Breaking of 1RSB in random regular MAX-NAE-SAT

Zsolt Bartha*, Nike Sun†, and Yumeng Zhang‡
*Berkeley Statistics Department, Berkeley, California 94720-3860, USA
Email: bartha@berkeley.edu
†MIT Mathematics Department, Cambridge, Massachusetts 02139-4307, USA
Email: nsun@mit.edu
‡Stanford Statistics Department, Stanford, California 94305-4020, USA
Email: ymzhang@berkeley.edu

***Abstract*—For several models of random constraint satisfaction problems, it was conjectured by physicists and later proved that a sharp satisfiability transition occurs. In the *unsatisfiable* regime, it is natural to consider the problem of *max-satisfiability*: violating the least number of constraints. This is a combinatorial optimization problem on the random energy landscape defined by the problem instance. In the bounded density regime, a very precise estimate of the max-sat value was obtained by Achlioptas, Naor, and Peres (2007), but it is not sharp enough to indicate the nature of the energy landscape. Later work (Sen, 2016; Panchenko, 2016) shows that for very large but bounded density, the max-sat value approaches the *mean-field* (complete graph) limit: this is conjectured to have an "FRSB" structure where near-optimal configurations form clusters within clusters, in an ultrametric hierarchy of infinite depth inside the discrete cube. A stronger form of FRSB was shown in several recent works to have algorithmic implications (again, in complete graphs). Consequently we find it of interest to understand how the model transitions from 1RSB near the satisfiability threshold, to (conjecturally) FRSB at large density. In this paper we show that in the random regular NAE-SAT model, the 1RSB description breaks down by a certain threshold density that we estimate rather precisely. This is proved by an explicit perturbation in the 2RSB parameter space. The choice of perturbation is inspired by the "bug proliferation" mechanism proposed by physicists (Montanari and Ricci-Tersenghi, 2003; Krzakala, Pagnani, and Weigt, 2004), corresponding roughly to a percolation-like threshold for a subgraph of dependent variables.**

***Keywords*-constraint satisfaction problems; energy landscape**

I. Introduction

A **random constraint satisfaction problem** (random CSP), broadly construed, is any problem specified by N variables subject to M random constraints. We shall consider a prototypical example, **random regular k-NAE-SAT**, where an instance $\mathscr{G}_N$ involves N binary variables $x_i \in \{\mathbf{0},\mathbf{1}\}$, subject to $M = N\alpha$ random constraints such that each constraint involves a subset of k variables (the formal definition is below). In the **satisfiable** regime $0 \le \alpha \le \alpha_{\text{sat}}$, with high probability the solution space is a nonempty (random) subset $\mathsf{S}(\mathscr{G}_N) \subseteq \{\mathbf{0},\mathbf{1}\}^N$. It is predicted by physicists [1] to undergo a precise series of sharp structural transitions as α increases between zero and α_{sat}. Several of these predictions have now been supported by rigorous results: for example, we point to works on solution geometry [2]–[4], the exact satisfiability threshold α_{sat} [5]–[7], the number of solutions [8], and associated inference problems [9]. In particular it is known that $\alpha_{\text{sat}} = 2^{k-1}\ln 2 - O(1)$.

In this paper we consider the **unsatisfiable** regime $\alpha > \alpha_{\text{sat}}$, where with high probability the solution space $\mathsf{S}(\mathscr{G}_N)$ is empty. It then becomes natural to study the **max-satisfiable value** (or **ground state energy**)

$$\mathbf{e}_{\min}(\mathscr{G}_N) \equiv \min \frac{1}{N} \min \left\{ \mathscr{H}(\underline{x}) : \underline{x} \in \{\mathbf{0},\mathbf{1}\}^N \right\},$$
$$\mathscr{H}(\underline{x}) \equiv \#\Big\{\text{constraints violated by } \underline{x}\Big\}.$$

In the computer science literature on this problem, much attention has focused on *strong refutation* [10]: is there an efficiently computable bound $\mathbf{e}_{\text{alg}}(\mathscr{G}_N)$ such that $\mathbf{e}_{\text{alg}}(\mathscr{G}_N) \le \mathbf{e}_{\min}(\mathscr{G}_N)$ for *any* $\mathscr{G}_N$, and $\mathbf{e}_{\text{alg}}(\mathscr{G}_N) \ge \alpha_N\delta$ with high probability (with δ a positive constant) for *random* $\mathscr{G}_N$?

In the regime where the clause density $\alpha = \alpha_N$ **diverges** in N, an easy union bound gives

$$\mathbf{e}_{\min}(\mathscr{G}_N) = \frac{(1-o_N(1))\alpha_N}{2^{k-1}}.$$

This allows a simple phrasing of an even more stringent version of strong refutation: is there an efficiently computable bound $\mathbf{e}_{\text{alg}}(\mathscr{G}_N) \le \mathbf{e}_{\min}(\mathscr{G}_N)$ (again, for any $\mathscr{G}_N$) such that

$$\mathbf{e}_{\text{alg}}(\mathscr{G}_N) = \frac{(1+o_N(1))\alpha_N}{2^{k-1}}$$

with high probability for random $\mathscr{G}_N$? For the closely related problem of random k-SAT, an efficient spectral algorithm of this type exists above $\alpha_N \approx N^{k/2-1}$ [11] (and extended by [12]). On the other hand, within a large family of convex programming algorithms (as defined by the *sum-of-squares hierarchy*) it has been shown that many problems of this kind (including random k-SAT, and random k-NAE-SAT for k even) are solvable in subexponential but not polynomial time for $1 \ll \alpha_N \ll N^{k/2-1}$ [13]–[16].

In the regime where α does **not** diverge with N, very strong bounds on $\mathbf{e}_{\min}(\mathscr{G}_N)$ are given by [17], as we will review below. However, the bounds are not quite precise

2575-8454/19/$31.00 ©2019 IEEE
DOI 10.1109/FOCS.2019.00086

IEEE computer society

enough to give information about the nature of the energy landscape. More recent results in the spin glass literature [18], [19] show that for α very large (roughly, $\Omega(64^k)$) the max-sat value approaches the **mean-field** (complete graph) limit, which is given by a Parisi-type variational formula [20] (in the physics literature see [21], [22]). The solution of the mean-field variational formula is conjectured to be "**full replica symmetry breaking**" (FRSB), e.g. by analogy with the zero-temperature Sherrington–Kirkpatrick model [23]. A stronger version of FRSB has been shown in several recent works (in mean-field settings) to have algorithmic implications [24]–[26]. By contrast, results near the satisfiability threshold [7], [8] are consistent only with "**one-step replica symmetry breaking**" (1RSB). This is to say that as α increases from α_{sat} to ∞, the model must transition from 1RSB to FRSB; and one may even speculate further on whether the $N^{k/2-1}$ threshold in the algorithmic literature relates to a transition in the type of FRSB.

In this paper we study a phenomenon which is proposed in the physics literature as the first transition beyond α_{sat} in the type of RSB. It is predicted to occur at an explicit value α_{Ga} [27], [28] (termed the **Gardner transition**, after [29]) — by a mechanism of **bug proliferation**, which we describe below. A simple consequence of this prediction is that the ground state energy would coincide with the 1RSB value $\mathbf{e}_{\text{1RSB}}$ up to α_{Ga}, but not thereafter. Our main result is a rigorous upper bound on this transition:

Theorem 1.1: For all $k \geq k_0$ (where k_0 is an absolute constant), if $\mathscr{G}_N$ is an instance of random regular k-NAE-SAT on N variables subject to $N\alpha$ constraints (Definition 1.2), and $\mathbb{E}$ is expectation over $\mathscr{G}_N$, then the quantity

$$\liminf_{N\to\infty} \left\{ \mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)] - \mathbf{e}_{\text{1RSB}}(\alpha) \right\} \tag{1}$$

is well-defined and nonnegative for all $\alpha_{\text{sat}} \leq \alpha \leq 4^k/k$. It is strictly positive for all $\alpha_{\text{Ga}} \leq \alpha \leq 4^k/k$ where $\alpha_{\text{Ga}} \asymp 4^k/k^3$. The formal characterizations of $\mathbf{e}_{\text{1RSB}}(\alpha)$ and α_{Ga} appear below in Propositions 7.2 and 8.1.

We will see soon (§II) that the ground state energy is naturally parametrized as

$$\mathbf{e}_{\min} = \frac{\alpha(1-\mathbf{p}_{\max})}{2^{k-1}}$$

for $0 \leq \mathbf{p}_{\max} \leq 1$. The first assertion of the theorem, the nonnegativity of (1), improves on the best previous upper bound on $\mathbf{p}_{\max}$ by a factor $1-\Omega(x)$ where the correction $\Omega(x)$ reflects the typical sizes of **clusters** of near-max-satisfiable configurations. We give the basic intuition for this correction in §III, and show in (12) that in the regime $2^k k^2 \ll_k \alpha \ll_k 4^k/k$ we expect a correction $x \geq \Omega(1/d^{1/2})$. In the full version of the paper we state a more precise bound for all $\alpha_{\text{sat}} \leq \alpha \leq 4^k/k$.

The result relies on an abstract "interpolation bound" proved in [8], which was adapted from a combination of prior works [30]–[34]. Its main consequence, for our purpose, is stated in Proposition 9.1 below; it involves an optimization over parameters $0 \leq y_1 \leq y_2$ and over a large space of probability measures Q. We prove Theorem 1.1 by direct analysis of the bound in a specific region of (y_1, y_2, Q). This seems to bear some resemblance to approaches of [23], [35], although only at a high level. Our explicit choice of perturbation is based on the "bug proliferation" mechanism proposed by physicists [27], [28], which we detail in the introductory section below. We leave as an open question to prove the matching lower bound, i.e., to show that $\lim_N \mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)] = \mathbf{e}_{\text{1RSB}}(\alpha)$ for all $\alpha_{\text{sat}} \leq \alpha \leq \alpha_{\text{Ga}}$.

In the remainder of this introductory section we present some guiding heuristics for this model, leading to the formal definitions of $\mathbf{e}_{\text{1RSB}}(\alpha)$ and α_{Ga}. Our discussion is based primarily on [17], together with the two papers from the physics literature that describe the bug proliferation mechanism: of the latter, one studies a similar model as here for $k = 3, 4$ [27], while the other studies the q-coloring model [28]. We will focus on the combinatorial intuition for k-NAE-SAT which simplifies when k is large. At the end of this section we outline the proof of Theorem 1.1. Before proceeding further, we formally define the model:

Definition 1.2 (random regular NAE-SAT*):* Let d, k be positive integers, and assume N is a positive integer such that $M = Nd/k$ is also integer. A **random d-regular k-NAE-SAT** instance on N variables is encoded by a random bipartite graph $\mathscr{G}_N$. The vertex set of $\mathscr{G}_N$ is partitioned into $V = \{v_1, \ldots, v_N\}$ (**variables**) and $F = \{a_1, \ldots, a_M\}$ (constraints or **clauses**). The two sets V, F are joined by a set E of random edges, generated according to the "configuration model": give d half-edges to each $v \in V$, give k half-edges to each $a \in F$, then take a uniformly random matching between the V-incident and F-incident half-edges to form a total of $Nd = Mk$ edges. Note that the sampling procedure can result in multi-edges, so $\mathscr{G}_N$ is more precisely a multigraph. Finally, assign to each $e \in E$ an independent label $\mathbf{L}_e$ sampled uniformly from $\{\mathbf{0}, \mathbf{1}\}$. We denote the instance as $\mathscr{G}_N = (V, F, E, \underline{\mathbf{L}})$. For $e \in E$ we write $v(e)$ for the incident variable, and $a(e)$ for the incident clause. We write

$$\begin{aligned}\hat{\nabla} e &\equiv \delta a(e) \setminus e \\ &= \{\text{edges incident to } e \text{ through a clause}\}, \\ \dot{\nabla} e &\equiv \delta v(e) \setminus e \\ &= \{\text{edges incident to } e \text{ through a variable}\}.\end{aligned}$$

For any variable $v \in V$ we write δv for the ordered d-tuple of edges incident to v, and ∂v for the ordered d-tuple of clauses $(a(e))_{e \in \delta v}$. For any clause $a \in F$ we write δa for the ordered k-tuple of edges incident to a, and ∂a for the ordered k-tuple of variables $(v(e))_{d \in \delta a}$. If $a \in F$ and $v \in V$ are neighbors joined by a single edge e (as will most often be the case) then we write $e \equiv (av)$. Given a variable assignment $\underline{x} \in \{\mathbf{0}, \mathbf{1}\}^N$, a clause $a \in F$ is **violated** if and

only if the k-tuple $(\mathbf{L}_e \oplus x_{v(e)})_{e\in\delta a}$ is **all equal** (all $\mathbf{0}$ entries or all $\mathbf{1}$ entries). A **solution** of $\mathscr{G}_N$ is a variable assignment $\underline{x} \in \{\mathbf{0},\mathbf{1}\}^N$ that violates no clauses.

Definition 1.3 (energy lanscape, max-satisfiable value): Given an instance $\mathscr{G}_N$ generated as in Definition 1.2, its **energy landscape** or **Hamiltonian** is simply the total count of violated clauses: for $\underline{x} \in \{\mathbf{0},\mathbf{1}\}^N$,

$$\mathscr{H}_N(\underline{x}) \equiv \sum_{a\in F} \mathscr{H}_a(\underline{x})\,, \tag{2}$$

$$\mathscr{H}_a(\underline{x}) \equiv \mathbf{1}\Big\{\#\Big\{e\in\delta a : \mathbf{L}_e \oplus x_{v(e)} = \mathbf{1}\Big\} \in \{0,k\}\Big\}\,.$$

Note that $\mathscr{H}_N$ is a random function on $\{\mathbf{0},\mathbf{1}\}^N$ determined by the instance $\mathscr{G}_N$. The solutions of $\mathscr{G}_N$ are precisely the zeroes of $\mathscr{H}_N$. The **max-satisfiable value** (**ground state energy**) of $\mathscr{G}_N$ is

$$\begin{aligned}\mathbf{e}_{\min}(\mathscr{G}_N) &\equiv \frac{\mathbf{E}_{\min}(\mathscr{G}_N)}{N}\\ &\equiv \frac{1}{N}\min\Big\{\mathscr{H}_N(\underline{x}) : \underline{x}\in\{\mathbf{0},\mathbf{1}\}^N\Big\}\,.\end{aligned}$$

Note that $0 \le \mathbf{e}_{\min}(\mathscr{G}_N) \le \alpha \equiv d/k$, and $\mathbf{e}_{\min}(\mathscr{G}_N)$ is positive if and only if $\mathscr{G}_N$ has no proper solutions. Let

$$\begin{aligned}\frac{\alpha(1-\mathbf{p}^\bullet(\alpha))}{2^k} &\equiv \mathbf{e}_\bullet(\alpha) \equiv \liminf_{N\to\infty}\mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)]\\ &\le \limsup_{N\to\infty}\mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)] \equiv \mathbf{e}^\bullet(\alpha) \equiv \frac{\alpha(1-\mathbf{p}_\bullet(\alpha))}{2^k}\,.\end{aligned}$$

(If the two sides are equal we write $\mathbf{e}_\star(\alpha) \equiv \mathbf{e}_\bullet(\alpha) = \mathbf{e}^\bullet(\alpha)$.) We also write "$\alpha_{\max}(\mathbf{p}) \le \alpha$" to mean that $\mathbf{p}^\bullet(\alpha') < \mathbf{p}$ for all $\alpha' > \alpha$, and similarly "$\alpha_{\max}(\mathbf{p}) \ge \alpha$" to mean that $\mathbf{p}_\bullet(\alpha') > \mathbf{p}$ for all $\alpha' < \alpha$.

Remark 1.4: Physicists predict that a broad family of random CSPs (including (NAE-)SAT, proper coloring, and independent set) exhibit qualitatively similar phase diagrams ([1] and refs. therein). The existing rigorous literature has proved different aspects of these predictions in different models, including for at least six closely related variants of the model specified in Definition 1.2: namely, random regular k-NAE-SAT, random k-NAE-SAT, random regular k-SAT, random k-SAT, random regular k-hypergraph bicoloring, random k-hypergraph bicoloring. Throughout this introduction, to simplify the discussion we will (nonrigorously) transfer all existing results to the setting of random regular k-NAE-SAT. It is not unreasonable to expect that a result proved in any of the other models can be reproved in random regular k-NAE-SAT, which is mathematically the simplest of all the six. Certainly, however, none of our formal results relies on this assumption.

To explain the basic intuitions underlying this paper, in §II we review the first moment bound of [17] in the setting of random regular k-NAE-SAT. We then explain in §III why the first moment bound is loose, and a rough heuristic correction. In §IV we explain that when α is not too large the heuristic correction is a reasonable approximation, but it should fail beyond some threshold $\alpha \asymp 4^k/k^3$. In §V and §VI we explain the more refined heuristic provided by the 1RSB combinatorial framework. This leads to the formal definitions of $\mathbf{e}_{\text{1RSB}}(\alpha)$ and α_{Ga}, in §VII and §VIII respectively. Finally, in §IX we state the interpolation bound and describe the proof approach.

II. First moment bound

Throughout this paper we write $f_{n,k} \asymp g_{n,k}$ to indicate that $C^{-1} \le f_{n,k}/g_{n,k} \le C$ for a constant C not depending on n,k. We write $f \ll_k g$ to indicate that $\lim_{k\to\infty} f/g = 0$. We parametrize

$$\mathbf{c} \equiv \frac{\alpha}{2^{k-1}\ln 2}\,,\quad \mathbf{e} \equiv \frac{\alpha(1-\mathbf{p})}{2^{k-1}}\,. \tag{3}$$

To explain the above parametrization of $\mathbf{e}$, consider an instance $\mathscr{G}_N$ of d-regular random k-NAE-SAT, and let $\mathscr{H}_N$ be its Hamiltonian defined by (2) above. For any *fixed* $\underline{x} \in \{\mathbf{0},\mathbf{1}\}^N$, the number of constraints that it violates is distributed as $\mathscr{H}_N(\underline{x}) \sim \mathrm{Bin}(M, 1/2^{k-1})$, so $\mathbb{E}\mathscr{H}_N(\underline{x})/N = \alpha/2^{k-1}$. Therefore it is certainly the case that $\mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)] \le \alpha/2^{k-1}$, so it is natural to parametrize energies as in (3). Now, following [17], for any given energy level $0 \le \mathbf{e} \le \alpha$, and any $0 < \eta \le 1$, we can consider

$$\begin{aligned}X_{\mathbf{e},\eta} &= \sum_{\underline{x}\in\{\mathbf{0},\mathbf{1}\}^N} \mathbf{1}\Big\{\frac{\mathscr{H}_N(\underline{x})}{N} \le \mathbf{e}\Big\}\frac{\eta^{\mathscr{H}_N(\underline{x})}}{\eta^{N\mathbf{e}}}\\ &\ge \#\Big\{\underline{x}\in\{\mathbf{0},\mathbf{1}\}^N : \frac{\mathscr{H}_N(\underline{x})}{N} \le \mathbf{e}\Big\} \equiv Y_{\mathbf{e}}\,.\end{aligned} \tag{4}$$

If $\mathbb{E}$ is expectation over the random instance $\mathscr{G}_N$, then

$$\begin{aligned}\mathbb{E}X_{\mathbf{e},\eta} &\le \mathbb{E}\sum_{\underline{x}\in\{\mathbf{0},\mathbf{1}\}^N}\frac{\eta^{\mathscr{H}_N(\underline{x})}}{\eta^{N\mathbf{e}}} = \frac{2^N}{\eta^{N\mathbf{e}}}\Big\{1-\frac{2}{2^k}+\frac{2\eta}{2^k}\Big\}^{N\alpha}\\ &= \exp\Big\{N\mathsf{f}_\eta(\alpha,\mathbf{e})\Big\}\,.\end{aligned} \tag{5}$$

If $\alpha,\mathbf{e}$ are fixed, a stationary point of f as a function of η is given by

$$\begin{aligned}\eta = \eta(\alpha,\mathbf{e}) &= \frac{\mathbf{e}(2^{k-1}-1)}{\alpha-\mathbf{e}}\\ &= \frac{(1-\mathbf{p})(2^{k-1}-1)}{2^{k-1}-(1-\mathbf{p})} \equiv \eta(\mathbf{p})\,.\end{aligned} \tag{6}$$

Setting $\mathsf{f}_{\eta(\mathbf{p})}(\alpha,\mathbf{e}) = 0$ gives $\alpha = \alpha_{\text{ubd}}(\mathbf{p}) \equiv \mathbf{c}(\mathbf{p})\cdot 2^{k-1}\ln 2$ where

$$\mathbf{c}(\mathbf{p}) \le \frac{1}{\mathbf{p}+(1-\mathbf{p})\ln(1-\mathbf{p})}\,. \tag{7}$$

Note that $\mathbf{c}(\mathbf{p})$ is strictly decreasing with respect to $\mathbf{p}$, since

$$\frac{d}{d\mathbf{p}}\frac{1}{\mathbf{c}(\mathbf{p})} = \ln\frac{2^{k-1}-(1-\mathbf{p})}{(2^{k-1}-1)(1-\mathbf{p})} > 0\,.$$

We have $\mathbf{c}(\mathbf{p}) \uparrow \infty$ as $\mathbf{p} \downarrow 0$, and

$$\mathbf{c}(1) = \frac{1}{-2^{k-1}\ln(1-1/2^{k-1})} = 1 - \Theta(2^{-k})\,.$$

Therefore the inverse function is well-defined for all $\alpha \geq \mathbf{c}(1)2^{k-1}\ln 2$, and we denote it as

$$\mathbf{e}_{\text{lbd}}(\alpha) \equiv \frac{\alpha(1-\mathbf{p}_{\text{ubd}}(\alpha))}{2^{k-1}} \equiv (\alpha_{\text{ubd}})^{-1}(\alpha)\,. \tag{8}$$

Since f_η is decreasing in α, we conclude that $\mathsf{f}_{\eta(\mathbf{p})}(\alpha, \mathbf{e}) < 0$ for all $\alpha > \alpha_{\text{ubd}}(\mathbf{p})$. For any such α, Markov's inequality gives that $\mathbb{P}(Y_{\mathbf{e}} > 0) \leq \mathbb{E}Y_{\mathbf{e}} \leq \mathbb{E}X_{\mathbf{e},\eta}$ is exponentially small with respect to N. We can summarize the above as

Lemma 2.1: If $\mathscr{G}_N$ is random regular k-NAE-SAT on N variables subject to $N\alpha$ constraints, then

$$\begin{aligned}&\liminf_{N\to\infty} \mathbb{E}[\mathbf{e}_{\min}(\mathscr{G}_N)]\\ &\quad = \liminf_{N\to\infty} \mathbb{E}\Big[\min\Big\{\mathbf{e} \geq 0 : Y_{\mathbf{e}} > 0\Big\}\Big] \geq \mathbf{e}_{\text{lbd}}(\alpha)\end{aligned} \tag{9}$$

as defined by (8). (In the shorthand of Definition 1.2, we have $\mathbf{e}_\bullet(\alpha) \geq \mathbf{e}_{\text{lbd}}(\alpha)$.)

Lemma 2.1 is the first moment bound from [17], transferred to the setting of regular NAE-SAT. Recalling (3), if $\mathbf{c}$ is large then the expression (8) for $\mathbf{p}_{\text{ubd}}$ can be approximated by

$$\begin{aligned}\frac{1}{\mathbf{c}} &= g(\mathbf{p}_{\text{ubd}}) = \mathbf{p}_{\text{ubd}} + (1-\mathbf{p}_{\text{ubd}})\ln(1-\mathbf{p}_{\text{ubd}})\\ &= \frac{(\mathbf{p}_{\text{ubd}})^2}{2}\Big\{1 + O(\mathbf{p}_{\text{ubd}})\Big\}\,,\end{aligned}$$

so that $\mathbf{p}_{\text{ubd}} = (2/\mathbf{c})^{1/2} + O(1/\mathbf{c})$. We point out that [17] studies the more difficult model of random k-SAT, and their main result is a much more challenging lower bound, which is done by the second moment method. Translating their full result to our model would give

$$\alpha_{\text{ubd}}(\mathbf{p})\Big\{1 + O\Big(\frac{k}{2^{k/2}}\Big)\Big\} \leq \alpha_{\max}(\mathbf{p}) \leq \alpha_{\text{ubd}}(\mathbf{p})\,. \tag{10}$$

We will not seek to rigorously prove the lower bound in (10), since we expect it to be easier than the lower bound already achieved by [17]. The more interesting open problem is to establish that $\mathbf{e}_{1\text{RSB}}(\alpha)$ is tight for $\alpha \leq \alpha_{\text{Ga}}$.

III. Clustering of near-max-satisfiable configurations

We next describe the intuition for why the first moment bound (9) cannot be exactly sharp. Suppose for the sake of argument that it is. Let $\mathbf{e} = \mathbf{e}_{\text{lbd}}$ and $\eta = \eta(\mathbf{p}_{\text{ubd}})$ as above. Any $\underline{x} \in \{\mathbf{0},\mathbf{1}\}^N$ that contributes to $X_{\mathbf{e},\eta}$ will be max-satisfiable, so it certainly must satisfy the weaker condition of being **locally max-satisfiable**, in the sense that flipping any single variable x_v cannot decrease the number of violated constraints. Explicitly, let $F_{\mathbf{0}}$ be the number of clauses incident to v which are satisfied only if $x_v = \mathbf{0}$:

$$F_{\mathbf{0}} = \#\Big\{e \in \delta v : \#\Big\{\begin{matrix} g \in \hat{\nabla}(e) \text{ such that} \\ \mathbf{L}_e \oplus \mathbf{L}_g \oplus x_{v(g)} = \mathbf{1}\end{matrix}\Big\} = k-1\Big\}\,,$$

and similarly $F_{\mathbf{1}}$. The spin $x_v = \mathbf{0}$ is locally max-satisfiable if and only if $F_{\mathbf{0}} \geq F_{\mathbf{1}}$. Let $X_{\mathbf{e},\eta}(x, \ell_{\mathbf{0}}, \ell_{\mathbf{1}})$ denote the contribution to $X_{\mathbf{e},\eta}$ from configurations $\underline{x}$ with $(x_v, F_{\mathbf{0}}, F_{\mathbf{1}}) = (x, \ell_{\mathbf{0}}, \ell_{\mathbf{1}})$. By taking expectation only over the edge labels $\mathbf{L}_e$ around the clauses neighboring v, we find

$$\begin{aligned}\mathbb{E}\Big\{X_{\mathbf{e},\eta}(\mathbf{0}, \ell_{\mathbf{0}}, \ell_{\mathbf{1}})\Big\} &= C_N \mathbf{1}\{\ell_{\mathbf{0}} \geq \ell_{\mathbf{1}}\}\binom{d}{\ell_{\mathbf{0}}, \ell_{\mathbf{1}}}\\ &\quad\times\Big(1-\frac{4}{2^k}\Big)^{d-\ell_{\mathbf{0}}-\ell_{\mathbf{1}}}\Big(\frac{2}{2^k}\Big)^{\ell_{\mathbf{0}}}\Big(\frac{2\eta}{2^k}\Big)^{\ell_{\mathbf{1}}}\end{aligned} \tag{11}$$

where C_N is a factor not depending on $\ell_{\mathbf{0}}, \ell_{\mathbf{1}}$, and for any $a_1 + \ldots + a_t \leq b$ we abbreviate

$$\binom{b}{a_1,\ldots,a_t} = \frac{b!}{a_1!\cdots a_t!(b-a_1-\ldots-a_t)!}\,.$$

Summing (11) over $\ell_{\mathbf{0}} \geq \ell_{\mathbf{1}}$, we find that the total expected contribution to $X_{\mathbf{e},\eta}$ from configurations with $x_v = \mathbf{0}$ is

$$\begin{aligned}&\mathbb{E}\Big\{X_{\mathbf{e},\eta}(x_v = \mathbf{0})\Big\}\\ &\quad = C_N \sum_{0\leq\ell\leq d}\binom{d}{\ell}\Big(1-\frac{4}{2^k}\Big)^{d-\ell}\Big(\frac{2(1+\eta)}{2^k}\Big)^{\ell} P_{\eta,\ell}\end{aligned}$$

where

$$P_{\eta,\ell} = \mathbb{P}\Big(\text{Bin}\Big(\ell, \frac{\eta}{1+\eta}\Big) \leq \frac{\ell}{2}\Big)\,.$$

Simply using the crude bound $1/2 \leq P_{\eta,\ell} \leq 1$ gives

$$\mathbb{E}\Big\{X_{\mathbf{e},\eta}(x_v = \mathbf{0})\Big\} \asymp C_N\Big(1 - \frac{4}{2^k} + \frac{2(1+\eta)}{2^k}\Big)^d\,.$$

Now note that if $F_{\mathbf{0}} = F_{\mathbf{1}}$ then variable v is *free*, meaning that flipping x_v alone does not change the total number of violated constraints. Summing (11) over $\ell_{\mathbf{0}} = \ell_{\mathbf{1}} = \ell/2$ gives

$$\begin{aligned}&\mathbb{E}\Big\{X_{\mathbf{e},\eta}(x_v = \mathbf{0}, v \text{ is free})\Big\}\\ &\quad = C_N \sum_{\ell \text{ even}}\binom{d}{\ell}\Big(1-\frac{4}{2^k}\Big)^{d-\ell}\Big(\frac{4\eta^{1/2}}{2^k}\Big)^{\ell} P_\ell\end{aligned}$$

where $P_\ell = \mathbb{P}(\text{Bin}(\ell, 1/2) = \ell/2) \asymp 1/\ell^{1/2}$. Now assume that $\mathbf{c}$ is large, so $\mathbf{p}$ is small and we see from (6) that $\eta \asymp 1$. Without the factor P_ℓ, the above sum is dominated by $\ell \asymp d\eta^{1/2}/2^k \asymp d/2^k$. Accounting for P_ℓ results in

$$\begin{aligned}&\mathbb{E}\Big\{X_{\mathbf{e},\eta}(x_v = \mathbf{0}, v \text{ is free})\Big\}\\ &\quad \asymp \frac{C_N}{(d/2^k)^{1/2}}\Big(1 - \frac{4}{2^k} + \frac{4\eta^{1/2}}{2^k}\Big)^d\,.\end{aligned}$$

(A more careful version of this calculation appears in the full version of the paper.) This would suggest the typical fraction of variables that are free is something like

$$\begin{aligned}\pi_{\mathbf{f}} &\asymp \frac{\mathbb{E}[X_{\mathbf{e},\eta}(x_v = \mathbf{0}, v \text{ is free})]}{\mathbb{E}[X_{\mathbf{e},\eta}(x_v = \mathbf{0})]} \\ &\asymp \frac{1}{(d/2^k)^{1/2}}\left(1 - \frac{4}{2^k} + \frac{4\eta^{1/2}}{2^k}\right)^d \\ &\qquad\qquad \div \left(1 - \frac{4}{2^k} + \frac{2(1+\eta)}{2^k}\right)^d \\ &= \frac{1}{(d/2^k)^{1/2}}\left(1 - \frac{2(1-\eta^{1/2})^2/2^k}{1 - 4/2^k + 2(1+\eta)/2^k}\right)^d ,\end{aligned}$$

If we assume that $k^2 \ll_k \mathbf{c} \ll_k 2^k/k$, then the above simplifies to

$$\begin{aligned}\pi_{\mathbf{f}} &\asymp \frac{1}{(d/2^k)^{1/2}}\left(1 - \frac{2}{2^k\mathbf{c}}\left[1 + O\left(\frac{1}{\mathbf{c}^{1/2}}\right)\right]\right)^d \\ &= \frac{1}{(d/2^k)^{1/2}} \exp\left\{ - \frac{d}{2^{k-1}\mathbf{c}}\left[1 + O\left(\frac{1}{\mathbf{c}^{1/2}}\right)\right]\right\} \\ &\asymp \frac{1}{d^{1/2}} .\end{aligned}$$

Suppose the configuration $\underline{x}$ has order $N/d^{1/2}$ free variables. Suppose for simplicity that they *do not interact*, meaning that *flipping any subset of free variables does not change the total number of violated constraints*. We will examine the validity of this supposition in §IV, but we simply grant it for now. This would mean that for a typical max-satisfiable configuration $\underline{x}$ we can find at least $2^{N\pi_{\mathbf{f}}}$ nearby configurations $\underline{x}'$ with $\mathscr{H}_N(\underline{x}) = \mathscr{H}_N(\underline{x}')$. But this would mean $\mathbb{E}X_{\mathbf{e},\eta} \geq 2^{N\pi_{\mathbf{f}}}$, in contradiction with our choice of $\mathbf{e} = \mathbf{e}_{\text{lbd}}$ and $\eta = \eta(\mathbf{p}_{\text{ubd}})$ which ensures that $\mathbb{E}X_{\mathbf{e},\eta}$ is exponentially small in N. This suggests that $\mathbf{e}_{\text{lbd}}$ (or equivalently its inverse α_{ubd}) cannot be tight bounds; our main theorem verifies this by establishing the lower bound $\mathbf{e}_\bullet(\alpha) > \mathbf{e}_{\text{lbd}}(\alpha)$. The above calculation suggests that $\exp\{N\mathsf{f}_\eta(\alpha, \mathbf{e})\}$ overestimates the typical value of $X_{\mathbf{e},\eta}$ by at least a factor $2^{N\pi_{\mathbf{f}}}$ where $\pi_{\mathbf{f}} \asymp 1/d^{1/2}$, which suggests, in the regime $k^2 \ll_k \mathbf{c} \ll_k 2^k/k$ (equivalently $2^k k^2 \ll_k \alpha \ll 4^k/k$), that

$$\alpha_{\max}(\mathbf{p}) \leq \left\{1 - \Omega\left(\frac{1}{d^{1/2}}\right)\right\}\alpha_{\text{ubd}}(\mathbf{p}) . \tag{12}$$

In the full version of the paper, we prove a rigorous bound which covers the full regime $\alpha_{\text{sat}} \leq \alpha \leq 4^k/k$, and agrees with (12) for $k^2 \ll_k \mathbf{c} \ll_k 2^k/k$. In fact in this regime we conjecture the estimate $\Omega(1/d^{1/2})$ to be tight.

IV. Percolation of dependent free variables

We now revisit the above assumption that the free variables do not interact. Take a clause a with no incident multi-edges (as will be the case for most clauses), and suppose it neighbors two free variables $v \neq w$. If the values of $\mathbf{L}_{au} \oplus x_u$ for $u \in \partial a \setminus \{v, w\}$ are all $\mathbf{0}$ or all $\mathbf{1}$, then x_v and x_w are **linked**, meaning they cannot both be arbitrarily flipped without increasing the number of violated constraints. For a free variable v, the number of linked free variables w sharing a clause with v is (on average, heuristically)

$$r = \Big\{(d-1)(k-1)\Big\} \times \pi_{\mathbf{f}} \times \frac{1}{2^{k-2}} \asymp \frac{d^{1/2}k}{2^k} , \tag{13}$$

where the factor $(d-1)(k-1)$ accounts for the branching factor of the underlying graph $\mathscr{G}_N$. We view the process of linked free variables as a dependent percolation on $\mathscr{G}_N$ spreading at rate r given by (13). As long as the rate is small, corresponding to $d \ll_k 4^k/k^2$ or

$$\alpha \ll_k \frac{4^k}{k^3} ,$$

we would expect the percolation to be **subcritical**, in the sense that the free subgraph — the subgraph of $\mathscr{G}_N$ induced by free variables and linking clauses — is mostly a **forest** of $\tilde{O}(1)$-sized trees. Moreover, roughly a $(1-r)$-fraction of free variables should be **isolated** (not linked to any other frees), so for small r it is a reasonable approximation to assume that none of the free variables interact.

As we detail in §V below, in the context of the current problem, the 1RSB framework is simply a convenient combinatorial model for the free subgraph, which captures the effect of free variables on the total energy in a well-organized manner. It yields the prediction that the limiting ground state energy is *exactly* $\mathbf{e}_{\text{1RSB}}(\alpha)$, where $\mathbf{e}_{\text{1RSB}}(\alpha)$ is an explicit function defined below in Proposition 7.2. The threshold α_{Ga}, given formally by Proposition 8.1, is an explicit prediction of the exact percolation threshold for the 1RSB combinatorial model. The derivation of $\mathbf{e}_{\text{1RSB}}(\alpha)$ relies crucially on the assumption that the free subgraph *is* essentially a forest, which should not be the case beyond α_{Ga}. This is the basic intuition for our main result which verifies that $\mathbf{e}_{\text{1RSB}}(\alpha)$ is indeed incorrect beyond α_{Ga}.

We remark that it is a much more challenging problem to obtain a sharper estimate of the asymptotic ground state energy in the regime $\alpha > \alpha_{\text{Ga}}$. The main result that we know of was obtained for the random k-SAT model [19] (see also [18]) by comparison with mean-field limits [21], [22]; from the discussion in [19] the estimate requires roughly $\alpha \geq \Omega(64^k)$. A related result was obtained for the max-cut problem by [36], for random graphs of large degree. It remains a difficult challenge to understand the regime between the mean-field (i.e., complete graph) limits and α_{Ga}.

Having laid out the basic intuitions for the model, we next proceed to define the 1RSB combinatorial framework. We emphasize that the 1RSB model itself is a heuristic, which plays no formal role in the proof of our main result. We introduce it because it is the quickest way to motivate the exact definitions of $\mathbf{e}_{\text{1RSB}}$ and α_{Ga}. We point to [37] for an introductory account and further references on the 1RSB framework.

V. Combinatorial model of near-max-satisfiable clusters

Following our earlier discussion, we now restrict attention to the subspace $\mathsf{Q}(\mathscr{G}_N) \subseteq \{\mathbf{0},\mathbf{1}\}^N$ of configurations that are locally max-satisfiable. Define a graph on vertex set Q by putting an edge between $\underline{x}$ and $\underline{x}'$ if and only if they differ in a single coordinate and $\mathscr{H}_N(\underline{x}) = \mathscr{H}_N(\underline{x}')$. A **(locally max-satisfiable) cluster** is any subset $\omega \subseteq \mathsf{Q}$ that constitutes a maximal connected component in that graph. The 1RSB heuristic models a cluster as follows:

Definition 5.1 (warning configurations): Suppose $\mathscr{G}_N = (V, F, E, \underline{\mathbf{L}})$ is any d-regular k-NAE-SAT problem instance. A **warning configuration** on $\mathscr{G}_N$ is an element $\underline{\mathbf{w}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}^{2E}$ which assigns a pair $\mathbf{w}_e \equiv (\dot{\mathbf{w}}_e, \hat{\mathbf{w}}_e)$ to each edge $e \in E$, satisfying conditions that we now specify. We take the convention throughout that $x \oplus \mathbf{f} \equiv \mathbf{f}$. Define

$$\ell_x(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{d-1}) \equiv \#\Big\{1 \le i \le d-1 : \hat{\mathbf{w}}_i = x\Big\}.$$

Then $\underline{\mathbf{w}}$ is a valid warning configuration if and only if it satisfies variable relations

$$\dot{\mathbf{w}}_e = \dot{\mathsf{WP}}\Big(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)\Big) \tag{14}$$
$$= \begin{cases} \mathbf{0} & \ell_{\mathbf{0}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)) > \ell_{\mathbf{1}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)), \\ \mathbf{1} & \ell_{\mathbf{0}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)) < \ell_{\mathbf{1}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)), \\ \mathbf{f} & \ell_{\mathbf{0}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)) = \ell_{\mathbf{1}}(\hat{\mathbf{w}}_g : g \in \dot{\nabla}(e)), \end{cases}$$

as well as clause relations

$$\hat{\mathbf{w}}_e = \hat{\mathsf{WP}}\Big(\dot{\mathbf{w}}_g : g \in \hat{\nabla}(e)\Big) \tag{15}$$
$$= \begin{cases} \mathbf{0} & \mathbf{L}_g \oplus \dot{\mathbf{w}}_g = \mathbf{L}_e \oplus \mathbf{1} \text{ for all } g \in \hat{\nabla}(e), \\ \mathbf{1} & \mathbf{L}_g \oplus \dot{\mathbf{w}}_g = \mathbf{L}_e \oplus \mathbf{0} \text{ for all } g \in \hat{\nabla}(e), \\ \mathbf{f} & \text{otherwise}, \end{cases}$$

for all $e \in E$. We may write $\hat{\mathsf{WP}} \equiv \hat{\mathsf{WP}}_{k-1}$ and $\dot{\mathsf{WP}} \equiv \dot{\mathsf{WP}}_{d-1}$ to emphasize the number of arguments. Given $\underline{\mathbf{w}}$ let

$$\underline{\mathfrak{y}} \equiv \underline{\mathfrak{y}}(\underline{\mathbf{w}}) \equiv \Big(\dot{\mathsf{WP}}_d(\hat{\mathbf{w}}_e : e \in \delta v)\Big)_{v \in V} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}^N.$$

If $\#\{v \in V : \mathfrak{y}_v = \mathbf{f}\} \le N/k^2$ then we say that $\underline{\mathbf{w}}$ is **near-frozen**.

Under the 1RSB heuristic, there is essentially a bijective correspondence

$$\left\{\begin{array}{c}\text{locally max-satisfiable}\\ \text{clusters } \omega \subseteq \mathsf{Q} \subseteq \{\mathbf{0},\mathbf{1}\}^N\end{array}\right\} \tag{16}$$
$$\leftrightarrow \left\{\begin{array}{c}\text{near-frozen warning}\\ \text{configurations } \underline{\mathbf{w}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}^{2E}\end{array}\right\}$$

between clusters ω and near-frozen warning configurations $\underline{\mathbf{w}}$. A loose characterization of the correspondence is that $\underline{\mathfrak{y}} \equiv \underline{\mathfrak{y}}(\underline{\mathbf{w}})$ encodes the smallest subcube of $\{\mathbf{0},\mathbf{1}\}^N$ containing ω: $\mathfrak{y}_v \in \{\mathbf{0},\mathbf{1}\}$ if and only if $x_v = \mathfrak{y}_v$ for all $\underline{x} \in \omega$, and $\mathfrak{y}_v = \mathbf{f}$ if and only if x_v takes both values $\{\mathbf{0},\mathbf{1}\}$. A more precise interpretation is that

$$\begin{aligned}\dot{\mathbf{w}}_e &= \textbf{variable-to-clause warning} \text{ along } e\\ &= \text{locally optimal choice within } \omega\\ &\qquad \text{of } x_{v(e)} \text{ in absence of edge } e,\\ \hat{\mathbf{w}}_e &= \textbf{clause-to-variable warning} \text{ along } e\\ &= \text{locally optimal choice within } \omega\\ &\qquad \text{of } x_{v(e)} \text{ “in absence of” edges } \dot{\nabla}(e),\end{aligned}$$

where $\mathbf{f}$ means that both spins $\{\mathbf{0},\mathbf{1}\}$ are locally optimal. Under this interpretation, the $\dot{\mathbf{w}}, \hat{\mathbf{w}}$ must then satisfy local consistency relations, which are the so-called **warning propagation** (WP) equations (15) and (14). The near-frozen restriction rules out configurations such as $\underline{\mathbf{w}} = \mathbf{f}^{2E}$ (all messages $\mathbf{f}$) which we do not expect to correspond to any actual cluster.

VI. Tree formula for the max-satisfiable value

To give an explicit calculation, let $\tau = (V', F', E')$ be a finite bipartite tree (representing an $O(1)$-sized subgraph of $\mathscr{G}_N$) with variables at its leaves. Say τ has a **frozen boundary**, in the sense that $\dot{\mathbf{w}}_e \in \{\mathbf{0},\mathbf{1}\}$ is fixed at every leaf edge e. By applying the maps $\hat{\mathsf{WP}}, \dot{\mathsf{WP}}$ recursively inwards from the leaves, we see that there is exactly one valid warning configuration $\underline{\mathbf{w}}$ on τ that is consistent with the boundary condition. Let $\mathbf{E}_{\min}(\tau)$ be the minimum number of clauses violated by any configuration $\underline{x} \in \{\mathbf{0},\mathbf{1}\}^{V'}$ with $x_{v(e)} = \dot{\mathbf{w}}_e$ at the leaves. We next explain that $\mathbf{E}_{\min}(\tau)$ can be computed by a simple dynamic-programming-type method.

Let E'' be the set of non-leaf edges of τ. For any $e \in E''$ we let $\hat{\tau}_e$ be the component containing $a(e)$ in $\tau \setminus \dot{\nabla}(e)$, and let $\dot{\tau}_e$ be the component containing $v(e)$ in $\tau \setminus e$. Let $\hat{\mathbf{E}}_e = \mathbf{E}_{\min}(\hat{\tau}_e)$ and $\dot{\mathbf{E}}_e = \mathbf{E}_{\min}(\dot{\tau}_e)$. If V'' denotes the non-leaf variables of τ, around any $v \in V''$ we have

$$\mathbf{E}_{\min}(\tau) = \dot{\varphi}(\underline{\hat{\mathbf{w}}}_{\delta v}) + \sum_{e \in \delta v} \hat{\mathbf{E}}_e, \tag{17}$$
$$\dot{\varphi}(\underline{\hat{\mathbf{w}}}_{\delta v}) \equiv \min\Big\{\ell_{\mathbf{0}}(\underline{\hat{\mathbf{w}}}_{\delta v}), \ell_{\mathbf{1}}(\underline{\hat{\mathbf{w}}}_{\delta v})\Big\}.$$

Similarly, around any clause $a \in F'$, we have

$$\mathbf{E}_{\min}(\tau) = \hat{\varphi}(\underline{\mathbf{L} \oplus \dot{\mathbf{w}}}_{\delta a}) + \sum_{e \in \delta a} \dot{\mathbf{E}}_e, \tag{18}$$
$$\hat{\varphi}(\underline{\mathbf{L} \oplus \dot{\mathbf{w}}}_{\delta a}) \equiv \mathbf{1}\Big\{\hat{\mathsf{WP}}_k(\underline{\mathbf{L} \oplus \dot{\mathbf{w}}}_{\delta a}) \neq \mathbf{f}\Big\}.$$

We sometimes write $\dot{\varphi} \equiv \dot{\varphi}_d$ and $\hat{\varphi} \equiv \hat{\varphi}_k$ to emphasize the number of arguments. Finally, for any $e \in E''$ we have

$$\mathbf{E}_{\min}(\tau) = \bar{\varphi}(\mathbf{w}_e) + \dot{E}_e + \hat{E}_e, \tag{19}$$
$$\bar{\varphi}(\mathbf{w}_e) \equiv \mathbf{1}\Big\{\dot{\mathbf{w}}_e \oplus \hat{\mathbf{w}}_e = \mathbf{1}\Big\}.$$

By summing over the internal vertices and subtracting over the internal edges, we arrive at

$$\sum_{v\in V''}\dot\varphi(\underline{\hat{\mathbf{w}}}_{\delta v})+\sum_{a\in F'}\hat\varphi(\underline{\mathbf{L}\oplus\dot{\mathbf{w}}_{\delta a}})-\sum_{e\in E''}\bar\varphi(\mathbf{w}_e)$$
$$=\Big(\#V''+\#F'-\#E''\Big)\mathbf{E}_{\min}(\tau)=\mathbf{E}_{\min}(\tau)\,,\quad(20)$$

where the last equality uses that τ is a tree. Thus the max-satisfiable value of a tree with frozen boundary is a sum of local functionals $\dot\varphi$, $\hat\varphi$, $\bar\varphi$ of the warning configuration.

The 1RSB heuristic further assumes that for near-frozen warning configurations, the entire graph $\mathscr{G}_N=(V,F,E,\underline{\mathbf{L}})$ can essentially be carved into trees with frozen boundaries. (In reality, even in the regime where free variables do not percolate, a typical warning configuration may contain a bounded number of small cycles of free warnings which do not admit a tree decomposition. However these few cycles should only affect the number of violated clauses by $O(1)$, so can be ignored in the heuristic analysis.) Then, by summing (20) over the components of the tree decomposition, we conclude that $\underline{\mathbf{w}}$ corresponds to a cluster $\omega\subseteq\{\mathbf{0},\mathbf{1}\}^N$ at energy level

$$\mathbf{E}_{\min}(\omega;\mathscr{G}_N)\equiv\min\Big\{\mathscr{H}_N(\underline{x}):\underline{x}\in\omega\Big\}=\varphi(\underline{\mathbf{w}})$$
$$\equiv\sum_{v\in V}\dot\varphi(\underline{\hat{\mathbf{w}}}_{\delta v})+\sum_{a\in F}\hat\varphi(\underline{\mathbf{L}\oplus\dot{\mathbf{w}}_{\delta a}})-\sum_{e\in E}\bar\varphi(\mathbf{w}_e)\,.\quad(21)$$

This is the main advantage of the $\underline{\mathbf{w}}$ encoding; it allows us to read off $\mathbf{E}_{\min}(\omega;\mathscr{G}_N)$ as a sum of local terms.

VII. Explicit 1RSB prediction

Going back to the bijection (16), we can take a parameter $y\geq 0$ and consider

$$\mathfrak{Z}(y)=\sum_{\omega}\exp\Big\{-y\mathbf{E}_{\min}(\omega;\mathscr{G}_N)\Big\}$$
$$=\sum_{\underline{\mathbf{w}}}\exp\Big\{-y\varphi(\underline{\mathbf{w}})\Big\}\quad(22)$$

where the first sum goes over clusters, while the second sum goes over near-frozen warning configurations. The corresponding probability measure on warning configurations is given by

$$\mu_y(\underline{\mathbf{w}})=\frac{\exp\{-y\varphi(\underline{\mathbf{w}})\}}{\mathfrak{Z}(y)}\,.\quad(23)$$

This is sometimes called the **survey propagation** or SP_y model, and can be viewed as a refinement of the reweighting $\eta^{\mathscr{H}_N(\underline{x})}$ discussed in §II. The "lifting" from $\eta^{\mathscr{H}_N(\underline{x})}$ to $e^{-y\varphi(\underline{\mathbf{w}})}$ represents one level of **replica symmetry breaking**. The **1RSB solution** to the original model is given by the **replica symmetric** solution to the "lifted" model (23). This sometimes goes by the name of **survey propagation** (SP). In particular, the **1RSB (SP) equations** are simply the **replica symmetric** or **belief propagation** (BP) equations for the lifted model. They can be defined as a pair of mappings on the space

$$\mathcal{M}\equiv\left\{\begin{matrix}\text{probability measures } q \text{ on } \{\mathbf{0},\mathbf{1},\mathbf{f}\}\\ \text{satisfying } q(\mathbf{0})=q(\mathbf{1})\end{matrix}\right\}.\quad(24)$$

The clause survey propagation takes $\dot q\in\mathcal{M}$ and outputs

$$[\dot{\mathsf{SP}}_y(\dot q)](\hat{\mathbf{w}})=\sum_{\underline{\dot{\mathbf{w}}}}\mathbf{1}\Big\{\hat{\mathbf{w}}=\dot{\mathsf{WP}}_{k-1}(\underline{\dot{\mathbf{w}}})\Big\}\prod_{i=1}^{k-1}\dot q(\dot{\mathbf{w}}_i)\,,\quad(25)$$

where the sum goes over $\underline{\dot{\mathbf{w}}}\in\{\mathbf{0},\mathbf{1},\mathbf{f}\}^{k-1}$, and $\dot{\mathsf{SP}}_y(\dot q)$ is a probability measure on $\{\mathbf{0},\mathbf{1},\mathbf{f}\}$, and in fact $\dot{\mathsf{SP}}_y(\dot q)\in\mathcal{M}$. The variable survey propagation takes $\hat q\in\mathcal{M}$ and outputs

$$[\hat{\mathsf{SP}}_y(\hat q)](\dot{\mathbf{w}})\quad(26)$$
$$=\frac{1}{\dot z}\sum_{\underline{\hat{\mathbf{w}}}}\mathbf{1}\Big\{\dot{\mathbf{w}}=\dot{\mathsf{WP}}_{d-1}(\underline{\hat{\mathbf{w}}})\Big\}\frac{\prod_{i=1}^{d-1}\hat q(\hat{\mathbf{w}}_i)}{\exp\{y\dot\varphi_{d-1}(\underline{\hat{\mathbf{w}}})\}}\,,$$

where the sum is over $\underline{\hat{\mathbf{w}}}\in\{\mathbf{0},\mathbf{1},\mathbf{f}\}^{d-1}$, and $\dot z$ is the normalization such that $\hat{\mathsf{SP}}_y(\hat q)\in\mathcal{M}$. Let $\mathsf{SP}_y\equiv\hat{\mathsf{SP}}_y\circ\dot{\mathsf{SP}}_y$. Now, recalling (3), we hereafter restrict consideration to parameters $y\geq 0$ satisfying

$$\gamma\equiv 2\mathbf{c}\Big(1-\frac{1}{e^{y/2}}\Big)^2\asymp 1\,.\quad(27)$$

Note $\gamma\asymp\mathbf{c}\min\{1,y^2\}$. If $\mathbf{c}$ is large then (27) forces $y\asymp 1/\mathbf{c}^{1/2}$. If $\mathbf{c}\asymp 1$ then it only forces that $y\geq\Omega(1)$. Define

$$\mathcal{M}^\bullet\equiv\Big\{q\in\mathcal{M}:q(\mathbf{f})\leq\frac{1}{k^2}\Big\},\quad(28)$$
$$\mathcal{M}^\gamma\equiv\Big\{q\in\mathcal{M}^\bullet:q(\mathbf{f})\asymp\frac{2^{-k\gamma/2}}{(\max\{\mathbf{c}ke^{-y/2},1\})^{1/2}}\Big\}.\quad(29)$$

We prove the following result on fixed points of the SP_y recursion:

Proposition 7.1 (proved in full version): Suppose $\alpha=\mathbf{c}2^{k-1}\ln 2$ with $\alpha_{\text{sat}}\leq\alpha\leq 4^k/k$, and suppose $y\geq 0$ satisfies (27). Then in the set $\mathcal{M}^\bullet$ there is a unique $\dot q_y$ satisfying the fixed-point equation $\dot q_y=\mathsf{SP}_y(\dot q_y)$. It must further lie in the smaller domain $\mathcal{M}^\gamma$.

Let $\dot q=\dot q_y$ be as given by Proposition 7.1, and denote $\hat q\equiv\hat q_y\equiv\dot{\mathsf{SP}}_y(\dot q_y)$. Recall the local functionals $\dot\varphi$, $\hat\varphi$, $\bar\varphi$ from (17), (18), (19). We can define three probability measures — $\dot\nu_y$ over $\underline{\hat{\mathbf{w}}}\in\{\mathbf{0},\mathbf{1},\mathbf{f}\}^d$, $\hat\nu_y$ over $\underline{\dot{\mathbf{w}}}\in\{\mathbf{0},\mathbf{1},\mathbf{f}\}^k$, and lastly $\bar\nu_y$ over $\mathbf{w}=(\dot{\mathbf{w}},\hat{\mathbf{w}})\in\{\mathbf{0},\mathbf{1},\mathbf{f}\}^2$ — as follows:

$$\dot\nu_y(\underline{\hat{\mathbf{w}}})=\frac{1}{\dot{\mathfrak{z}}_y(\hat q)}\exp\Big\{-y\dot\varphi_d(\underline{\hat{\mathbf{w}}})\Big\}\prod_{i=1}^{d}\hat q(\hat{\mathbf{w}}_i)\,,\quad(30)$$

$$\hat\nu_y(\underline{\dot{\mathbf{w}}})=\frac{1}{\hat{\mathfrak{z}}_y(\dot q)}\exp\Big\{-y\hat\varphi_k(\dot{\mathbf{w}})\Big\}\prod_{i=1}^{k}\dot q(\dot{\mathbf{w}}_i)\,,\quad(31)$$

$$\bar\nu_y(\mathbf{w})=\frac{1}{\bar{\mathfrak{z}}_y(\dot q,\hat q)}\exp\Big\{-\bar\varphi(\dot{\mathbf{w}},\hat{\mathbf{w}})\Big\}\dot q(\dot{\mathbf{w}})\hat q(\hat{\mathbf{w}})\,.\quad(32)$$

Under the SP heuristic, the local marginals of the measure (23) are approximately given by the ν: for instance,

$$\mu_y\left(\underline{\mathbf{w}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}^{2E} : \underline{\hat{\mathbf{w}}}_{\delta v} = \underline{\hat{\mathbf{w}}}\right) \approx \dot{\nu}_y(\underline{\hat{\mathbf{w}}})\,.$$

The corresponding energy level can be obtained by averaging (21) with respect to the ν: this gives

$$\mathbf{e}(y) = \sum_{\underline{\hat{\mathbf{w}}}} \dot{\varphi}_d(\underline{\hat{\mathbf{w}}})\dot{\nu}_y(\underline{\hat{\mathbf{w}}}) + \alpha \sum_{\underline{\dot{\mathbf{w}}}} \hat{\varphi}_k(\underline{\dot{\mathbf{w}}})\hat{\nu}_y(\underline{\dot{\mathbf{w}}}) - d\sum_{\mathbf{w}} \bar{\varphi}(\mathbf{w})\bar{\nu}_y(\mathbf{w})\,. \tag{33}$$

The SP heuristic further predicts that $N^{-1}\ln \mathfrak{Z}(y)$ converges (for a suitable range of y) to the replica symmetric formula,

$$\mathfrak{F}(y) = \ln \dot{\mathfrak{z}}_y(\hat{q}) + \alpha\left\{\ln \hat{\mathfrak{z}}_y(\dot{q}) - k\ln \bar{\mathfrak{z}}_y(\dot{q},\hat{q})\right\}, \tag{34}$$

where $\dot{\mathfrak{z}}_y$, $\hat{\mathfrak{z}}_y$, and $\bar{\mathfrak{z}}_y$ are the normalizing constants from (30), (31), and (32). Now, returning to (22), suppose that we had an "energetic complexity function" function Σ such that

$$\mathfrak{Y}_{\mathbf{e}} \equiv \#\left\{\underline{\mathbf{w}} : \varphi(\underline{\mathbf{w}}) \approx N\mathbf{e}\right\} \approx \mathbb{E}\mathfrak{Y}_{\mathbf{e}} \approx \exp\left\{N\Sigma(\mathbf{e})\right\}$$

where the interpretation for $\Sigma(\mathbf{e}) < 0$ is that $\mathbb{E}\mathfrak{Y}_{\mathbf{e}}$ is exponentially small with respect to N so $\mathfrak{Y}_{\mathbf{e}} = 0$ whp. Then we would expect

$$\exp\left\{N\mathfrak{F}(y)\right\} \approx \mathfrak{Z}(y) \approx \exp\left\{N\max_{\mathbf{e}}\left\{\Sigma(\mathbf{e}) - y\mathbf{e} : \Sigma(\mathbf{e}) \geq 0\right\}\right\},$$

that is to say, given Σ we can obtain $\mathfrak{F}$ by taking the Legendre dual. Of course, we are in the opposite situation: we already obtained explicit expressions (33) and (34) for $\mathbf{e}(y)$ and $\mathfrak{F}(y)$, but we do not know Σ. We therefore formally define the **energetic complexity function** as $\mathfrak{S}(y) = \mathfrak{F}(y) + y\mathbf{e}(y)$. (While the informal complexity Σ is a function of $\mathbf{e}$, the formal complexity $\mathfrak{S}$ is a function of y.) Recall (27) and let

$$\Gamma(y) \equiv \mathbf{c}\left(1 - \frac{1+y}{e^y}\right). \tag{35}$$

It is straightforward to verify that $\gamma(y)/2 \leq \Gamma(y) \leq \gamma(y)$ for all $y \geq 0$; see Figure 1a. For small y (corresponding, via (27), to large $\mathbf{c}$) we have

$$\Gamma(y) = \gamma(y)\left\{1 - O(y)\right\}.$$

For $y \geq \Omega(1)$ (corresponding, via (27), to $\mathbf{c} \asymp 1$) we have instead

$$2\Gamma(y) = \gamma(y)\left\{1 + O(e^{-y/2})\right\}.$$

The following proposition formally defines the 1RSB formula.

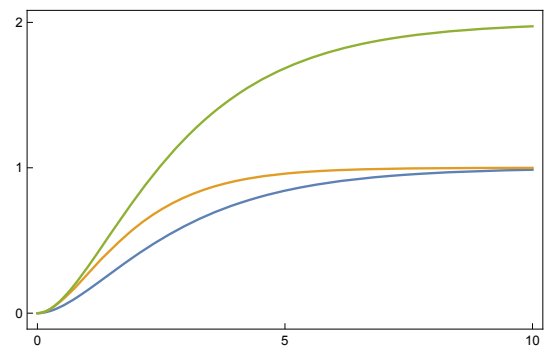
Proposition 7.2 (proved in full version): Suppose $k \geq k_0$ and $\alpha_{\text{sat}} \leq \alpha \leq 4^k/k$; and denote $\mathbf{c} = \alpha/(2^{k-1}\ln 2)$. Then, on the range of y satisfying (27), the function $\mathfrak{S}(y) = \mathfrak{F}(y) + y\mathbf{e}(y)$ is smooth and strictly decreasing, with a unique root $y_\star = y_\star(\alpha)$. With $\Gamma(y)$ as defined by (35), this root satisfies the estimate

$$\Gamma(y_\star) = 1 + O\left(\frac{1}{e^{\Omega(k)}}\right). \tag{36}$$

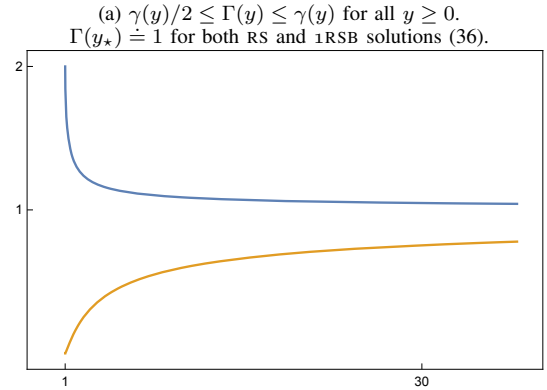
The **1RSB ground state energy** can be defined as

$$\mathbf{e}_{\text{1RSB}}(\alpha) \equiv \mathbf{e}(y_\star(\alpha)) = -\inf\left\{\frac{\mathfrak{F}(y)}{y} : y \geq 0\,, y \text{ satisfies (27)}\right\} \tag{37}$$

(the equivalence of the last two quantities will be proved).



(a) $\gamma(y)/2 \leq \Gamma(y) \leq \gamma(y)$ for all $y \geq 0$. $\Gamma(y_\star) \doteq 1$ for both RS and 1RSB solutions (36).



(b) Upper curve: $\gamma(y_\star)$ as a function of $\mathbf{c} \geq 1$. Lower curve: $\exp(-y_\star)$ as a function of $\mathbf{c} \geq 1$.

Figure 1: Approximate parameters of the 1RSB solution. At clause density $\alpha = \mathbf{c}2^{k-1}\ln 2$, the max-satisfiable value is $\mathbf{e} = \alpha(1-\mathbf{p})/2^{k-1}$ where $1 - \mathbf{p} \doteq \eta \doteq e^{-y}$ is given approximately by the lower curve in panel (B). At this precision it is consistent with the replica symmetric (RS) solution (cf. the estimate of [17]). A more precise comparison between RS and 1RSB is given in the full version of the paper.

To show that $\mathfrak{S}$ is decreasing in y, we will in fact show that $\mathfrak{S}'(y) = y\mathbf{e}'(y) = -y\mathfrak{F}''(y) < 0$. In the full version of the paper, we also review the physical interpretation of $\mathfrak{F}''(y)$.

We remark that the estimate (36) is a rather lossy approximation of $\mathbf{e}_{\text{1RSB}}(\alpha)$. In fact, on its own it does not carry more

information than the first moment [17] bound: observe from (6) that $\eta = \eta(\mathbf{p}) = (1-\mathbf{p})[1+O(\mathbf{p}/2^k)]$. Substituting into (7) gives

$$\frac{1}{\mathbf{c}(\mathbf{p})} = \left(1-\eta+\eta\ln\eta\right)\left\{1+O\left(\frac{1}{e^{\Omega(k)}}\right)\right\} = \left(1-\frac{1+y}{e^y}\right)\left\{1+O\left(\frac{1}{e^{\Omega(k)}}\right)\right\}, \tag{38}$$

simply by taking $y = -\ln\eta$. Thus more care is needed to obtain a comparison such as (12) with the first moment. Towards this end, let us comment briefly on what Proposition 7.2 implies for $\dot{q}(\mathbf{f})$. Recall from Proposition 7.1 that $\dot{q} = \dot{q}_y \in \mathcal{M}^\gamma$, meaning (see (29)) that

$$\begin{aligned}\dot{q}(\mathbf{f}) &\asymp \frac{2^{-k\gamma/2}}{(\max\{\mathbf{c}ke^{-y/2},1\})^{1/2}} \\ &\asymp \frac{1}{2^{k\gamma/2}}\left(\min\left\{\frac{e^{y/2}\min\{1,y^2\}}{k},1\right\}\right)^{1/2} \\ &\le O\left(\frac{1}{2^{k\gamma/2}}\right)\end{aligned}$$

for $\gamma = \gamma(y)$. It follows from (27) that $\mathbf{c}ke^{-y/2} = O(1)$ if and only if $y \ge 2\ln k - O(1)$. For such $y_\star \ge 2\ln k - O(1)$, the estimate (36) implies

$$\mathbf{c} = \Gamma(y_\star)\left(1-\frac{1+y_\star}{e^{y_\star}}\right)^{-1} = 1 - \frac{O(\ln k)}{k^2},$$

meaning α is only slightly above the satisfiability threshold. In this regime

$$\begin{aligned}\frac{k\gamma(y_\star)}{2} &= \frac{k\gamma(y_\star)}{2\Gamma(y_\star)}\Gamma(y_\star) \\ &= k\left\{1+O\left(\frac{1}{e^{y_\star/2}}\right)\right\}\left\{1+O\left(\frac{1}{e^{\Omega(k)}}\right)\right\} = k + O(1),\end{aligned}$$

so $\dot{q}(\mathbf{f}) \asymp 2^{-k}$. This is consistent with estimates slightly below the satisfiability threshold obtained by [7].

We now discuss $y \le 2\ln k + O(1)$. In general, for any fixed $\mathbf{c}$ the value $\Gamma(y)$ is strictly increasing in y, therefore $y_\star$ must be roughly decreasing with $\mathbf{c}$ (modulo the error in the estimate (36)). The ratio $\gamma(y)/\Gamma(y)$ is a function of y alone, and is increasing in y. Therefore, as $\mathbf{c}$ increases, $\gamma(y_\star) \doteq \gamma(y_\star)/\Gamma(y_\star)$ decreases smoothly, from $\gamma \doteq 2$ to $\gamma \doteq 1$ (Figure 1b). For $\Omega(1/k^2) \le y \le 2\ln k + O(1)$ we have $\dot{q}(\mathbf{f}) = k^{O(1)}/2^{k\gamma/2}$, which is roughly increasing as $\mathbf{c}$ decreases if we ignore the $k^{O(1)}$ factor. Finally, if $y = O(1/k^2)$ then

$$\begin{aligned}\frac{k\gamma(y_\star)}{2} &= \frac{k\gamma(y_\star)}{2\Gamma(y_\star)}\Gamma(y_\star) \\ &= \frac{k}{2}\left\{1+O(y)\right\}\left\{1+O\left(\frac{1}{e^{\Omega(k)}}\right)\right\} = \frac{k}{2} + O(1),\end{aligned}$$

so in this regime we have

$$\dot{q}(\mathbf{f}) \asymp \frac{y}{(2^k k)^{1/2}} \asymp \frac{1}{(2^k\mathbf{c}k)^{1/2}} \asymp \frac{1}{d^{1/2}}, \tag{39}$$

which matches with (12).

VIII. Explicit Gardner threshold

We now describe the exact predicted threshold α_{Ga} for the stability of the 1RSB solution. Recall the loose calculation (13) of the branching rate of linked frees. One can refine this by considering the rate of "bug proliferation" [27], [28] in the warning model: if a warning incoming to a vertex is changed, it may change an outgoing warning, and one can calculate the branching rate of this process. Explicitly, let

$$\left(\dot{\mathbf{w}}_{ai} : 2 \le a \le d, 2 \le i \le k\right) \equiv (\dot{\mathbf{w}}_j)_{1\le j\le \boldsymbol{b}} \equiv \dot{\mathbf{w}}_{1:\boldsymbol{b}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}^{\boldsymbol{b}}$$

where we have abused notation and made the identification $\dot{\mathbf{w}}_{ai} \equiv \dot{\mathbf{w}}_{(a-2)(k-1)+(i-1)}$. Recall the mappings $\dot{\mathsf{WP}}$ and $\hat{\mathsf{WP}}$ defined in (14) and (15). Define $\hat{\mathbf{w}}_a \equiv \hat{\mathsf{WP}}(\dot{\mathbf{w}}_{a,2},\ldots,\dot{\mathbf{w}}_{a,k})$ for each $2 \le a \le d$, and then let

$$\begin{aligned}\mathsf{WP}(\dot{\mathbf{w}}_{1:\boldsymbol{b}}) &\equiv \dot{\mathsf{WP}}(\hat{\mathbf{w}}_2,\ldots,\hat{\mathbf{w}}_d) \\ \varphi(\dot{\mathbf{w}}_{1:\boldsymbol{b}}) &\equiv \dot{\varphi}(\hat{\mathbf{w}}_2,\ldots,\hat{\mathbf{w}}_d)\end{aligned}$$

Let $\dot{q}_y$ be as given by Proposition 7.1. Then, for $\dot{\mathbf{v}},\dot{\mathbf{r}},\dot{\mathbf{w}},\dot{\mathbf{s}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}$, let

$$\boldsymbol{B}_{\dot{\mathbf{v}}\dot{\mathbf{r}},\dot{\mathbf{w}}\dot{\mathbf{s}}} \equiv \frac{\displaystyle\sum_{\dot{\underline{\mathbf{w}}}_\circ} \mathbf{1}\left\{\begin{matrix}\dot{\mathbf{v}} = \mathsf{WP}(\dot{\mathbf{w}},\dot{\mathbf{w}}_{2:\boldsymbol{b}}) \\ \dot{\mathbf{r}} = \mathsf{WP}(\dot{\mathbf{s}},\dot{\mathbf{w}}_{2:\boldsymbol{b}})\end{matrix}\right\}\frac{\dot{q}_y(\dot{\mathbf{w}})\prod_{i=2}^{\boldsymbol{b}}\dot{q}_y(\dot{\mathbf{w}}_i)}{\exp\{y\varphi(\dot{\mathbf{s}},\dot{\mathbf{w}}_{2:\boldsymbol{b}})\}}}{\displaystyle\sum_{\dot{\mathbf{w}}_{1:\boldsymbol{b}}} \mathbf{1}\left\{\dot{\mathbf{v}} = \mathsf{WP}(\dot{\mathbf{w}}_{1:\boldsymbol{b}})\right\}\frac{\prod_{i=1}^{\boldsymbol{b}}\dot{q}_y(\dot{\mathbf{w}}_i)}{\exp\{y\varphi(\dot{\mathbf{w}}_{1:\boldsymbol{b}})\}}}. \tag{40}$$

This defines a 9×9 matrix $\boldsymbol{B}$, which is the **stability matrix** for our model. We let $\boldsymbol{B}_{\neq}$ be the 6×6 submatrix with row and column indices in $\{(\dot{\mathbf{w}},\dot{\mathbf{s}}) : \dot{\mathbf{w}} \neq \dot{\mathbf{s}}\}$, and let $\lambda \equiv \lambda_y(\alpha)$ be the largest eigenvalue of $\boldsymbol{B}_{\neq}$. The physics literature [27], [28] proposes that the 1RSB solution is correct as long as $\boldsymbol{b}\lambda_y(\alpha)$ (a refinement of (13)) is less than one at $y = y_\star(\alpha)$. We extract its large-k behavior in the following:

Proposition 8.1 (proved in full version): The **Gardner threshold** α_{Ga} can be formally defined as

$$\alpha_{\text{Ga}} \equiv \sup\left\{\alpha \le \frac{4^k}{k} : \boldsymbol{b}\lambda_{y_\star(\alpha)}(\alpha) \le 1\right\},$$

where $y_\star(\alpha)$ is the root given by Proposition 7.1. The large-k behavior is given by $\alpha_{\text{Ga}} \asymp 4^k/k^3$.

IX. Interpolation bound

As mentioned before, our proof of Theorem 1.1 is based on a general interpolation upper bound, in the spirit of [30]–[34]. The precise bound that we use, as we now describe, is a generalization of a similar result in [8]. Let Ω be the space of probability measures on $\{\mathbf{0},\mathbf{1},\mathbf{f}\}$. We write ρ for elements

of Ω, and Q for probability measures over Ω. Similarly as above, we will abuse notation and write

$$\left(\dot{\mathbf{w}}_{ai} : 1 \le a \le d, 2 \le i \le k\right) \equiv (\dot{\mathbf{w}}_j)_{1 \le j \le D} \equiv \dot{\mathbf{w}}_{1:D}$$

where $D \equiv d(k-1)$. Let $\hat{\mathbf{w}}_a \equiv \hat{\mathsf{WP}}(\dot{\mathbf{w}}_{a,2}, \ldots, \dot{\mathbf{w}}_{a,k})$ and $\varphi(\dot{\mathbf{w}}_{1:D}) \equiv \dot{\varphi}_d(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_d)$. Define

$$\mathcal{G}(y_1, y_2, Q) \tag{41}$$
$$= \int \left\{ \sum_{\dot{\mathbf{w}}_{1:k}} \frac{\prod_{j=1}^k \rho_j(\dot{\mathbf{w}}_j)}{\exp(y_2 \hat{\varphi}(\dot{\mathbf{w}}_{1:k}))} \right\}^{y_1/y_2} \prod_{j=1}^k dQ(\rho_j)\,,$$

$$\mathcal{W}(y_1, y_2, Q) \tag{42}$$
$$= \int \left\{ \sum_{\dot{\mathbf{w}}_{1:D}} \frac{\prod_{i=1}^D \rho_i(\dot{\mathbf{w}}_i)}{\exp(y_2 \varphi(\dot{\mathbf{w}}_{1:D}))} \right\}^{y_1/y_2} \prod_{i=1}^D dQ(\rho_i)\,.$$

For $0 \le y_1 \le y_2$, the **zero-temperature 2RSB functional** is defined by

$$\Phi^{\text{2RSB}}(y_1, y_2, Q) \equiv \frac{1}{y_1} \ln \mathcal{W}(y_1, y_2, Q) - \frac{\alpha(k-1)}{y_1} \ln \mathcal{G}(y_1, y_2, Q)\,. \tag{43}$$

A heuristic derivation of Φ^{2RSB} is presented in more detail in the full version of the paper, but we briefly describe it here. For simplicity assume $\alpha \equiv d/k$ is an integer, and let $\mathscr{G}_N$ be an instance of random d-regular k-NAE-SAT on N variables. Remove $\alpha(k-1)$ clauses and their incident edges at random, and call the resulting graph $\mathscr{G}_{N+1/2}$: it is still a k-NAE-SAT instance on N variables, but is no longer d-regular since some variables have open "slots" (missing edges). Then introduce a new variable $v \equiv v_{N+1}$, together with d new clauses. For each new clause, add one new edge connecting the clause to v, and $k-1$ new edges connecting the clause to the open "slots" in $\mathscr{G}_{N+1/2}$. Then the resulting graph $\mathscr{G}_{N+1}$ is an instance of random d-regular k-NAE-SAT on $N+1$ variables. For $\beta \ge 0$ we can consider

$$\mu_\beta(\underline{x}) \equiv \frac{\exp\{-\beta \mathscr{H}_{N+1/2}(\underline{x})\}}{Z_{N+1/2}(\beta)} \tag{44}$$

where $Z_{N+1/2}(\beta)$ is the normalizing constant that makes μ_β a probability measure over $\underline{x} \in \{\mathbf{0}, \mathbf{1}\}^N$. The structure of μ_β is not known. However, by analogy with other models [38]–[42], a natural simplifying assumption is that it has a **hierarchichal (ultrametric) structure** with Poisson–Dirichlet weights on each level of the hierarchy. This means that the ℓ-point marginals of μ_β, for bounded ℓ, converge in the large-N limit to an explicit form: for a two-level hierarchy,

$$\mu_\beta(x_1, \ldots, x_\ell) \approx \sum_{s,t \ge 1} \nu_{st} \prod_{i=1}^{\ell} w_{st,i}(x_i)\,, \tag{45}$$

where the $w_{st,i}$ are sampled recursively as follows. Let $\mathscr{P}_0 \equiv \mathscr{P}$ be the space of probability measures over $\{\mathbf{0}, \mathbf{1}\}$, and for $r \ge 1$ let $\mathscr{P}_r$ be the space of probability measures over $\mathscr{P}_{r-1}$. Let $Q_\beta \in \mathscr{P}_2$. Let $(r_{s,i})_{s,i}$ be i.i.d. samples from law Q_β. For each i and each s, let $(w_{st,i})_{t \ge 1}$ be a sequence of i.i.d. samples from $r_{s,i}$. Note $r_{s,i} \in \mathscr{P}_1$ so $w_{st,i} \in \mathscr{P}$. Independently, $(\nu_{st})_{s,t \ge 1}$ are random weights sampled from the law of a **Ruelle probability cascade** (RPC) with parameters $0 < m_1 < m_2 < 1$ — a two-level version of the standard Poisson–Dirichlet process (see [43, Ch. 2] and the full version of this paper). Under assumption (45), and taking $\beta \to \infty$ with $m_i \beta \to y_i$, one has

$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln \frac{Z_N(\beta)}{Z_{N+1/2}(\beta)} \approx \frac{\alpha(k-1)}{y_1} \ln \mathcal{G}(y_1, y_2, Q)\,,$$
$$\lim_{\beta \to \infty} \frac{1}{\beta} \ln \frac{Z_{N+1}(\beta)}{Z_{N+1/2}(\beta)} \approx \frac{1}{y_1} \ln \mathcal{W}(y_1, y_2, Q)\,,$$

where Q is a probability measure over Ω, obtained as a projection of Q_β. The basic idea is as follows: a.

1) Project $w \in \mathscr{P}$ to $\mathbf{w} \in \{\mathbf{0}, \mathbf{1}, \mathbf{f}\}$ where $\{w \text{ near } \mathbf{1_0}\}$ maps to $\mathbf{w} = \mathbf{0}$, $\{w \text{ near } \mathbf{1_1}\}$ maps to $\mathbf{w} = \mathbf{1}$, and the remaining $w \in \mathscr{P}$ map to $\mathbf{w} = \mathbf{f}$. Denote this mapping $\pi : \mathscr{P} \to \{\mathbf{0}, \mathbf{1}, \mathbf{f}\}$.
2) Project $r \in \mathscr{P}_1$ to $\rho \in \Omega$ via the pushforward, $\rho(\mathbf{w}) = (\pi_\sharp r)(\mathbf{w}) = r(\pi^{-1}(\mathbf{w}))$.
3) Project $Q_\beta \in \mathscr{P}_2$ to a probability measure Q over Ω via another pushforward, $Q = (\pi_\sharp)_\sharp Q_\beta = Q_\beta \circ (\pi_\sharp)^{-1}$.

The details are given in the full version of the paper. Combining the above relations gives the heuristic approximation

$$-\mathbf{e}_\star = \lim_{\beta \to \infty} \frac{1}{N\beta} \ln Z_N(\beta) = \lim_{\beta \to \infty} \frac{1}{\beta} \ln \frac{Z_{N+1}(\beta)}{Z_N(\beta)} = \Phi^{\text{2RSB}}(y_1, y_2, Q)\,.$$

The following proposition shows that one side of the approximation can be made rigorous:

Proposition 9.1 (proved in full version): For any parameters $0 \le y_1 \le y_2$ and any probability measure Q over Ω, we have a corresponding zero-temperature 2RSB bound $-\mathbf{e}_\bullet \le \Phi^{\text{2RSB}}(y_1, y_2, Q)$.

The detailed heuristic derivation of Φ^{2RSB}, as well as the proof of Proposition 9.1, are given in the full version of the paper. There are two simple ways in which Φ^{2RSB} can degenerate: I.

1) The probability measure Q is fully supported on a single element $\rho \in \Omega$. In this case $\Phi^{\text{2RSB}}(y_1, y_2, Q)$ depends only on y_2 and ρ, so we can define $\Phi^{\text{2RSB}}(y_1, y_2, Q) \equiv \Phi^{\text{1RSB}}(y_2, \rho)$.
2) The probability measure Q decomposes as $Q = \rho_{\mathbf{0}} Q_{\mathbf{0}} + \rho_{\mathbf{1}} Q_{\mathbf{1}} + \rho_{\mathbf{f}} Q_{\mathbf{f}}$ where each $Q_{\dot{\mathbf{w}}}$ is fully supported on the single element $\mathbf{1}_{\dot{\mathbf{w}}} \in \Omega$. In this case we have $\Phi^{\text{2RSB}}(y_1, y_2, Q) = \Phi^{\text{1RSB}}(y_1, \rho)$.

Let $\dot{q}_y$ be the solution of Proposition 7.1, and define

$$Q_2 \equiv Q_{2,y} \equiv \sum_{\dot{\mathbf{w}} \in \{\mathbf{0},\mathbf{1},\mathbf{f}\}} \dot{q}_y(\dot{\mathbf{w}}) Q_{\dot{\mathbf{w}}} \tag{46}$$

where each $Q_{\dot{\mathbf{w}}}$ is the unit mass on $\mathbf{1}_{\dot{\mathbf{w}}} \in \Omega$. One can verify by straightforward algebraic manipulations that

$$\Phi^{\text{1RSB}}(y, \dot{q}_y) = \frac{\mathfrak{F}(y)}{y}\,.$$

Thus an immediate consequence of Proposition 9.1 is that $-\mathbf{e}_\bullet$ is upper bounded by

$$\Phi^{\text{2RSB}}(y, y, Q_{2,y}) = \Phi^{\text{1RSB}}(y, \dot{q}_y) = \frac{\mathfrak{F}(y)}{y}\,, \tag{47}$$

throughout the range of y where $\dot{q}_y$ is defined. It has been observed in the physics literature [27], [28] that linearizing the stationarity equations for the functional Φ^{2RSB} (equivalently, the **2RSB cavity equations**)

$$\frac{\partial \Phi^{\text{2RSB}}(y_1, y_2, Q)}{\partial Q} = 0$$

around $Q = Q_2$ gives rise to the stability matrix $\boldsymbol{B}_{\neq}$ introduced in §VIII. To prove Theorem 1.1, we show that an explicit perturbation of (y, y, Q_2) decreases the value of Φ^{2RSB} as soon as the top eigenvalue of $\boldsymbol{B}_{\neq}$ exceeds $1/\boldsymbol{b}$. While the physics literature certainly hints that this would be the case, to our knowledge this rigorous connection between the Gardner eigenvalue and the stability of the 2RSB functional has not been previously established.

Full version

The full version of this paper is arXiv:1904.08891.

Acknowledgment

We thank Florent Krzakala, Andrea Montanari, Guilhem Semerjian, and Allan Sly for many helpful conversations. This research was supported in part by NSF CAREER grant DMS-1752728.

References

[1] F. Krzakala, 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.*, vol. 104, no. 25, pp. 10 318–10 323, 2007.

[2] D. Achlioptas and F. Ricci-Tersenghi, "On the solution-space geometry of random constraint satisfaction problems," in *Proc. 38th STOC*. ACM, 2006, pp. 130–139.

[3] D. Achlioptas and A. Coja-Oghlan, "Algorithmic barriers from phase transitions," in *Proc. 49th FOCS*. IEEE, 2008, pp. 793–802.

[4] A. Montanari, R. Restrepo, and P. Tetali, "Reconstruction and clustering in random constraint satisfaction problems," *SIAM J. Discrete Math.*, vol. 25, no. 2, pp. 771–808, 2011. [Online]. Available: https://doi.org/10.1137/090755862

[5] D. Achlioptas and C. Moore, "The asymptotic order of the random k-SAT threshold," in *Proc. 43rd FOCS*. IEEE, 2002, pp. 779–788.

[6] A. Coja-Oghlan and K. Panagiotou, "Catching the k-NAESAT threshold," in *Proc. 44th STOC*. ACM, 2012, pp. 899–908.

[7] J. Ding, A. Sly, and N. Sun, "Satisfiability threshold for random regular NAE-SAT," *Comm. Math. Phys.*, vol. 341, no. 2, pp. 435–489, 2016. [Online]. Available: https://doi.org/10.1007/s00220-015-2492-8

[8] A. Sly, N. Sun, and Y. Zhang, "The number of solutions for random regular NAE-SAT," in *Proc. 57th FOCS*. IEEE, 2016, pp. 724–731.

[9] A. Coja-Oghlan, F. Krzakala, W. Perkins, and L. Zdeborová, "Information-theoretic thresholds from the cavity method," *Adv. Math.*, vol. 333, pp. 694–795, 2018. [Online]. Available: https://doi.org/10.1016/j.aim.2018.05.029

[10] U. Feige, "Relations between average case complexity and approximation complexity," in *Proc. 34th STOC*. ACM, New York, 2002, pp. 534–543. [Online]. Available: https://doi.org/10.1145/509907.509985

[11] A. Coja-Oghlan, A. Goerdt, and A. Lanka, "Strong refutation heuristics for random k-SAT," *Combin. Probab. Comput.*, vol. 16, no. 1, pp. 5–28, 2007. [Online]. Available: https://doi.org/10.1017/S096354830600784X

[12] S. R. Allen, R. O'Donnell, and D. Witmer, "How to refute a random CSP," in *Proc. 56th FOCS*. IEEE, 2015, pp. 689–708.

[13] D. Grigoriev, "Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity," *Theoret. Comput. Sci.*, vol. 259, no. 1-2, pp. 613–622, 2001. [Online]. Available: https://doi.org/10.1016/S0304-3975(00)00157-2

[14] G. Schoenebeck, "Linear level Lasserre lower bounds for certain k-CSPs," in *Proc. 49th FOCS*. IEEE, 2008, pp. 593–602.

[15] P. Raghavendra, S. Rao, and T. Schramm, "Strongly refuting random CSPs below the spectral threshold," in *Proc. 49th STOC*. ACM, New York, 2017, pp. 121–131.

[16] P. K. Kothari, R. Mori, R. O'Donnell, and D. Witmer, "Sum of squares lower bounds for refuting any CSP," in *Proc. 49th STOC*. ACM, 2017, pp. 132–145.

[17] D. Achlioptas, A. Naor, and Y. Peres, "On the maximum satisfiability of random formulas," *J. ACM*, vol. 54, no. 2, 2007. [Online]. Available: https://doi.org/10.1145/1219092.1219098

[18] S. Sen, "Optimization on sparse random hypergraphs and spin glasses," *Random Structures Algorithms*, vol. 53, no. 3, pp. 504–536, 2018. [Online]. Available: https://doi.org/10.1002/rsa.20774

[19] D. Panchenko, "On the k-SAT model with large number of clauses," *Random Struct. Algor.*, vol. 52, no. 3, pp. 536–542, 2018. [Online]. Available: https://doi.org/10.1002/rsa.20748

[20] A. Auffinger and W.-K. Chen, "Parisi formula for the ground state energy in the mixed p-spin model," *Ann. Probab.*, vol. 45, no. 6B, pp. 4617–4631, 2017. [Online]. Available: https://doi.org/10.1214/16-AOP1173

[21] L. Leuzzi and G. Parisi, "The k-SAT problem in a simple limit," *J. Stat. Phys.*, vol. 103, no. 5-6, pp. 679–695, 2001.

[22] A. Crisanti, L. Leuzzi, and G. Parisi, "The 3-SAT problem with large number of clauses in the ∞-replica symmetry breaking scheme," *J. Phys. A*, vol. 35, no. 3, p. 481, 2002.

[23] A. Auffinger, W.-K. Chen, and Q. Zeng, "The SK model is full-step replica symmetry breaking at zero temperature," arXiv:1703.06872, 2017.

[24] L. Addario-Berry and P. Maillard, "The algorithmic hardness threshold for continuous random energy models," arXiv:1810.05129, 2018.

[25] E. Subag, "Following the ground-states of full-RSB spherical spin glasses," arXiv:1812.04588, 2018.

[26] A. Montanari, "Optimization of the Sherrington–Kirkpatrick Hamiltonian," arXiv:1812.10897, 2018.

[27] A. Montanari and F. Ricci-Tersenghi, "On the nature of the low-temperature phase in discontinuous mean-field spin glasses," *Eur. Phys. J. B*, vol. 33, no. 3, pp. 339–346, 2003.

[28] F. Krzakala, A. Pagnani, and M. Weigt, "Threshold values, stability analysis, and high-q asymptotics for the coloring problem on random graphs," *Phys. Rev. E*, vol. 70, no. 4, p. 046705, 2004.

[29] E. Gardner, "Spin glasses with p-spin interactions," *Nucl. Phys. B*, vol. 257, pp. 747–765, 1985.

[30] F. Guerra and F. L. Toninelli, "The thermodynamic limit in mean field spin glass models," *Comm. Math. Phys.*, vol. 230, no. 1, pp. 71–79, 2002. [Online]. Available: https://doi.org/10.1007/s00220-002-0699-y

[31] F. Guerra, "Broken replica symmetry bounds in the mean field spin glass model," *Comm. Math. Phys.*, vol. 233, no. 1, pp. 1–12, 2003. [Online]. Available: https://doi.org/10.1007/s00220-002-0773-5

[32] D. Panchenko and M. Talagrand, "Bounds for diluted mean-fields spin glass models," *Probab. Theory Related Fields*, vol. 130, no. 3, pp. 319–336, 2004. [Online]. Available: https://doi.org/10.1007/s00440-004-0342-2

[33] M. Bayati, D. Gamarnik, and P. Tetali, "Combinatorial approach to the interpolation method and scaling limits in sparse random graphs," *Ann. Probab.*, vol. 41, no. 6, pp. 4080–4115, 2013. [Online]. Available: https://doi.org/10.1214/12-AOP816

[34] D. Gamarnik, "Right-convergence of sparse random graphs," *Probab. Theory Related Fields*, vol. 160, no. 1-2, pp. 253–278, 2014. [Online]. Available: https://doi.org/10.1007/s00440-013-0528-6

[35] A. Coja-Oghlan, C. Efthymiou, N. Jaafari, M. Kang, and T. Kapetanopoulos, "Charting the replica symmetric phase," *Comm. Math. Phys.*, vol. 359, no. 2, pp. 603–698, 2018. [Online]. Available: https://doi.org/10.1007/s00220-018-3096-x

[36] A. Dembo, A. Montanari, and S. Sen, "Extremal cuts of sparse random graphs," *Ann. Probab.*, vol. 45, no. 2, pp. 1190–1217, 2017. [Online]. Available: https://doi.org/10.1214/15-AOP1084

[37] M. Mézard and A. Montanari, *Information, physics, and computation*, ser. Oxford Graduate Texts. Oxford University Press, Oxford, 2009. [Online]. Available: https://doi.org/10.1093/acprof:oso/9780198570837.001.0001

[38] B. Derrida, "Random-energy model: An exactly solvable model of disordered systems," *Phys. Rev. B*, vol. 24, no. 5, p. 2613, 1981.

[39] ——, "A generalization of the random energy model which includes correlations between energies," *J. Phys. Lett.–Paris*, vol. 46, no. 9, pp. 401–407, 1985.

[40] D. Ruelle, "A mathematical reformulation of Derrida's REM and GREM," *Comm. Math. Phys.*, vol. 108, no. 2, pp. 225–239, 1987. [Online]. Available: http://projecteuclid.org/euclid.cmp/1104116461

[41] G. Parisi, "Infinite number of order parameters for spin-glasses," *Phys. Rev. Lett.*, vol. 43, no. 23, p. 1754, 1979.

[42] D. Panchenko, "The Parisi ultrametricity conjecture," *Ann. of Math. (2)*, vol. 177, no. 1, pp. 383–393, 2013. [Online]. Available: https://doi.org/10.4007/annals.2013.177.1.8

[43] ——, *The Sherrington–Kirkpatrick model*, ser. Springer Monographs in Mathematics. Springer, New York, 2013. [Online]. Available: https://doi.org/10.1007/978-1-4614-6289-7