, d^- independently from the Poisson($\alpha k/2$) distribution, and write $\underline{d} \equiv (d^+, d^-)$. Write $\mathscr{P}$ for the space of probability measures on $[0,1]$, and define a mapping $\boldsymbol{R} : \mathscr{P} \to \mathscr{P}$ as follows. Given $\mu \in \mathscr{P}$, generate (independently of $\underline{d}$) an array $\underline{\eta} \equiv [(\eta_j)_{j\geqslant 1}, (\eta_{ij}^+, \eta_{ij}^-)_{i,j\geqslant 1}]$ of i.i.d. samples from μ. Define $\boldsymbol{R}\mu \in \mathscr{P}$ to be the law of $R \equiv R(\underline{d}, \underline{\eta})$, defined by

$$R \equiv \frac{(1-\Pi^-)\Pi^+}{\Pi^+ + \Pi^- - \Pi^+\Pi^-}, \text{ where } \Pi^{\pm} \equiv \prod_{i=1}^{d^{\pm}} \Big(1 - \prod_{j=1}^{k-1} \eta_{ij}^{\pm}\Big).$$

PROPOSITION 1. *Let $\mu_\ell \in \mathscr{P}$ $(\ell \geqslant 0)$ be the sequence of probability measures defined by $\mu_0 = \delta_{1/2}$, and $\mu_\ell = \boldsymbol{R}\mu_{\ell-1}$ for all $\ell \geqslant 1$. For $k \geqslant k_0$ and $\alpha_{\mathrm{lbd}} \leqslant \alpha \leqslant \alpha_{\mathrm{ubd}}$, this sequence converges to a limit $\mu = \mu_\alpha \in \mathscr{P}$, satisfying $\boldsymbol{R}\mu = \mu$. Let*

$$\Phi(\alpha) = \mathbb{E}\Big[\ln \frac{\Pi^+ + \Pi^- - \Pi^+\Pi^-}{(1 - \prod_{j=1}^k \eta_j)^{(k-1)\alpha}} \Big],$$

where $\mathbb{E}$ denotes expectation over $(\underline{d}, \underline{\eta})$ where entries of $\underline{d}$ are i.i.d. from Poisson$(\alpha k/2)$ *and entries of $\underline{\eta}$ are i.i.d. from μ. This function is well-defined and strictly decreasing in α, with a unique zero $\alpha_\star = \alpha_\star(k)$.*

Our proof is heavily guided by insights emerging from the statistical physics analysis of random k-SAT and related problems. In the remainder of this introductory section we briefly survey some of this literature, beginning with a discussion of the main obstacles in determining α_{s}.

1.1 Obstacles in computing the threshold

Let $\mathrm{SOL}(\phi)$ denote the set $\phi^{-1}(\mathtt{+}) \subseteq \{\mathtt{+},\mathtt{-}\}^n$ of satisfying assignments of ϕ. Clearly, ϕ is SAT if and only if $Z \equiv |\mathrm{SOL}(\phi)|$ is positive. For any non-negative integer random variable X, we have the first and second moment inequalities:

$$\frac{(\mathbb{E}X)^2}{\mathbb{E}[X^2]} \leqslant \mathbb{P}(X > 0) \leqslant \mathbb{E}X \tag{2}$$

where $\mathbb{E}$ denotes expectation under $\mathbb{P}$. A natural approach to bounding α_{s} is to apply (2) with $X = Z$ and $\mathbb{P} = \mathbb{P}^n_\alpha$. For example, $\mathbb{E}Z = \exp\{n[\ln 2 - \alpha/2^k]\}$, giving $\alpha_{\mathrm{s}} \leqslant 2^k \ln 2$ which is within $O(1)$ of the true threshold. (We have abused notation somewhat, since *a priori* a sharp threshold α_{s} may not exist. Throughout this paper, $\alpha_{\mathrm{s}} \leqslant \alpha$ means formally that $\lim_n \mathbb{P}(\mathrm{SAT}) = 0$ at any fixed density above α. Likewise, $\alpha_{\mathrm{s}} \geqslant \alpha$ means formally that $\lim_n \mathbb{P}(\mathrm{SAT}) = 1$ at any fixed density below α.) If $\mathbb{E}[Z^2] \lesssim (\mathbb{E}Z)^2$ holds at density α, then the inequality on the left-hand side in (2), combined with Friedgut's theorem [22], gives $\alpha_{\mathrm{s}} \geqslant \alpha$.

It is known however that the distribution of Z is highly non-concentrated, such that $\mathbb{E}[Z^2] \gg (\mathbb{E}Z)^2$ at *any* positive α, meaning the second moment inequality fails to give any non-trivial lower bound on α_{s}. It is also known that the first moment upper bound is not sharp. The central difficulty of random k-SAT is that this non-concentration has two distinct sources, as we now explain:

Clustering and condensation

A key insight from the statistical physics research on random k-SAT is that the non-concentration of $Z \equiv |\mathrm{SOL}|$ is caused in part by a peculiarity in the (typical) geometry of the random set SOL: for α in a non-trivial *condensation regime* $(\alpha_{\mathrm{c}}, \alpha_{\mathrm{s}})$, a dominating contribution to $\mathbb{E}Z$ comes from the rare event of seeing an atypically large *cluster* of very similar solutions [27, 35]. This results in non-concentration with $Z \gg \mathbb{E}Z$ on the rare event, versus $Z \ll \mathbb{E}Z$ in the typical picture.

The condensation phenomenon is one aspect of a detailed phase diagram which is conjectured for a class of random CSPs that includes random k-SAT, as well as the coloring and independent set problems on sparse random graphs. We refer the reader to [27] for details and further references, summarizing here only some salient features: up to some density α_{d}, almost all of the mass in SOL lies within a single well-connected subset of the Hamming cube $\{\mathtt{+},\mathtt{-}\}^n$. The geometry changes abruptly at α_{d}, above which most of the mass in SOL is roughly equidistributed among exponentially many clusters. The *clustering threshold* α_{d} empirically matches the density above which most algorithms fail; there is some rigorous support for this link [2, 24, 38]. A further transition occurs at the *condensation threshold* $\alpha_{\mathrm{c}} \in (\alpha_{\mathrm{d}}, \alpha_{\mathrm{s}})$, above which $\mathbb{E}Z$ becomes dominated by rare large clusters — this has been rigorously confirmed in the coloring model, with explicit α_{c} [6].

This conjectural phase diagram is derived largely on the basis of an analogy between random CSPs and *spin glasses*, classical models of disordered magnets [32]. Physicists have observed this analogy since the 1980s [31], and the study of random CSPs within the spin glass framework has yielded rich insights. Given a random k-SAT instance, let ν be the uniform measure on its solution space SOL. Since SOL is a random set, ν is a random measure on $\{\mathtt{+},\mathtt{-}\}^n$, and is what is termed a *dilute* spin glass [21], where *dilute* refers to the sparsity (bounded density) of constraints.

The measure ν exhibits *replica symmetry breaking* (RSB) if two independent samples $\underline{x}, \underline{x}'$ from ν (two *replicas*) have non-trivial overlap structure. For sparse random CSPs, it is conjectured [27, 35] that the condensation threshold α_{c} marks the onset of RSB. Indeed, below condensation, SOL consists of either a single large cluster, or else exponentially many well-separated clusters of roughly equal size. In either case the (normalized) overlap $n^{-1}\langle \underline{x}, \underline{x}' \rangle$ is expected to be concentrated near zero. In contrast, in the condensation regime $\alpha \in (\alpha_{\mathrm{c}}, \alpha_{\mathrm{s}})$, it is believed that a *bounded* number of clusters carry most of the mass of SOL. The two replicas may still lie in distinct clusters (with trivial overlap), but now they lie with non-negligible probability in the same cluster, corresponding to a large overlap.

It is believed that in the *replica symmetric* (RS) regime below α_{c}, the measure ν has correlation decay, and is well-approximated by RS heuristic methods (belief propagation). It is further conjectured that random k-SAT in the RSB regime $(\alpha_{\mathrm{c}}, \alpha_{\mathrm{s}})$ exhibits *one-step replica symmetry breaking* (1RSB): this means that the overlap distribution is supported on two values, or alternatively that *clusters are replica symmetric*. Writing $\boldsymbol{\Omega}$ for the number of clusters in SOL,[1] a heuristic RS calculation yields the prediction

$$\boldsymbol{\Omega} \approx \exp\{n\Phi(\alpha)\} \tag{3}$$

with Φ as in the statement of Thm. 1. The 1RSB threshold $\alpha_\star$, defined as the root of Φ, is thus the predicted threshold for the existence of clusters.

Graphical fluctuations

Throughout the following, a k-CNF formula will be represented by a bipartite factor graph $\mathscr{G} = (V, F, E)$ with vertex set $V \cup F$ partitioned into variables V and clauses F, and with (undirected) edges E joining variable to clauses. Write $n \equiv |V|$ and $m \equiv |F|$; we generically denote variables u, v, w, clauses a, b, c, and edges $e = (av) = (va)$. Edge $e = (av)$ is equipped with a sign $\mathtt{L}_e = \mathtt{L}_{av} \in \{\mathtt{+},\mathtt{-}\}$ indicating whether the inclusion of variable v in clause a is affirmative ($\mathtt{L}_{av} = \mathtt{+}$) or negative ($\mathtt{L}_{av} = \mathtt{-}$). We will take all edge lengths to be $1/2$, so two variables are within unit distance if they participate in the same clause.

We have already explained above that for $\alpha \in (\alpha_{\mathrm{c}}, \alpha_{\mathrm{s}})$, the rare event of an atypically large cluster is a source of non-

[1] For the moment, take $\boldsymbol{\Omega}$ to be the number of connected components of SOL.

concentration for Z. For the random k-SAT model, however, it is well known that *at any positive* α, there is already non-concentration caused by fluctuations in the graph structure. We refer the reader to [3, 4, 15, 16] for further discussions of this issue. To give a simple example, for a variable v let $\partial v(\pm)$ denote the set of clauses in which v appears with sign $\pm$. The *degree distribution* $\mathscr{D}$ of $\mathscr{G}$ is the empirical distribution of pairs $(|\partial v(+)|, |\partial v(-)|)$: for each pair (d^+, d^-) of non-negative integers, $\mathscr{D}(d^+, d^-)$ records the fraction of variables v with $|\partial v(\pm)| = d^\pm$. Then $\mathscr{D}$ fluctuates across different samples of $\mathscr{G}$, so we can decompose

$$\mathbb{E}Z = \sum_{\mathscr{D}} \mathbb{P}(\mathscr{D})\mathbb{E}[Z|\mathscr{D}]. \tag{4}$$

In the limit of large n, $\mathscr{D}$ is concentrated near the typical degree distribution, a product of Poisson$(\alpha k/2)$:

$$\mathscr{D}^{\text{typ}}(d^+, d^-) = \frac{e^{-k\alpha}(k\alpha/2)^{d^+ + d^-}}{(d^+)!(d^-)!},$$

with gaussian fluctuations: $\mathbb{P}(\mathscr{D}) \approx \exp\{-n\|\mathscr{D} - \mathscr{D}^{\text{typ}}\|^2\}$.[2] In contrast, $\mathbb{E}[Z|\mathscr{D}]$ has no reason to be stationary at $\mathscr{D}^{\text{typ}}$, and in fact we expect behavior of the form

$$\mathbb{E}[Z|\mathscr{D}] \approx \mathbb{E}[Z|\mathscr{D}^{\text{typ}}] \exp\{n\langle c, \mathscr{D} - \mathscr{D}^{\text{typ}}\rangle\}.$$

Comparing these approximations, we see that it is always advantageous to pay a large deviations cost in $\mathbb{P}(\mathscr{D})$ to gain in $\mathbb{E}[Z|\mathscr{D}]$, and as a result the first moment (4) will be dominated by an *atypical* degree distribution $\mathscr{D}^\star \neq \mathscr{D}^{\text{typ}}$, where $\mathbb{P}(\mathscr{D}^\star)$ is exponentially small but $\mathbb{E}[Z|\mathscr{D}^\star]$ is exponentially large compared with $\mathbb{E}Z$.

In fact this issue goes far beyond the degree fluctuations: for any R, we can define $\mathscr{D}_R$ to be the empirical distribution of R-neighborhood types in the graph (recalling that we set edge lengths to be $1/2$, the degree distribution corresponds to $R = 1/2$). Under the random k-SAT probability measure $\mathbb{P} = \mathbb{P}_n^\alpha$, $\mathscr{G}$ is a sparse random graph with few short cycles, that is to say, it locally has the structure of a tree. In fact we can explicitly describe the local structure by the PGW^α (Poisson Galton–Watson) random tree: this tree is rooted at a variable, and has alternating layers of variables and clauses generated in a random manner: each variable independently produces Poisson(αk) child clauses; each clause produces $k-1$ child variables; and each edge is labelled with a random sign $\mathsf{L} \in \{+, -\}$. In the limit $n \to \infty$, $\mathscr{G}$ *converges locally in distribution*, in the formal sense defined by [5, 7], to the PGW^α tree. This means equivalently that for each fixed R, as $n \to \infty$, $\mathscr{D}_R$ concentrates near $\mathscr{D}_R^{\text{typ}}$ which is the law of the first R levels of PGW^α (the tree can be finite, but it is infinite with positive probability). However, by analogous considerations as above, the first moment $\mathbb{E}Z$ *is dominated by atypical* $\mathscr{D}_R$ *for any depth* R. Further, for $\ell > R$ we can decompose

$$\mathbb{E}[Z|\mathscr{D}_R] = \sum_{\mathscr{D}_\ell} \mathbb{P}(\mathscr{D}_\ell|\mathscr{D}_R)\mathbb{E}[Z|\mathscr{D}_\ell],$$

where the sum is taken over all $\mathscr{D}_\ell$ compatible with $\mathscr{D}_R$. Even if we condition on $\mathscr{D}_R$ near $\mathscr{D}_R^{\text{typ}}$, $\mathbb{E}[Z|\mathscr{D}_R]$ is dominated by atypical $\mathscr{D}_\ell$ for all $\ell > R$, meaning $Z \ll \mathbb{E}[Z|\mathscr{D}_R^{\text{typ}}]$ with high probability for all R.

[2]This approximation holds with an appropriate choice of $\|\cdot\|$, comparable with the standard ℓ^2 norm for fixed k, α.

In the condensation regime, the 1RSB heuristic asserts that if we replace Z with the number $\mathbf{\Omega}$ of clusters, then we will remove the non-concentration caused by atypically large clusters. The graphical fluctuations remain, causing

$$\mathbf{\Omega} \ll \mathbb{E}[\mathbf{\Omega}|\mathscr{D}_R^{\text{typ}}] \text{ for all } R \tag{5}$$

(with high probability). Let us remark briefly that the non-concentration (5), as caused by graphical fluctuations, is not an issue in the physicists' prediction of $\alpha_\star$. Indeed, according to the 1RSB heuristic, the uniform measure on clusters exhibits correlation decay. Physicists therefore estimate $\mathbf{\Omega}$ by assuming this correlation decay, and directly working with an analogous model of clusters defined on the PGW^α tree. Of course, this completely circumvents (nonrigorously) the issue of atypical neighborhood profiles $\mathscr{D}_R$ in the random graph $\mathscr{G}$. The tree-based cluster model is characterized by the distributional fixed point $\boldsymbol{R}\mu = \mu$ of Propn. 1, and the corresponding exponent that governs the growth rate of $\mathbf{\Omega}$ (see (3)) is $\Phi(\alpha)$.

1.2 Prior rigorous results

As noted above, the k-SAT threshold for $k = 2$ was rigorously determined in a few independent works [11, 19, 25]; even finer results characterizing the scaling window were subsequently obtained [8]. The case $k = 2$ is unique in that there is no condensation regime, which is believed to appear for all $k \geqslant 3$; as a result the mechanisms governing the $k = 2$ satisfiability transition are quite different. Until recently, all exact satisfiability transitions obtained for sparse random CSPs have been obtained for models without condensation regimes, such as XOR-SAT [34, 37] and 1-in-k-SAT [1].

For $k \geqslant 3$, the first moment method has long been known to yield fairly accurate upper bounds [20]. A more subtle first moment calculation, restricted to "locally maximal" solutions, achieved a more accurate upper bound of

$$\alpha_{\text{s}} \leqslant 2^k \ln 2 - \tfrac{1}{2}(1 + \ln 2) + \epsilon_k \quad [26]$$

with $\lim_{k\to\infty} \epsilon_k = 0$. This bound is correct up to the second-order asymptotic term. By contrast, the early lower bounds for k-SAT, which were generally algorithmic, missed the true threshold by a large factor. Indeed, the best algorithmic lower bounds to date give

$$\alpha_{\text{s}} \geqslant \max\{1.817, (1 - \epsilon_k)\ln k\} \cdot 2^k/k \quad [23, 12],$$

which is off from the true threshold by a factor of order $k/(\ln k)$ when k is large.

In a major breakthrough, Friedgut [22] applied methods of discrete Fourier analysis to show that random k-SAT has a sharp threshold *sequence* $\alpha_{\text{s}}(n)$: for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}_n^{\alpha_{\text{s}}(n)-\epsilon}(\text{SAT}) = 1 = \lim_{n\to\infty} \mathbb{P}_n^{\alpha_{\text{s}}(n)+\epsilon}(\text{UNSAT}).$$

Friedgut's result leaves open the possibility that one needs to take $\alpha_{\text{s}}(n)$ non-convergent with n. In contrast, the conjecture (1) states that the above holds with $\alpha_{\text{s}}(n) = \alpha_{\text{s}}$ independent of n.

Subsequent advances in lower bounding the satisfiability transition have all followed the same basic approach, which we also take in this paper. First, the second moment bound (left-hand side of (2)) is used to prove satisfiability with positive probability: $\liminf_{n\to\infty} \mathbb{P}(\text{SAT}) > 0$ at some density α. Friedgut's theorem immediately implies $\lim_{n\to\infty} \mathbb{P}(\text{SAT}) = 1$ at any density less than α, therefore $\alpha_{\text{s}} \geqslant \alpha$.

As discussed above, applying the second moment bound from (2) with the most obvious choice $X = Z$ fails to give *any* non-trivial lower bound on α_{s}. Improved lower bounds have been obtained by increasingly sophisticated choices for the random variable X. Achlioptas and Moore obtained the first satisfiability lower bound to achieve the correct asymptotic order (asymptotic in the limit $k \to \infty$),

$$\alpha_{\mathrm{s}} \geqslant 2^{k-1} \ln 2 - O(1) \quad [3].$$

This was proved by taking X to be Z_{NAE}, the number of NAE-SAT solutions for the random k-CNF. This restriction symmetrizes the SAT problem, thereby eliminating the local fluctuations issue from the second moment calculation. On the other hand, being NAE-SAT is roughly "twice as difficult" as being SAT, so the restriction is quite prohibitive, costing roughly a factor 2 in the lower bound.

Achlioptas and Peres applied the second moment method with a more subtle symmetrization technique which is much less prohibitive, yielding the greatly improved lower bound

$$\alpha_{\mathrm{s}} \geqslant 2^k \ln 2 - O(k) \quad [4].$$

In the limit of large k, this is correct in the leading term. In random k-SAT, the non-stationarity of $\mathbb{E}[Z|\mathscr{D}]$ at $\mathscr{D}^{\text{typ}}$ is closely tied to the fact that variables lean towards majority: conditioned on the random k-CNF, if $\underline{x}$ is sampled uniformly at random from SOL, it is typically the case that x_v is positively correlated with $|\partial v(\texttt{+})| - |\partial v(\texttt{-})|$ for each variable v. The major innovation of Achlioptas–Peres was to apply the second moment method on a *weighted* number of solutions, where the weighting penalizes for the total number of satisfied literal occurrences. This effectively *balances* the variable spins x_v, decoupling them from the degree fluctuations and allowing the second moment approach to succeed. The improved lower bound [4] results because the weighted count captures a much larger slice of SOLcompared with the NAE-SAT solutions.

Coja-Oghlan and Panagiotou subsequently improved this approach by introducing a step of conditioning on the degree distribution. This makes it possible to incorporate the typical correlation between x_v and $|\partial v(\pm)|$, capturing an even greater slice of SOL: they proved

$$\alpha_{\mathrm{s}} \geqslant 2^k \ln 2 - \tfrac{3}{2} \ln 2 - \epsilon_k \quad [15],$$

within $O(1)$ of the true threshold. A key idea was to identify a subset $\overline{\text{SOL}} \subseteq \text{SOL}$ of *judicious* configurations, where the variable spins x_v are non-trivially correlated with $|\partial v(\pm)|$, but are decoupled from the neighborhood structure beyond the degrees. Thus the judicious condition [15] is a significant generalization of the Achlioptas–Peres weighting scheme [4].

All lower bounds up to this point essentially applied the second moment method to restricted versions of SOL, and all remained slightly below the (conjectural) condensation threshold [27]. The 1RSB heuristic suggests that the count of *clusters* is well-concentrated while the count of assignments is not, indicating that one should instead apply the second moment method on the number of clusters. Among sparse CSPs expected to exhibit a non-trivial condensation regime, the first satisfiability lower bound to surpass the condensation barrier was obtained for random k-NAE-SAT [14], by the second moment method applied to a certain (rough) proxy for the number of clusters.

More recent work established exact satisfiability thresholds in random regular versions of k-NAE-SAT [18] and MAX-IND-SET [17], as well as a quasi-satisfiability threshold in symmetrized random regular k-SAT [13]. All these models exhibit non-trivial condensation regimes, but do not have the problem of graphical fluctuations, as the neighborhood of every variable looks like the (same) regular tree. These results relied on combinatorial models that that give extremely precise, yet reasonably tractable, encodings of clusters.

Coja-Oghlan and Panagiotou subsequently obtained the best random k-SAT lower bound prior to our work,

$$\alpha_{\mathrm{s}} \geqslant 2^k \ln 2 - (1 + \ln 2)/2 - \epsilon_k \quad [16],$$

matching the upper bound [26] up to ϵ_k error. This advance was significant in implementing the idea of counting clusters while accounting for degree fluctuations, which was done by extending the judicious condition [15] to the combinatorially encoded k-SAT clusters. Below we will review their program and explain the obstacles in attaining a sharp lower bound.

1.3 Proof overview

Recall Propn. 1 that the 1RSB threshold prediction $\alpha_\star$ is defined as a certain root of an explicit function $\Phi(\alpha)$. Our proof of Propn. 1 entails a rather involved recursive analysis, and we only prove $\alpha_\star$ is well-defined provided $k \geqslant k_0$.

Upper bound

Having established that $\alpha_\star$ is well-defined, the sharp upper bound $\alpha_{\mathrm{s}} \leqslant \alpha_\star$ is a straightforward consequence of interpolative free energy bounds [21, 36] for dilute spin-glass models (where "dilute" essentially means sparse). These results concern the *positive-temperature* version of SAT, where violated clauses are penalized rather than forbidden: each variable assignment $\underline{x} \in \{\pm\}^n$ is assigned weight $\exp\{-\beta H(\underline{x})\}$ where $H(\underline{x})$ counts the number of clauses unsatisfied by $\underline{x}$. Let

$$Z(\beta) \equiv \sum_{\underline{x} \in \{\pm\}^n} \exp\{-\beta H(\underline{x})\}$$

denote the total mass of the binary cube $\{\pm\}^n$ under this weighting, so that clearly $Z = |\text{SOL}| \leqslant Z(\beta)$ for any $\beta \geqslant 0$. For any fixed $\beta < \infty$, $\ln Z(\beta)$ is well-defined, and is well-concentrated about its mean by a standard argument (take the Doob martingale of $\ln Z(\beta)$ with respect to the clause-revealing filtration, and apply the Azuma–Hoeffding bound). For a class of models including positive-temperature k-SAT, it is shown [21, 36] that

$$\limsup_{n \to \infty} n^{-1} \mathbb{E} \ln Z(\beta) \leqslant \inf\{\Phi_\beta(\zeta) : \zeta \in \mathscr{P}\}$$

for a functional Φ_β defined on the space $\mathscr{P}$ of probability measures on $[0, 1]$. For $\alpha > \alpha_\star$, guided by the 1RSB predictions one can choose a particular $\zeta = \zeta_\beta$ and show

$$\lim_{\beta \to \infty} \Phi_\beta(\zeta_\beta) = -\infty.$$

Take β sufficiently large (depending on α) so that $\Phi_\beta(\zeta_\beta)$ is negative. Since $\ln Z(\beta)$ is well-concentrated, it follows that $Z(\beta) < 1$ with high probability. Since $Z \leqslant Z(\beta)$ and Z is integer-valued, we conclude $Z = 0$ with high probability, proving $\alpha_{\mathrm{s}} \leqslant \alpha_\star$.

Lower bound

The main content of this paper is the matching lower bound $\alpha_{\mathrm{s}} \geqslant \alpha_\star$. As the 1RSB heuristic suggests, we apply the second moment method on clusters of k-SAT solutions. We use a

particularly concise combinatorial encoding [16] which represents clusters as $\{\texttt{red}, \texttt{yellow}, \texttt{green}, \texttt{blue}\}$ colorings on the edges of $\mathscr{G} = (V, F, E)$, subject to certain local rules that we review below. The set of all clusters is then represented by the set $\mathrm{COL} \subseteq \{\texttt{red}, \texttt{yellow}, \texttt{blue}, \texttt{green}\}^E$ of valid colorings.

As mentioned above, a key step of [15, 16] is to condition on the degree sequence of a random k-SAT instance. This makes it possible to incorporate the correlation between the variables and degrees. In fact, variables are typically correlated not only with their degrees, but also with their neighborhood structures to arbitrary depths.[3] However, for the second moment method to go through, one must restrict to a particular subset $\overline{\mathrm{COL}} \subseteq \mathrm{COL}$ of judicious configurations, where the dependence on any local structure beyond the degrees is forcibly removed. This lack of dependence is highly atypical, meaning $\overline{\mathrm{COL}}$ captures only a tiny fraction of COL. This eventually incurs an ϵ_k loss in the lower bound on α_{s}.

In this work we condition on the empirical distribution of depth-R neighborhood types, which we regard as a generalized degree distribution. This type of conditioning was previously implemented in work of Bordenave and Caputo [9]. The plan is to identify, for each fixed R, a subset

$$\overline{\mathrm{COL}}_R \subseteq \mathrm{COL}$$

which captures the correlation between variables and their neighborhood structures up to depth R, but eliminates the dependence beyond depth R (judicious). The idea is that as R grows, we incorporate more and more of the correlation, thereby capturing larger and larger slices of COL. For each fixed R, we apply the second moment method on $|\overline{\mathrm{COL}}_R|$ to establish a satisfiability lower bound $\alpha_{\mathrm{s}} \geqslant \alpha_{\mathrm{lbd}}(R)$. We then show that $\alpha_{\mathrm{lbd}}(R) \to \alpha_\star$ in the limit $R \to \infty$, concluding the proof of Thm. 1.

The analysis of Coja-Oghlan and Panagiotou [16] is greatly simplified by an initial preprocessing step which removes an ϵ_k fraction of the most "atypical" variables from the graph. Note that as k grows, the k-SAT graph in the relevant regime $\alpha \asymp 2^k$ becomes more and more regular. By removing an ϵ_k fraction of variables, one can easily ensure that *all* remaining variables v have degrees $|\partial v(\pm)|$ very near to average:

$$\left| \frac{|\partial v(\pm)|}{k\alpha/2} - 1 \right| \leqslant \frac{k^{O(1)}}{2^{k/2}} \text{ for all } v \in V.$$

Many estimates in [16] rely on this explicit control. Clearly, in order to achieve a sharp threshold, we must remove a fraction of vertices tending to zero as $R \to \infty$ — meaning we cannot hope to avoid including increasingly "bad" vertices as R grows. Instead, we carry out a preprocessing step where the goal is only to ensure that bad vertices are surrounded by a sufficient buffer of nice vertices which will help to enforce the desired behavior in the second moment. This preprocessing step is rather involved, and will be described in more detail below.

[3] In random k-SAT, variables are correlated with neighborhood structures to arbitrary depth in both the original model and the cluster model, as can be seen in [15] and [16] respectively. In random k-NAE-SAT, thanks to the additional symmetry, variables are uncorrelated with neighborhood structures under the original model [3]. However, the NAE-SAT cluster model will exhibit correlations to arbitrary depth, similarly as in SAT. We expect that the methods of this paper can be applied to obtain the sharp satisfiability threshold in random k-NAE-SAT.

Having completed the preprocessing, we condition on the empirical distribution $\mathscr{D}$ of depth-R neighborhood types in the processed graph. We then identify a subset $\overline{\mathrm{COL}}_R \subseteq \mathrm{COL}$ of good colorings of the processed graph, and perform second moment method on the random variable $\overline{\boldsymbol{Z}} = |\overline{\mathrm{COL}}_R|$: with $\mathbb{E}_{\mathscr{D}}$ denoting expectation conditional on $\mathscr{D}$, we will show

$$\mathbb{E}_{\mathscr{D}}[\overline{\boldsymbol{Z}}^2] \lesssim (\mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}})^2. \tag{6}$$

The second moment can be cast as an optimization problem over a vector ω that represents the empirical distribution of edge colors for a typical pair $(\underline{\sigma}^1, \underline{\sigma}^2)$ of independent uniform samples from $\overline{\mathrm{COL}}_R$. The entries of ω are indexed by edge types; and each entry of ω is a probability distribution over $\{\texttt{red}, \texttt{yellow}, \texttt{green}, \texttt{blue}\}^2$ which gives the empirical distribution of colors for that edge type. The second moment bound (6) amounts to showing that the optimal ω has each entry equal to product measure.

A central idea in this paper is to update ω in *blocks* that correspond to trees of bounded (though growing with R) depth. More precisely, the block will correspond to all induced subgraphs of the (processed) k-SAT graph that are isomorphic to a given tree T of depth $\leqslant R$. We consider the optimization problem over the entries of ω corresponding to the internal edges of the tree, keeping all other entries fixed. This reduces an optimization problem on large finite graphs to an optimization problem on finite trees subject to certain boundary conditions. We then carry out the tree optimization by a system of weights that act as Lagrange multipliers for the boundary constraints. The preprocessing step was specifically tailored for this tree optimization problem.

In the remainder of this extended abstract we describe in further detail some of the main innovations in our proof. The full version of the paper has been made available online at http://arxiv.org/abs/1411.0650.

2. FROZEN MODEL AND COLOR MODEL

We first describe the combinatorial encoding of clusters. Recall $\mathscr{G} \equiv (V, F, E)$ is the bipartite factor graph representing the k-SAT instance. Write ∂ for the neighbors of a vertex *with multiplicity*, and δ for the incident edges. For a variable $v \in V$ we regard $\partial v, \delta v$ as unordered multisets, while for a clause $a \in F$ we regard $\partial a, \delta a$ as ordered tuples: that is, each edge $(av) \in E$ comes with a label $j(v; a) \in [k]$, indicating the position of v in ∂a.

2.1 Frozen model

Recall a k-SAT solution is a configuration $\underline{x} \in \{\texttt{+}, \texttt{-}\}^V$ such that every clause $a \in F$ is satisfied, meaning $(\mathtt{L}_{av} x_v)_{v \in \partial a}$ is not identically $\texttt{-}$. We now introduce a new spin $\texttt{f} \equiv \texttt{free}$ to encode the k-SAT solution clusters:

Definition 1. On a given k-SAT instance $\mathscr{G} = (V, F, E)$, a *frozen configuration* is a vector $\underline{x} \in \{\texttt{+}, \texttt{-}, \texttt{f}\}^V$ such that

(i) each clause $a \in F$ is satisfied, meaning $(\mathtt{L}_{av} x_v)_{v \in \partial a}$ is not identically $\texttt{-}$; and

(ii) for each variable $v \in V$, $x_v \neq \texttt{f}$ if and only if v is *forced*, meaning that for some $a \in \partial v$, the product of $\mathtt{L}_{au}$ and x_u is $\texttt{-}$ for all $u \in \partial a \backslash v$.

In the above definition and throughout what follows, we adopt the convention that the product of $\texttt{+}$ with $\texttt{f}$, or the product of $\texttt{-}$ with $\texttt{f}$, is $\texttt{f}$.

2.2 Tree recursions for frozen model

Let T be a bipartite factor graph which is a finite tree, such that the leaf vertices ∂T are variables. (Eventually we will take T to be a subgraph of $\mathscr{G}$, given by the r-neighborhood $B_r(v)$ of a variable $v \in V$.) We shall often consider the frozen model on finite trees of this form with i.i.d. rigid balanced input at the boundary. Roughly speaking, this will be the measure on frozen configurations of T induced by setting all leaf variables to be forced ("rigid") to `+` or `-`, independently and uniformly ("i.i.d. balanced").

More formally, for any variable $v \in T$ with neighboring clause $a \in T$, let T_{va} be the component of $T \backslash a$ containing v (T_{va} includes the half-edge in δv that was previously matched to a half-edge in δa). Define likewise T_{av} to be the component of $T \backslash v$ containing a. We shall write

$$\eta_{va} \equiv \text{"probability for } v \text{ to negate } \mathtt{L}_{av}\text{, under the frozen model on } T_{va} \text{ with i.i.d. rigid balanced input on } T_{va} \cap \partial T\text{."}$$

Explicitly, the η_{va} are defined as follows. For a leaf variable $v \in \partial T$ with (unique) neighboring clause $a \in T$, set $\eta_{va} = 1/2$. The interpretation is that T is a subgraph of some large graph $\mathscr{G}$, and the rest of the graph $\mathscr{G} \backslash T$ forces v to `±` or `-`. The forcing is independent of the sign $\mathtt{L}_{av}$, and so $\eta_{va} = 1/2$. We then calculate η_{va} at internal edges by recursing up the tree, treating the "branches" $(T_{bv})_{b \in \partial v \backslash a}$ as independent inputs. For variable v with neighboring clause a, write

$$\partial v(\mathtt{+}a) \equiv \{b \in \partial v \backslash a : \mathtt{L}_{bv} = \mathtt{+L}_{av}\},$$
$$\partial v(\mathtt{-}a) \equiv \{b \in \partial v \backslash a : \mathtt{L}_{bv} = \mathtt{-L}_{av}\}.$$

On the finite tree, η_{va} is expressed in terms of the η_{wb} ($b \in \partial v \backslash a$, $w \in \partial b \backslash v$) by the recursive relation

$$\eta_{va} = \frac{\Pi^{+}_{va}(1 - \Pi^{-}_{va})}{\Pi^{+}_{va} + \Pi^{-}_{va} - \Pi^{+}_{va}\Pi^{-}_{va}}, \text{ where}$$
$$\Pi^{\mathtt{\pm}}_{va} \equiv \prod_{b \in \partial v(\mathtt{\pm}a)} \Big(1 - \prod_{u \in \partial v \backslash b} \eta_{ub}\Big)$$

— note the clear resemblance with the recursion of §1.

2.3 Color model (warning propagation)

It is known that the frozen model can be conveniently re-expressed as a *Gibbs measure* (also termed factor model or Markov random field) with spins on variable-clause edges, subject to constraints defined by clauses and variables. This sometimes goes by the name of the "warning propagation" model; the reader is referred to [33, 10, 28, 29, 30] for more background. We use the "color model" [16], which is an efficient projection of the standard warning propagation model. In this model, an edge $e = (av)$ ($a \in F$, $v \in V$) is colored

`red`	if v is forced to satisfy a;
`blue`	if v satisfies a but is not forced by it;
`yellow`	if v is forced to negate a (by some other clause $b \in \partial v \backslash a$);
`green`	if v is free.

Given $\mathscr{G} = (V, F, E)$ there is a bijective correspondence

$$\{\text{frozen configurations } \underline{x} \in \{\mathtt{+}, \mathtt{-}, \mathtt{f}\}^V\} \longleftrightarrow \{\text{valid colorings } \underline{\sigma} \in \{\mathtt{red}, \mathtt{yellow}, \mathtt{green}, \mathtt{blue}\}^E\}. \quad (7)$$

The uniform measure ν on valid colorings of $\mathscr{G}$ can be written as a Gibbs measure

$$\nu(\underline{\sigma}) = \frac{1}{Z} \prod_{v \in V} \varphi_v(\underline{\sigma}_{\delta v}) \prod_{a \in F} \varphi_a(\underline{\sigma}_{\delta a})$$

where φ_x ($x \in V \cup F$) is an indicator function whose definition involves only δx. If we define tree recursions for this model analogously to the frozen model recursions (§2.2), we arrive precisely at the standard *belief propagation* (BP) recursions for this Gibbs measure. These recursions are expressed in terms of *messages* $\dot{q}, \hat{q}$ which are probability measures over the set $\{\mathtt{red}, \mathtt{yellow}, \mathtt{blue}, \mathtt{green}\}$ of edge spins:

$\dot{q}_{va}$ = message v to a, "law of σ_{av} in absence of a";
$\hat{q}_{av}$ = message a to v, "law of σ_{av} in absence of v".

We omit the recursions here as their derivation is standard. Thanks to the bijection (7), the frozen model tree recursions can be retrieved as a special case of the color model belief propagation recursions. This is detailed in the appendix; we summarize here that on a finite tree T as in §2.2, there is a set of messages $\dot{q}_{va}, \hat{q}_{va}$ that are given by a simple transformation of the η_{va}, and solve the BP recursions on T. It turns out that these particular messages $\dot{q}_{va}$ and $\hat{q}_{av}$ are functions respectively of T_{va} and T_{av}, so we will denote

$$\dot{q}_{va} \equiv \dot{\boldsymbol{q}}(T_{va}),\ \hat{q}_{av} \equiv \hat{\boldsymbol{q}}(T_{av}).$$

3. PREPROCESSING

We present here a simplified version of our preprocessing step that highlights the key features.

3.1 Simple types and niceness

Our basic definitions of neighborhood types are as follows. Fix a large integer R, and condition on the event that the graph has girth $\geqslant 2R$ (which occurs with probability $\asymp_R 1$).

Definition 2. In a graph $\mathscr{G} = (V, F, E)$, the *simple type* t_e of a clause-variable edge $e \equiv (av) \equiv (va) \in E$ is the isomorphism class of $(B_R(v), e)$, the R-neighborhood around v rooted at edge e.[4] We write $j(t_e) \equiv j(v; a)$ to indicate the position of the variable in the clause.

The *simple type* of a vertex $x \in V \cup F$ is the multi-set of all incident edge types $\{t_e : e \in \delta x\}$. This has a slightly different representation according to whether x is a clause or a variable:

1. If $x \in F$, its simple type L_x has no repeated elements, since each edge $e \in \delta x$ has a distinct index $j(t_e) \in [k]$. Thus L_x is equivalently represented as the ordered k-tuple $(L_x(1), \ldots, L_x(k))$ where $L_x(j)$ is the type of the j-th edge in δx.

2. If $x \in V$, its simple type T_x may have repeated elements. It is equivalently represented as the isomorphism class of $B_R(v)$ regarded as a graph rooted at v.

In what follows, it will be convenient to denote

$$R \equiv 10^2 R' \equiv 10^4 r. \quad (8)$$

We can assume that r is a large positive integer.

[4]The edge-rooted graphs (T_i, e_i), $i = 1, 2$, are isomorphic if there is a bijective graph homomorphism $\iota : T_1 \to T_2$ with $e_1 \mapsto e_2$, preserving edge labels $\mathtt{L}_{av} \in \{\mathtt{+}, \mathtt{-}\}$ and $j(v; a) \in [k]$.

Definition 3. For $v \in V$ let $T = B_r(v)$. For each $a \in \partial v$, the *canonical messages* on edge $e = (av)$ are defined by

$$\star\dot{q}_{va} \equiv \dot{\boldsymbol{q}}(T_{va}),\ \star\hat{q}_{av} \equiv \hat{\boldsymbol{q}}(T_{av}) \text{ for each } a \in \partial v.$$

The *canonical marginal* on e is defined by

$$\star\pi_{av}(\sigma) = \frac{\star\dot{q}_{va}(\sigma)\star\hat{q}_{av}(\sigma)}{\sum_\tau \star\dot{q}_{va}(\tau)\star\hat{q}_{av}(\tau)}.$$

Each of $\star\pi_{av}, \star\dot{q}_{va}, \star\hat{q}_{av}$ is a probability distribution over the colors $\{\texttt{red}, \texttt{yellow}, \texttt{blue}, \texttt{green}\}$.

For α near $\alpha_{\rm s}$, each cluster has only a small number of free variables. It is reasonable to expect that as R increases, $\star\pi$ will be an increasingly good approximation of the true edge marginals, provided the marginal does not depend too sensitively on the structure of the faraway (beyond R) levels. We quantify this by a *stability* property which is explained in the appendix.

Definition 4. A variable v is *nice* if it is stable (appendix), has degrees satisfying $||\partial v^{\pm}| - 2^{k-1}k\ln 2| \leqslant 2^{2k/3}$, and has canonical messages $\star\dot{q}, \star\hat{q}$ satisfying

$$\left\{\begin{array}{r} |\star\hat{q}_{av}(\texttt{yellow}) - \frac{1}{3}[1 - \frac{2}{3}(\frac{1}{2})^k]| \leqslant 2^{-k/8}, \\ |\star\dot{q}_{va}(\texttt{yellow}) - \frac{1}{3}[1 - \frac{2}{3}(\frac{1}{2})^k]| \leqslant 2^{-k/8}, \\ 2^k|\star\hat{q}_{av}(\texttt{red}) - \frac{1}{3}(\frac{1}{2})^{k-1}| \leqslant 2^{-k/8}, \\ 2^k|\star\dot{q}_{va}(\texttt{green}) - \frac{1}{3}(\frac{1}{2})^k| \leqslant 2^{-k/8}. \end{array}\right\} \text{ for all } a \in \partial v.$$

(The values specified above are roughly typical for $\dot{q}, \hat{q}$.)

3.2 Bootstrap percolation of defects

In a general bipartite factor graph $\mathscr{G} = (V, F, E)$, given some subset of variables $D_0 \subseteq V$, for $t \geqslant 1$ set $D_t \supseteq D_{t-1}$ to be the union of D_{t-1} together with all variables having at least two neighboring variables in $D_{t-1} \cap V$. The set

$$\mathsf{BSP}(D_0; \mathscr{G}) \equiv D_\infty = \text{union of } (D_t)_{t \geqslant 0} \tag{9}$$

will be termed the *bootstrap percolation* of D_0 in $\mathscr{G}$. We identify defective regions of the graph by a certain "localized" bootstrap percolation:

Definition 5. Let $D_* \equiv \{v \in V : v \text{ is not nice}\}$. Let κ be a large absolute constant, and let D_0 be the κ-neighborhood of D_* (the union of κ-neighborhoods of all variables in D_*). A variable v is *defective* if $v \in \mathsf{BSP}(D_0 \cap B_{R'/2}(v), B_{R'/2}(v))$.

Importantly, whether a variable is defective is determined by its R'-neighborhood — that is, being defective is a local property. By construction, each defect has at its boundary a buffer of nice variables of depth at least κ. A clause is considered part of a defect if and only if all its incident variables belong to the defect — otherwise, it will follow from our preprocessing procedure that for each remaining clause in the processed graph at most one incident variable can belong to a defect, and in this case the clause is not considered part of any defect.

3.3 Containment and enclosures

For variables $u, v \in V$, let $\mathfrak{B}(u, v)$ count the defective variables on the shortest path from u to v (inclusive), while $\mathfrak{H}(u, v)$ counts the non-defective variables. The following definition is at the heart of our second moment analysis.

Definition 6. Let δ^* be a small absolute constant. Define the *containment radius* of variable v to be

$$\mathrm{rad}(v) \equiv \min\left\{t \geqslant 0 : \sum_{s \leqslant t < 2R'} \mathfrak{R}_s(v)\right\} \leqslant 1/4, \text{ where} \tag{10}$$

$$\mathfrak{R}_s(v) \equiv \sum_{u : d(u,v) = s} \frac{\exp\{k(\delta^*)^{-1}\mathfrak{B}(u, v)\}}{\exp\{(k\ln 2)(1 + \delta^*)\mathfrak{H}(u, v)\}}$$

In particular, if v is defective then $\mathrm{rad}(v) > 0$. We say v is *self-contained* if

$$\mathrm{rad}(u) \leqslant d(u, v) \text{ for all } u \text{ with } d(u, v) \leqslant R';$$

this is a local property that can be determined from the $4R'$-neighborhood of the variable.

The central aim of preprocessing is to ensure that it is possible to carve up the graph into "enclosures": the formal definition is given below, but roughly speaking these will be regions of diameter at most R' such that every variable in a given enclosure has containment radius less than or equal to its minimal distance from the enclosure boundary (in particular, all the boundary variables must be self-contained). The following definitions are used to carve up the graph:

Definition 7. We say that a variable v is *perfect* if it is orderly and self-contained. We say v is *fair* if it is stable; its $5R'$-neighborhood contains no more than $\exp\{k^2(5R')\}$ variables; and lastly it does not belong in any length-R' path that fails to contain at least one perfect variable. Whether a variable is fair can be determined from its $5R'$-neighborhood.

3.4 Preprocessing algorithm

The following is a variant of the bootstrap percolation process defined in (9). Recall from (8) that $R = 10^2 R'$.

Definition 8. In a graph $\mathscr{G} = (V, F, E)$, let

$$\mathscr{A}(\mathscr{G}) \equiv \left\{\begin{array}{c} v \in V \text{ such that } B_{3R/10}(v) \text{ contains} \\ \geqslant 2 \text{ clauses of degree } k-1, \text{ or} \\ \geqslant 1 \text{ clause of degree } \leqslant k-2. \end{array}\right\}.$$

Given an initial subset of variables $A \subseteq V$, let

$$\begin{aligned} {}_0\mathscr{G}_A &\equiv \mathscr{G} \Big\backslash \Big\{B_R(v) : v \in A\Big\}, \text{ then} \\ {}_{t+1}\mathscr{G}_A &\equiv {}_t\mathscr{G}_A \Big\backslash \Big\{{}_tB_R(v) : v \in \mathscr{A}({}_t\mathscr{G}_A)\Big\} \text{ for } t \geqslant 0, \end{aligned}$$

where ${}_tB_R(v)$ is the R-neighborhood of v with respect to the graph ${}_t\mathscr{G}_A$. When $\mathscr{A}({}_t\mathscr{G}_A) = \varnothing$ the process has reached the final graph ${}_t\mathscr{G}_A \equiv {}_\infty\mathscr{G}_A$. Let $\mathsf{BSP}'(A; \mathscr{G})$ denote the set of all variables removed by this procedure.

Preprocessing algorithm on $\mathscr{G}$:
Let $A \subseteq V$ be the non-fair variables (Defn. 7).
Delete $\mathsf{BSP}'(A; \mathscr{G})$ and output ${}_{\rm pr}\mathscr{G} \equiv {}_\infty\mathscr{G}_A$.

Definition 9. Any perfect variable $v \in {}_{\rm pr}V$ constitutes a *singleton enclosure.* A *compound enclosure* is a subgraph of ${}_{\rm pr}\mathscr{G}$ induced by a subset of variables $U^\circ \cup U^\partial \subseteq {}_{\rm pr}V$, where U° is a (nonempty, maximal) connected component of non-perfect variables, and $U^\partial \equiv \{u \in {}_{\rm pr}V : d(u, U^\circ) = 1\}$ is its external boundary consisting of perfect variables.

A connected component in ${}_{\rm pr}\mathscr{G}$ of non-perfect variables must have diameter at most R', which means that every compound enclosure must be a tree. In fact a vast majority of variables will be singleton enclosures.

Definition 10. For each element (variable, clause, or edge) in $\mathscr{G}$, the *simple total type* records the simple type (Defn. 2) of that element both before and after preprocessing.

Definition 11. For each element (variable, clause, or edge) in a compound enclosure U, its *compound type* records the graph structure of U with the position of the element marked, together with the simple total type of every element in U. In particular, different elements appearing in the same compound enclosure must have different compound types, even if their simple total types match.

Definition 12. The *total type* of an element in $\mathscr{G}$ is defined to be its simple total type if the element does not lie in a compound enclosure. If the element lies in a compound enclosure, its total type is defined to be its simple type (defined with respect to the initial graph) together with its compound type (defined with respect to the processed graph). We use $\boldsymbol{T}$, $\boldsymbol{L}$, and $\boldsymbol{t}$ to denote the total types of variables, clauses, and edges respectively.

With some abuse of notation, we write $\mathscr{G} \equiv (V, F, E)$ from now on for the processed graph labelled with all total types. We denote $n \equiv |V|$ and $m \equiv |F|$ (where these can be smaller than the original values of n, m).

Definition 13. The *processed degree distribution* $\mathscr{D}$ is the empirical profile of total types in the graph $\mathscr{G}$: $\mathscr{D} \equiv (\hat{\mathscr{D}}, \hat{\hat{\mathscr{D}}})$ with $\hat{\mathscr{D}}$ (resp. $\hat{\hat{\mathscr{D}}}$) the empirical distribution of variable (resp. clause) total types. The empirical distribution $\overline{\mathscr{D}}$ of edge total types can be computed as a marginal of both $\hat{\mathscr{D}}, \hat{\hat{\mathscr{D}}}$.

The main result of our preprocessing analysis is as follows:

PROPOSITION 2. *The processed graph conditional on $\mathscr{D}$ is uniformly distributed over the set of all graphs consistent with $\mathscr{D}$. The following hold with high probability:*

(a) Preprocessing removes $\leqslant n/\exp\{2^{ck}R\}$ variables, for an absolute constant $c > 0$.

(b) Every total type present in the processed graph appears $\geqslant nc_1$ times, for a constant $c_1(k, R) > 0$.

Our definition of total type was chosen to guarantee the uniformity — it allows us now to sample $\mathscr{G}$ conditional on $\mathscr{D}$ using a generalization ([9]) of the standard configuration model for graphs with given degree sequence. For expository purposes, we have omitted from this abstract some more technical components of the preprocessing, including steps taken to ensure Propn. 2b.

4. PROOF OUTLINE

We can now supply a more detailed overview of our proof. First we analyze the distributional recursion and establish Propn. 1. Some of the estimates obtained in this analysis are applied to study the preprocessing algorithm and to prove Propn. 2. These estimates are technically rather challenging, and relied on the assumption of large k.

We show moreover that valid colorings of the processed graph can be mapped to valid k-SAT solutions on the original graph. Therefore, to prove our main result it suffices to establish the existence with high probability of valid colorings of the processed graph. Further, by Friedgut's theorem [22] it suffices to show existence with probability non-vanishing in the limit $n \to \infty$. This will be done by the second moment method applied to a particular subset of good colorings, as we now define.

4.1 Separable judicious colorings

Recall $\mathscr{G} \equiv (V, F, E)$ now refers to the processed graph, with $n = |V|$, $m = |F|$, and degree distribution (Defn. 13). Given a valid coloring $\underline{\sigma}$, let π, ω be defined by

$$\begin{aligned} n\overline{\mathscr{D}}(\boldsymbol{t})\pi_{\boldsymbol{t}}(\sigma) &\equiv |\{(av) \in E : \boldsymbol{t}_{av} = \boldsymbol{t}, \sigma_{av} = \sigma\}|; \\ m\hat{\mathscr{D}}(\boldsymbol{L})\omega_{\boldsymbol{L},j}(\sigma) &\equiv \left| \left\{ \begin{matrix} (av) \in E : \boldsymbol{L}_a = \boldsymbol{L}, \\ j(v;a) = j, \sigma_{av} = \sigma \end{matrix} \right\} \right|. \end{aligned} \tag{11}$$

Both π and ω are functions of the given coloring $\underline{\sigma}$; moreover, π is a linear function of ω. The following two definitions are adapted from [15, 16]:

Definition 14. A valid coloring $\underline{\sigma}$ on a processed graph $\mathscr{G}$ is *self-judicious* if $\omega_{\boldsymbol{L},j}$ depends only on $\boldsymbol{L}(j)$, that is to say,

$$\omega_{\boldsymbol{L},j} = \pi_{\boldsymbol{L}(j)} \text{ for all } \boldsymbol{L}, j.$$

The coloring $\underline{\sigma}$ is termed *judicious* if furthermore π agrees (up to rounding) with the canonical edge marginal $_\star\pi$ based on the variable r-neighborhood (Defn. 3). Note judicious is a stronger condition than self-judicious.

Definition 15. For a judicious coloring $\underline{\sigma}$, let $\underline{x}$ be the frozen configuration corresponding via (7) to $\underline{\sigma}$. We say $\underline{\sigma}$ is *separable* if there are $\leqslant \exp\{(\ln n)^5\}$ judicious colorings $\underline{\sigma}'$ such that the Hamming distance between $\underline{x}$ and $\underline{x}'$ lies outside the interval $[(1 - k^4 2^{-k/2})n/2, (1 + k^4 2^{-k/2})n/2]$.

Recalling the discussion of §1.3, we now set

$$\overline{\mathrm{COL}}_R \equiv \{\text{judicious separable colorings } \underline{\sigma}\}.$$

We will perform the second moment method on $\overline{\boldsymbol{Z}} \equiv |\overline{\mathrm{COL}}_R|$, conditional on $\mathscr{D}$:

THEOREM 2. *There is a constant $C(k, R)$ such that*

$$\mathbb{E}_{\mathscr{D}}[\overline{\boldsymbol{Z}}^2] \leqslant C(\mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}})^2 + e^{o(n)}\mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}} \tag{12}$$

with high probability over the random degree sequence $\mathscr{D}$.

The rightmost term in (12) is the contribution from pairs of colorings $\underline{\sigma}, \underline{\sigma}'$ with high correlation, which is directly controlled by the separability condition (Defn 15). By Thm. 2, if $\mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}}$ is exponentially large in n, then we have the bound (6). This implies $\overline{\boldsymbol{Z}} > 0$ with non-vanishing probability as $n \to \infty$, which in turn implies a satisfiability lower bound as explained above: $\alpha_{\mathrm{s}} \geqslant \alpha_{\mathrm{lbd}}(R)$ where

$$\alpha_{\mathrm{lbd}}(R) = \sup\{\alpha : \liminf n^{-1} \ln \mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}} > 0\}.$$

To prove $\alpha_{\mathrm{lbd}}(R) \to \alpha_\star$, we show that most judicious configurations are separable in the sense that, with high probability over $\mathscr{D}$, we have $\mathbb{E}_{\mathscr{D}}\overline{\boldsymbol{Z}} = [1 - o_n(1)]\mathbb{E}_{\mathscr{D}}\boldsymbol{Z}$ where

$$\overline{\boldsymbol{Z}} \leqslant \boldsymbol{Z} \equiv \#\text{ judicious colorings } \underline{\sigma}.$$

The advantage of working with $\boldsymbol{Z}$ rather than $\overline{\boldsymbol{Z}}$ is that $\boldsymbol{Z}$ is amenable to moment calculations under the generalized configuration model mentioned above. We can express $\mathbb{E}_{\mathscr{D}}\boldsymbol{Z}$ as a sum over products of multinomial coefficients, and thereby show that with high probability (using also Propn. 2a),

$$\mathbb{E}_{\mathscr{D}}\boldsymbol{Z} \geqslant \exp\{n[\Phi(\alpha) - \epsilon_R]\}$$

for $\lim_R \epsilon_R = 0$. Since Φ is decreasing (Propn. 1), it follows that $\lim_{R\to\infty} \alpha_{\mathrm{lbd}}(R) = \alpha_\star$ as required.

Outside of the highly-correlated regime captured by the second term on the right-hand side of (12), we shall drop the

separability condition and work with judicious colorings. We decompose

$$\mathbb{E}[\boldsymbol{Z}^2] = \sum_{\omega} \mathbb{E}[\boldsymbol{Z}^2(\omega)]$$

with $\boldsymbol{Z}^2(\omega)$ the contribution to $\boldsymbol{Z}^2$ from pairs $\underline{\sigma} \equiv (\underline{\sigma}^1, \underline{\sigma}^2)$ with edge empirical measure ω — ω is defined analogously as before, but now each entry is a probability measure over $\{\texttt{red}, \texttt{yellow}, \texttt{green}, \texttt{blue}\}^2$.

For each edge $e = (av)$ with $\boldsymbol{L}_a = \boldsymbol{L}$ and $j(v;a) = j$, write $\omega_e \equiv \omega_{\boldsymbol{L},j}$. For a measure p on $\{\texttt{red}, \texttt{yellow}, \texttt{blue}, \texttt{green}\}^2$, write p^i $(i = 1, 2)$ for its single-copy marginals. As $\boldsymbol{Z}$ counts only judicious colorings, any empirical measure ω giving a positive contribution to $\boldsymbol{Z}^2$ must satisfy

$$\omega_e^1 = {}_\star\pi_e = \omega_e^2 \text{ for all } e \in E, \tag{13}$$

where ${}_\star\pi_e$ is the canonical edge marginal defined according to the r-neighborhood of the incident variable v (Defn. 3). We refer to measures with property (13) as *judicious*. From the preceding discussion, we are interested in the second moment contribution outside the highly-correlated regime, that is, we wish to estimate

$$\mathbb{E}\boldsymbol{Z}^2(\boldsymbol{I}) = \sum_{\omega \in \boldsymbol{I}} \mathbb{E}\boldsymbol{Z}^2(\omega)$$

where $\boldsymbol{I}$ is a certain neighborhood of $\omega^\otimes \equiv {}_\star\pi \otimes {}_\star\pi$. We show that $\mathbb{E}[\boldsymbol{Z}^2(\omega)]$ is uniquely maximized over all $\omega \in \boldsymbol{I}$ precisely at $\omega^\otimes$, which immediately gives (with $\dot{\leqslant}$ meaning $\leqslant$ up to factors polynomial in n)

$$\mathbb{E}[\boldsymbol{Z}^2(\boldsymbol{I})] \mathrel{\dot{\leqslant}} (\mathbb{E}\boldsymbol{Z})^2.$$

We subsequently remove the polynomial factor by proving that $\mathbb{E}[\boldsymbol{Z}^2(\omega)]$ has the appropriate decay in a neighborhood of $\omega^\otimes$ to yield $\mathbb{E}[\boldsymbol{Z}^2(\boldsymbol{I})] \lesssim (\mathbb{E}\boldsymbol{Z})^2$, thereby concluding the proof of Thm. 2.

4.2 Single-site and block updates

To prove $\omega^{\mathrm{opt}} = \omega^\otimes$, assume not. We will take the entry of ω^{opt} furthest (by some metric) from the corresponding entry of $\omega^\otimes$, and re-optimize in this entry to obtain ω' which (i) is closer to $\omega^\otimes$ than ω^{opt}, and (ii) gives a larger contribution to the partition function than ω^{opt}. This contradicts the optimality of ω^{opt}, proving our claim.

The main work of implementing this strategy is in defining and analyzing our update procedure. A single-site update re-optimizes the marginal for a single edge type, keeping all the other edge marginals fixed. For nice types this update does indeed contract towards the product measure, but for non-nice types it may not. We therefore supplement the single-site updates with *block* updates where we re-optimize over the edge marginals for *all* types appearing within a compound enclosure, while keeping fixed the marginals in the rest of the graph. The definition of enclosure was tailored to ensure that these block updates contract towards $\omega^\otimes$.

4.3 Reduction to optimization on trees

The block update is more complicated than the single-site update, but a key observation is that the optimization factorizes in a simple manner due to our notion (Defn. 11) of compound types. Fix a tree T that is fully contained within some compound enclosure. There are $\mathtt{N}$ disjoint copies of the enclosure in the graph, hence $\mathtt{N}$ disjoint copies of T (where $\mathtt{N}$ is large by Propn. 2b). Let

$$\omega = \begin{pmatrix} \omega_{\delta T} \\ \omega_{\mathrm{int}} \\ \omega_{\mathrm{ext}} \end{pmatrix} = \begin{pmatrix} \omega_{\boldsymbol{L},j} : \boldsymbol{L}_j \text{ appears in leaves } \delta T \text{ of } T \\ \omega_{\boldsymbol{L},j} : \boldsymbol{L}_j \text{ appears in } T^\circ = T \backslash \delta T \\ \omega_{\boldsymbol{L},j} : \boldsymbol{L}_j \text{ does not appear in } T \end{pmatrix}$$

where δT denotes the leaf edges of T. Given $\omega_{\delta T}$, the configuration model within the copies of T is independent of the configuration model in the remainder of the graph: that is to say, we have the factorization

$$\mathbb{E}[\boldsymbol{Z}^2(\omega)] = \mathbb{E}[\boldsymbol{Z}^2_{\mathrm{int}}(\omega_{\mathrm{int}}, \omega_{\delta T})]\mathbb{E}[\boldsymbol{Z}^2_{\mathrm{ext}}(\omega_{\delta T}, \omega_{\mathrm{ext}})] \tag{14}$$

where $\boldsymbol{Z}^2_{\mathrm{int}}(\omega_{\mathrm{int}}, \omega_{\delta T})$ is the partition function on $\mathtt{N}$ disjoint copies of T subject to empirical measure $(\omega_{\mathrm{int}}, \omega_{\delta T})$, and $\boldsymbol{Z}^2_{\mathrm{ext}}(\omega_{\delta T}, \omega_{\mathrm{ext}})$ is the partition function on the graph with all the copies of T removed (leaving behind dangling edges). Let us emphasize again that the above factorization relies crucially on our definition of compound types. Then, with $\doteq$ denoting equality up to polynomial factors,

$$\begin{aligned}
\max_{\omega_{\mathrm{int}}} \mathbb{E}[\boldsymbol{Z}^2(\omega_{\mathrm{int}}, \omega_{\delta T}, \omega_{\mathrm{ext}})] &\doteq \sum_{\omega_{\mathrm{int}}} \mathbb{E}[\boldsymbol{Z}^2(\omega_{\mathrm{int}}, \omega_{\delta T}, \omega_{\mathrm{ext}})] \\
&\doteq \Big(\sum_{\omega_{\mathrm{int}}} \mathbb{E}[\boldsymbol{Z}^2_{\mathrm{int}}(\omega_{\mathrm{int}}, \omega_{\delta T})] \Big) \mathbb{E}[\boldsymbol{Z}^2_{\mathrm{ext}}(\omega_{\delta T}, \omega_{\mathrm{ext}})] \\
&\doteq \Big(\max_{\omega_{\mathrm{int}}} \mathbb{E}[\boldsymbol{Z}^2_{\mathrm{int}}(\omega_{\mathrm{int}}, \omega_{\delta T})] \Big) \mathbb{E}[\boldsymbol{Z}^2_{\mathrm{ext}}(\omega_{\delta T}, \omega_{\mathrm{ext}})].
\end{aligned}$$

That is to say, optimizing the partition function on the graph subject to fixed empirical measures outside a tree T can be reduced to optimizing the partition function on the graph on $\mathtt{N}$ disjoint copies of the tree T subject to fixed empirical measures at the leaves ∂T. As we now proceed to explain, the latter optimization problem can be analyzed by belief propagation in weighted models.

To solve the constrained optimization problem of maximizing the tree partition function subject to given marginals at the leaves, we introduce a system of Lagrange multipliers to arrive at an unconstrained optimization problem. The Lagrange multipliers are implemented by adding weights to our original unweighted model of judicious colorings on trees. We give a direct construction of these weights which allow us to estimate their sizes. It is well understood how to use belief propagation to solve the unconstrained optimization in the weighted tree model. We analyze the BP solution to show that the root marginal contracts towards the desired product measure.

This gives Thm. 2 with $n^{O(1)}$ in place of the constant C. Another key result of our reweighting approach is that the contraction can be used to deduce that $n^{-1} \ln \mathbb{E}[\boldsymbol{Z}^2(\omega)]$ has uniformly negative-definite Hessian at ω^{opt}, simplifying a step that is often very technically challenging. This allows us to improve the $n^{O(1)}$ to the required constant C, yielding Thm. 2. The main result Thm. 1 then follows by combining with Friedgut's sharp threshold theorem [22].

Acknowledgements

We wish to thank Amir Dembo, Elchanan Mossel, Andrea Montanari, and Lenka Zdeborová for many helpful conversations. We are grateful for the hospitality of the Theory Group at Microsoft Research Redmond where much of the key work was done. Our research was supported in part by NSF grant DMS-1313596 (J.D.); NSF grants DMS-1208338 and DMS-1352013, and Sloan Research Fellowship (A.S.); and NSF MSPRF grant DMS-1401123 (N.S.).

5. REFERENCES

[1] 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. ACM–SIAM, 2001.

[2] D. Achlioptas and A. Coja-Oghlan. Algorithmic barriers from phase transitions. In *Proc. 49th FOCS*, pages 793–802. IEEE, 2008.

[3] D. Achlioptas and C. Moore. Random k-SAT: two moments suffice to cross a sharp threshold. *SIAM J. Comput.*, 36(3):740–762 (electronic), 2006.

[4] D. Achlioptas and Y. Peres. The threshold for random k-SAT is $2^k \ln 2 - O(k)$. *J. Amer. Math. Soc.*, 17(4):947–973 (electronic), 2004.

[5] D. Aldous and R. Lyons. Processes on unimodular random networks. *Electron. J. Probab.*, 12(54):1454–1508, 2007.

[6] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann, and D. Vilenchik. The condensation phase transition in random graph coloring. In *Proc. 18th RANDOM*, 2014.

[7] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6(23):1–13 (electronic), 2001.

[8] B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson. The scaling window of the 2-SAT transition. *Random Structures Algorithms*, 18(3):201–256, 2001.

[9] C. Bordenave and P. Caputo. Large deviations of empirical neighborhood distribution in sparse random graphs. arXiv:1308.5725v1, 2013.

[10] A. Braunstein, M. Mézard, and R. Zecchina. Survey propagation: an algorithm for satisfiability. *Random Struct. Algor.*, 27(2):201–226, 2005.

[11] V. Chvátal and B. Reed. Mick gets some (the odds are on his side). In *Proc. 33rd FOCS*, pages 620–627. IEEE, 1992.

[12] A. Coja-Oghlan. A better algorithm for random k-SAT. *SIAM J. Comput.*, 39(7):2823–2864, 2010.

[13] A. Coja-Oghlan. Random regular k-sat. arXiv:1310.2728v1, 2013.

[14] A. Coja-Oghlan and K. Panagiotou. Catching the k-NAESAT threshold. In *Proc. 45th STOC*, pages 899–907. ACM, New York, 2012.

[15] A. Coja-Oghlan and K. Panagiotou. Going after the k-SAT threshold. In *Proc. 45th STOC*, pages 705–714, New York, NY, USA, 2013. ACM.

[16] A. Coja-Oghlan and K. Panagiotou. The asymptotic k-SAT threshold. arXiv:1310.2728v5, 2014.

[17] J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. arXiv:1310.4787v1, 2013.

[18] J. Ding, A. Sly, and N. Sun. Satisfiability threshold for random regular NAE-SAT. In *Proc. 46th STOC*. ACM, 2013.

[19] W. Fernandez de la Vega. Random 2-SAT: results and problems. *Theoret. Comp. Sci.*, 265(1):131–146, 2001.

[20] J. Franco and M. Paull. Probabilistic analysis of the Davis–Putnam procedure for solving the satisfiability problem. *Disc. Appl. Math.*, 5(1):77–87, 1983.

[21] S. Franz and M. Leone. Replica bounds for optimization problems and diluted spin systems. *J. Statist. Phys.*, 111(3-4):535–564, 2003.

[22] E. Friedgut. Sharp thresholds of graph properties, and the k-SAT problem. *J. Amer. Math. Soc.*, 12(4):1017–1054, 1999. With an appendix by Jean Bourgain.

[23] A. Frieze and S. Suen. Analysis of two simple heuristics on a random instance of k-SAT. *J. Algorithms*, 20(2):312–355, 1996.

[24] D. Gamarnik and M. Sudan. Limits of local algorithms over sparse random graphs. In *Proc. 5th ITCS*, pages 369–376, New York, NY, USA, 2014. ACM.

[25] A. Goerdt. A threshold for unsatisfiability. *J. Comput. System Sci.*, 53(3):469–486, 1996.

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

[27] F. Krząkała, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová. Gibbs states and the set of solutions of random constraint satisfaction problems. *Proc. Natl. Acad. Sci.*, 104(25):10318–10323, 2007.

[28] E. Maneva, E. Mossel, and M. J. Wainwright. A new look at survey propagation and its generalizations. In *Proc. 16th SODA*, pages 1089–1098. ACM, New York, 2005.

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

[30] M. Mézard and A. Montanari. *Information, physics, and computation*. Oxford Graduate Texts. Oxford University Press, Oxford, 2009.

[31] M. Mézard and G. Parisi. Replicas and optimization. *J. Phys. Lett.*, 46(17):771–778, 1985.

[32] M. Mézard, G. Parisi, and M. A. Virasoro. *Spin glass theory and beyond*, volume 9 of *World Scientific Lecture Notes in Physics*. World Scientific Publishing Co., Inc., Teaneck, NJ, 1987.

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

[34] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina. Two solutions to diluted p-spin models and XORSAT problems. *J. Statist. Phys.*, 111(3-4):505–533, 2003.

[35] A. Montanari, F. Ricci-Tersenghi, and G. Semerjian. Clusters of solutions and replica symmetry breaking in random k-satisfiability. *J. Stat. Mech. Theory E.*, 2008(04):P04004, 2008.

[36] D. Panchenko and M. Talagrand. Bounds for diluted mean-fields spin glass models. *Probab. Theory Rel. Fields*, 130(3):319–336, 2004.

[37] B. Pittel and G. B. Sorkin. The satisfiability threshold for k-XORSAT. arXiv:1212.1905v2, 2012.

[38] M. Rahman and B. Virag. Local algorithms for independent sets are half-optimal. arXiv:1402.0485v1, 2014.