The Annals of Probability
2013, Vol. 41, No. 6, 4080–4115
DOI: 10.1214/12-AOP816

COMBINATORIAL APPROACH TO THE INTERPOLATION METHOD AND SCALING LIMITS IN SPARSE RANDOM GRAPHS

BY MOHSEN BAYATI[1], DAVID GAMARNIK[2] AND PRASAD TETALI[3]

Stanford Univeristy, MIT and Georgia Tech

We establish the existence of free energy limits for several combinatorial models on Erdös–Rényi graph $\mathbb{G}(N, \lfloor cN \rfloor)$ and random r-regular graph $\mathbb{G}(N, r)$. For a variety of models, including independent sets, MAX-CUT, coloring and K-SAT, we prove that the free energy both at a positive and zero temperature, appropriately rescaled, converges to a limit as the size of the underlying graph diverges to infinity. In the zero temperature case, this is interpreted as the existence of the scaling limit for the corresponding combinatorial optimization problem. For example, as a special case we prove that the size of a largest independent set in these graphs, normalized by the number of nodes converges to a limit w.h.p. This resolves an open problem which was proposed by Aldous (Some open problems) as one of his six favorite open problems. It was also mentioned as an open problem in several other places: Conjecture 2.20 in Wormald [In *Surveys in Combinatorics*, 1999 (*Canterbury*) (1999) 239–298 Cambridge Univ. Press]; Bollobás and Riordan [*Random Structures Algorithms* **39** (2011) 1–38]; Janson and Thomason [*Combin. Probab. Comput.* **17** (2008) 259–264] and Aldous and Steele [In *Probability on Discrete Structures* (2004) 1–72 Springer].

Our approach is based on extending and simplifying the interpolation method of Guerra and Toninelli [*Comm. Math. Phys.* **230** (2002) 71–79] and Franz and Leone [*J. Stat. Phys.* **111** (2003) 535–564]. Among other applications, this method was used to prove the existence of free energy limits for Viana–Bray and K-SAT models on Erdös–Rényi graphs. The case of zero temperature was treated by taking limits of positive temperature models. We provide instead a simpler combinatorial approach and work with the zero temperature case (optimization) directly both in the case of Erdös–Rényi graph $\mathbb{G}(N, \lfloor cN \rfloor)$ and random regular graph $\mathbb{G}(N, r)$. In addition we establish the large deviations principle for the satisfiability property of the constraint satisfaction problems, coloring, K-SAT and NAE-K-SAT, for the $\mathbb{G}(N, \lfloor cN \rfloor)$ random graph model.

1. Introduction.

Consider two random graph models on nodes $[N] \triangleq \{1, \dots, N\}$, the Erdös–Rényi graph $\mathbb{G}(N, M)$ and the random r-regular graph $\mathbb{G}(N, r)$. The first model is obtained by generating M edges of the $N(N-1)/2$ possible

Received October 2011; revised October 2012.

[1]Supported in part by MBA Class of 1969 Faculty Scholarship.
[2]Supported in part by the NSF Grants CMMI-1031332.
[3]Supported in part by the NSF Grants DMS-07-01043 and CCR-0910584.

MSC2010 subject classifications. Primary 05C80, 60C05; secondary 82-08.
Key words and phrases. Constraint satisfaction problems, partition function, random graphs.

edges uniformly at random without replacement. Specifically, assume $M = \lfloor cN \rfloor$ where $c > 0$ is a constant (does not grow with N). The second model $\mathbb{G}(N, r)$ is a graph chosen uniformly at random from the space of all r-regular graphs on N nodes, where the integer r is a fixed integer constant. Consider the size $|\mathcal{I}_N|$ of a largest independent set $\mathcal{I}_N \subset [N]$ in $\mathbb{G}(N, \lfloor cN \rfloor)$ or $\mathbb{G}(N, r)$. It is straightforward to see that $|\mathcal{I}_N|$ grows linearly with N. It was conjectured in several papers including Conjecture 2.20 in [8, 22, 28], as well as [25] and [3], that $|\mathcal{I}_N|/N$ converges in probability as $N \to \infty$. Additionally, this problem was listed by Aldous as one of his six favorite open problems [2]. (For a new collection of Aldous's favorite open problems, see [1].) The fact that the actual value of $|\mathcal{I}_N|$ concentrates around its mean follows from a standard Azuma-type inequality. However, a real challenge is to show that the expected value of $|\mathcal{I}_N|$ normalized by N does not fluctuate for large N.

This conjecture is in fact just one of a family of similar conjectures. Consider, for example, the random MAX-K-SAT problem—the problem of finding the largest number of satisfiable clauses of size K in a uniformly random instance of a K-SAT problem on N variables with cN clauses. This problem can be viewed as an optimization problem over a sparse random *hypergraph*. A straightforward argument shows that asymptotically as $N \to \infty$, at least $1 - 2^{-K}$ fraction of the clauses can be satisfied with high probability (w.h.p.). Indeed any random assignment of variables satisfies each clause with probability $1-2^{-K}$. It was conjectured in [10] that the proportion of the largest number of satisfiable clauses has a limit w.h.p. as $N \to \infty$. As another example, consider the problem of partial q-coloring of a graph: finding a q-coloring of nodes which maximizes the total number of properly colored edges. It is natural to conjecture again that value of this maximum has a scaling limit w.h.p. (though we are not aware of any papers explicitly stating this conjecture).

Recently a powerful rigorous statistical physics method was introduced by Guerra and Toninelli [23] and further developed by Franz and Leone [16], Franz, Leone and Toninelli [17], Panchenko and Talagrand [27] and Montanari [26] in the context of the theory of spin glasses. The method is based on an ingenious interpolation between a random hypergraph model on N nodes on the one hand, and a disjoint union of random hypergraph models on N_1 and N_2 nodes, on the other hand, where $N = N_1 + N_2$. Using this method it is possible to show for certain spin glass models on random hypergraphs, that when one considers the expected log-partition function, the derivative of the interpolation function has a definite sign at *every value* of the interpolation parameter. As a result the expected log-partition function of the N-node model is larger (or smaller depending on the details of the model) than the sum of the corresponding expected log-partition functions on N_1 and N_2-node models. This super(sub)-additivity property is used to argue the existence of the (thermodynamic) limit of the expected log-partition function scaled by N. From this property the existence of the scaling limits for the ground states (optimization problems described above) can also be shown by taking a limit as positive temperature approaches zero temperature. In [16], the method was used to prove the scaling limit of log-partition functions corresponding to random K-SAT model for even K, and also for the so-called Viana–Bray models with random symmetric Hamiltonian functions. The case of odd K was apparently resolved later using the same method [18].

Results and technical contributions. The goal of the present work is to simplify and extend the applicability of the interpolation method, and we do this in several important ways. First, we extend the interpolation method to a variety of models on Erdös–Rényi graphs not considered before. Specifically, we consider independent set, MAX-CUT, Ising, graph coloring (henceforth referred to as coloring), K-SAT and Not-All-Equal K-SAT (NAE-K-SAT) models. The coloring model, in particular, is of special interest as it is the first nonbinary model to which interpolation method is applied.

Second, we provide a simpler and a more combinatorial interpolation scheme as well as analysis. Moreover, we treat the zero temperature case (optimization problem) directly and separately from the case of the log-partition function, and again the analysis turns out to be substantially simpler. As a result, we prove the existence of the limit of the appropriately rescaled value of the optimization problems in these models, including the independent set problem, thus resolving the open problem stated earlier.

Third, we extend the above results to the case of random regular graphs (and hypergraph ensembles, depending on the model). The case of random regular graphs has been considered before by Franz, Leone and Toninelli [17] for the K-SAT and Viana–Bray models with an *even* number of variables per clause, and Montanari [26] in the context of bounds on the performance of *low density parity check* (LDPC) codes. In fact, both papers consider general degree distribution models. The second of these papers introduces a multi-phase interpolation scheme. In this paper we consider a modification of the interpolation scheme used in [17] and apply it to the same six models we are focusing in the case of Erdös–Rényi graph.

Finally, we prove the large deviation principle for the satisfiability property for coloring, K-SAT and NAE-K-SAT models on Erdös–Rényi graph in the following sense. A well-known satisfiability conjecture [19] states that for each of these models there exists a (model dependent) critical value c^* such that for every $\varepsilon > 0$, when the number of edges (or clauses for a SAT-type problem) is at most $(c^* - \varepsilon)N$, the model is colorable (satisfiable) w.h.p., and when it is at least $(c^* + \varepsilon)N$, it is not colorable (not satisfiable) w.h.p. as $N \to \infty$. Friedgut [19] came close to proving this conjecture by showing that these models exhibit sharp phase transition: there exists a sequence c_N^* such that for every ε, the model is colorable (satisfiable) w.h.p. as $N \to \infty$ when the number of edges (clauses) is at most $(c_N^* - \varepsilon)N$, and is not colorable (satisfiable) w.h.p. when the number of edges (clauses) is at least $(c_N^* + \varepsilon)N$. It is also reasonable to conjecture (which in fact is known to be true in the case $K = 2$), that not only the satisfiability conjecture is valid, but, moreover, the probability of satisfiability $p(c, N)$ decays to zero exponentially fast when $c > c^*$.

In this paper we show that for these three models, namely coloring, K-SAT and NAE-K-SAT, the limit $r(c) \triangleq \lim_{N\to\infty} N^{-1} \log p(c, N)$ exists for every c. Namely, while we do not prove the satisfiability conjecture and the exponential rate of convergence to zero of the satisfiability probability above the critical threshold, we do prove that if the convergence to zero occurs exponentially fast, it does so at a well-defined rate $r(c)$. Assuming the validity of the satisfiability conjecture and the exponential rate of decay to zero above c^*, our result implies that $r(c) = 0$ when $c < c^*$ and $r(c) < 0$ when $c > c^*$. Moreover, we show that our results would imply the satisfiability conjecture, if one could strengthen Friedgut's result as follows: for every $\varepsilon > 0$, $p(c_N^* + \varepsilon, N)$ converges to zero exponentially fast, where c_N^* is the same sequence as in Friedgut's theorem.

Organization of the paper. The remainder of the paper is organized as follows. In the following section we introduce the sparse random (Erdös–Rényi) and random regular (hyper)-graphs and introduce various combinatorial models of interest. Our main results are stated in Section 3. The proofs for the case of Erdös–Rényi graphs are presented in Section 4 for results related to combinatorial optimization, and in Section 5 for results related to the log-partition function. The proofs of results for random regular graphs are presented in Section 6. Several auxiliary technical results are established in the Appendices A and B. In particular we state and prove a simple modification of a classical super-additivity theorem: if a sequence is nearly super-additive, it has a limit after an appropriate normalization.

Notations. We close this section with several notational conventions. $\mathbb{R}(\mathbb{R}_+)$ denotes the set of (nonnegative) real values, and $\mathbb{Z}(\mathbb{Z}_+)$ denotes the set of (nonnegative) integer values. The log function is assumed to be with a natural base. As before, $[N]$ denotes the set of integers $\{1, \dots, N\}$. $O(\cdot)$ stands for standard order of magnitude notation. Specifically, given two positive functions $f(N), g(N)$ defined on $N \in \mathbb{Z}_+$, $f = O(g)$ means $\sup_N f(N)/g(N) < \infty$. Also $f = o(g)$ means $\lim_{N\to\infty} f(N)/g(N) = 0$. Throughout the paper, we treat $[N]$ as a set of nodes, and we consider splitting this into two sets of nodes, namely $[N_1] = \{1, \dots, N_1\}$ and $\{N_1 + 1, \dots, N\}$. For symmetry, with some abuse of notation, it is convenient to denote the second set by $[N_2]$ where $N_2 = N - N_1$. Δ denotes the set-theoretic symmetric difference. $\mathrm{Bi}(N, \theta)$ denotes the binomial distribution with N trials and success probability θ. $\mathrm{Pois}(c)$ denotes a Poisson distribution with parameter c, $\stackrel{d}{=}$ stands for equality in distribution. A sequence of random variables X_N is said to converge to a random variable X with high probability (w.h.p.) if for every $\varepsilon > 0$, $\lim_{N\to\infty} \mathbb{P}(|X_N - X| > \varepsilon) = 0$. This is the usual convergence in probability.

2. Sparse random hypergraphs. Given a set of nodes $[N]$ and a positive integer K, a directed hyperedge is any ordered set of nodes $(i_1, \dots, i_K) \in [N]^K$. An undirected hyperedge is an unordered set of K not necessarily distinct nodes $i_1, \dots, i_K \in [N]$. A directed (undirected) K-uniform hypergraph on the node set $[N]$ is a pair $([N], E)$, where E is a set of directed (undirected) K-hyperedges $E = \{e_1, \dots, e_{|E|}\}$. Here uniformity corresponds to the fact that every hyperedge has precisely K nodes. A hypergraph is called simple if the nodes within each hyperedge e_m, $1 \le m \le |E|$, are distinct and all the hyperedges are distinct. A (directed or undirected) hypergraph is called r-regular if each node $i \in [N]$ appears in exactly r hyperedges. The necessary condition for such a hypergraph to exist is $Nr/K \in \mathbb{Z}_+$. A degree $\Delta_i = \Delta_i(\mathbb{G})$ of a node i is the number of hyperedges containing i. A (partial) matching is a set of hyperedges such that each node belongs to at most one hyperedge. A matching is perfect if every node of the graph belongs to exactly one hyperedge. In this paper we use the terms hypergraph and graph (hyperedge and edge) interchangeably.

In order to address a variety of models in a unified way, we introduce two random directed hypergraph models, namely the Erdös–Rényi random graph model $\mathbb{G}(N, M)$, $M \in \mathbb{Z}_+$, and the random regular graph $\mathbb{G}(N, r)$, $r \in \mathbb{Z}_+$. These two graph models, each consisting of N nodes, are described as follows. The first $\mathbb{G}(N, M, K)$ is obtained by selecting M directed hyperedges uniformly at random with replacement from the space of all $[N]^K$ hyperedges. A variant of this is the *simple* Erdös–Rényi graph also denoted for convenience by $\mathbb{G}(N, M)$, which is obtained by selecting M edges uniformly at random without replacement from the set of all undirected hyperedges each consisting of *distinct* K nodes. In this paper we will consider exclusively the case when $M = \lfloor cN \rfloor$, and c is a positive constant which does not grow with N. In this case the probability distribution of the degree of a typical node is $\mathrm{Pois}(cK) + O(1/N)$. For this reason we will also call it a *sparse* random Erdös–Rényi graph. Often a sparse random Erdös–Rényi graph is defined by including each hyperedge in $[N]^K$ into the hypergraph with probability c/N^{K-1}, and not including it with the remaining probability $1 - c/N^{K-1}$. The equivalence of two models is described using the notion of contiguity and is well described in a variety of books, for example, [4, 24].

The second model $\mathbb{G}(N, r, K)$ is defined to be an r-regular directed K-uniform hypergraph generated uniformly at random from the space of all such graphs. We assume $Nr/K \in \mathbb{Z}_+$, so that the set of such graphs is nonempty. A simple (directed or undirected) version of $\mathbb{G}(N, r, K)$ is defined similarly. In this paper we consider exclusively the case when r is a constant (as a function of N), and we call $\mathbb{G}(N, r, K)$ a *sparse* random regular graph.

REMARK 1. The reason for considering the more general case of hypergraphs is to capture combinatorial models with hyperedges. For example, in the case of K-SAT each clause contains $K \ge 2$ distinct nodes that can be considered as a hyperedge on K nodes (more detail is provided below).

REMARK 2. In all models studied in this paper, except for K-SAT and NAE-K-SAT, K satisfies $K = 2$. Therefore, to simplify the notation we drop the reference to K and throughout the paper use the shorter notation $\mathbb{G}(N, M)$ and $\mathbb{G}(N, r)$ for the two random graph models.

From nonsimple to simple graphs. While it is common to work with simple hypergraphs, for our purpose it is more convenient to establish results for directed nonsimple hypergraphs first. It is well known, however, that both $\mathbb{G}(N, M)$ and $\mathbb{G}(N, r)$ graphs are simple with probability which remains at least a constant as $N \to \infty$, as long as c, r, K are constants. Since we prove statements which hold w.h.p., our results have immediate ramification for simple Erdös–Rényi and regular graphs.

It will be useful to recall the so-called configuration method of constructing the random regular graph [5, 6, 20]. Each node i is associated with r nodes denoted by $j_1^i, \dots, j_r^i$. We obtain a new set of Nr nodes. Consider the K-uniform perfect matching $e_1, \dots, e_{Nr/K}$ generated uniformly at random on this set of nodes. From this set of edges we generate a graph on the original N nodes by projecting each edge to its representative. Namely an edge $(i_1, \dots, i_K)$ is created if and only if there is an edge of the form $(j_{k_1}^{i_1}, \dots, j_{k_K}^{i_K})$ for some $k_1, \dots, k_K \in [r]$. The resulting graph is a random r-regular (not necessarily simple) graph, which we again denote by $\mathbb{G}(N, r)$. From now on when we talk about configuration graph, we have in mind the graph just described on Nr nodes. It is known [24] that with probability bounded away from zero as $N \to \infty$ the resulting graph is in fact simple.

Given a hypergraph $\mathbb{G} = ([N], E)$, we will consider a variety of combinatorial structures on $\mathbb{G}$, which can be defined in a unified way using the notion of a Markov random field (MRF). The MRF is a hypergraph $\mathbb{G}$ together with an alphabet $\chi = \{0, 1, \dots, q - 1\}$, denoted by $[q^-]$, and a set of node and edge potentials $H_i, i \in [N], H_e, e \in E$. A node potential is a function $H_i : [q^-] \to \mathbb{R}$, and an edge potential is a function $H_e : [q^-]^K \to \{-\infty\} \cup \mathbb{R}$. Given a MRF $(\mathbb{G}, [q^-], H_i, H_e, i \in [N], e \in E)$ and any $x \in [q^-]^N$, let

$$H(x) = \sum_{i \in [N]} H_i(x_i) + \sum_{e \in E} H_e(x_e), \qquad H(\mathbb{G}) = \sup_{x \in [q^-]^N} H(x),$$

where $x_e = (x_i, i \in e)$. Namely, $H(x)$ is the value associated with a chosen assignment x, and H is the optimal value, or the groundstate in the statistical physics terminology. In many cases the node and edge potentials will be random functions generated i.i.d.; see examples below.

Associated with a MRF is the Gibbs probability measure $\mu_{\mathbb{G}}$ on the set of node values $[q^-]^N$ defined as follows. Fix a parameter $\lambda > 0$, and assign the probability mass $\mu_{\mathbb{G}}(x) = \lambda^{H(x)}/Z_{\mathbb{G}}$ to every assignment $x \in [q^-]^N$, where $Z_{\mathbb{G}} = \sum_x \lambda^{H(x)}$

is the normalizing *partition function*. Observe that $\lim_{\lambda\to\infty}(\log\lambda)^{-1}\log Z_{\mathbb{G}} = H(\mathbb{G})$. Sometimes one considers $\lambda = \exp(1/T)$ where T is *temperature*. The case $T = 0$, namely $\lambda = \infty$ then corresponds to the zero temperature regime, or equivalently the optimization (groundstate) problem. We distinguish this with a positive temperature case, namely $\lambda < \infty$.

We will consider in this paper a variety of MRF defined on sparse random graphs $\mathbb{G}(N, \lfloor cN \rfloor)$ and $\mathbb{G}(N, r)$. (In the statistical physics literature x_i are called spin values, and the corresponding MRF is called a *diluted spin glass model*.) We now describe some examples of concrete and well-known MRF and show that they fit the framework described above.

Independent set. $K = 2$ and $q = 2$. Define $H_i(1) = 1$, $H_i(0) = 0$ for all $i \in [N]$. Define $H_e(1, 1) = -\infty$, $H_e(1, 0) = H_e(0, 1) = H_e(0, 0) = 0$ for every edge $e = (i_1, i_2)$. Then for every vector $x \in \{0, 1\}^N$ we have $H(x) = -\infty$ if there exists an edge $e_j = (i_1, i_2)$ such that $x_{i_1} = x_{i_2} = 1$ and $H(x) = |\{i : x_i = 1\}|$, otherwise. Equivalently, $H(x)$ takes finite value only on x corresponding to independent sets, and in this case it is the cardinality of the independent set. $H(\mathbb{G})$ is the cardinality of a largest independent set. Note that one can have many independent sets with cardinality $H(\mathbb{G})$.

MAX-CUT. $K = 2$ and $q = 2$. Define $H_i(0) = H_i(1) = 0$. Define $H_e(1, 1) = H_e(0, 0) = 0$, $H_e(1, 0) = H_e(0, 1) = 1$. Every vector $x \in \{0, 1\}^N$ partitions nodes into two subsets of nodes taking values 0 and 1, respectively. $H(x)$ is the number of edges between the two subsets. $H(\mathbb{G})$ is the largest such number, also called maximum cut size. A more general case of this model is q-coloring; see below.

Anti-ferromagnetic Ising model. $K = 2$ and $q = 2$. Fix $\beta > 0$, $B \in \mathbb{R}$. Define $H_i(0) = -B$, $H_i(1) = B$. Define $H_e(1, 1) = H_e(0, 0) = -\beta$, $H_e(1, 0) = H_e(0, 1) = \beta$. It is more common to use alphabet $\{-1, 1\}$ instead of $\{0, 1\}$ for this model. We use the latter for consistency with the remaining models. The parameter B, when it is nonzero represents the presence of an external magnetic field.

q-coloring $K = 2$ and q is arbitrary. $H_i(x) = 0, \forall x \in [q^-]$ and $H_e(x, y) = 0$ if $x = y$ and $H_e(x, y) = 1$ otherwise. Therefore for every $x \in [q^-]^N$, $H(x)$ is the number of properly colored edges, and $H(\mathbb{G})$ is the maximum number of properly colored edges.

Random K-SAT. $K \geq 2$ is arbitrary, $q = 2$. $H_i = 0$ for all $i \in [N]$. The edge potentials H_e are defined as follows. For each edge $e \in E$ generate $a_e = (a_1, \dots, a_K)$ uniformly at random from $\{0, 1\}^K$, independently for all edges. For each edge e set $H_e(a_1, \dots, a_K) = 0$ and $H_e(x) = 1$ for all other $x = (x_1, \dots, x_K)$. Then for every $x \in \{0, 1\}^N$, $H(x)$ is the number of satisfied clauses (hyperedges), and $H(\mathbb{G})$

is the largest number of satisfiable clauses. Often this model is called (random) MAX-K-SAT model. We drop the MAX prefix in the notation.

NAE-K-SAT (*Not-All-Equal-K-SAT*). The setting is as above except now we set $H_e(a_1, \dots, a_K) = H_e(1 - a_1, \dots, 1 - a_K) = 0$ and $H_e(x) = 1$ for all other x for each e.

It is for the K-SAT and NAE-K-SAT models that considering directed, as opposed to undirected, hypergraphs is convenient, as for these models the order of nodes in edges matters. For the remaining models, however, this is not the case.

In several examples considered above we have had only two possible values for the edge potential H_e and one value for the node potential. Specifically, for the cases of coloring, K-SAT and NAE-K-SAT problems, H_e took only values 0 and 1. It makes sense to call instances of such problems "satisfiable" if $H(\mathbb{G}) = |E|$; namely every edge potential takes value 1. In the combinatorial optimization terminology this corresponds to finding a proper coloring, a satisfying assignment and a NAE satisfying assignment, respectively. We let $p(N, M) = \mathbb{P}(H(\mathbb{G}(N, M)) = M)$ denote the probability of satisfiability when the underlying graph is the Erdös–Rényi graph $\mathbb{G}(N, M)$. We also let $p(N, r) = \mathbb{P}(H(\mathbb{G}(N, r)) = rNK^{-1})$ denote the satisfiability probability for a random regular graph $\mathbb{G}(N, r)$.

3. Main results. We now state our main results. Our first set of results concerns the Erdös–Rényi graph $\mathbb{G}(N, \lfloor cN \rfloor)$.

THEOREM 1. *For every $c > 0$, and for every one of the six models described in Section 2, there exists (model dependent) $H(c)$ such that*

$$\lim_{N \to \infty} N^{-1} H\bigl(\mathbb{G}\bigl(N, \lfloor cN \rfloor\bigr)\bigr) = H(c), \tag{1}$$

w.h.p. Moreover, $H(c)$ is a Lipschitz continuous function with Lipschitz constant 1. It is a nondecreasing function of c for MAX-CUT, coloring, K-SAT and NAE-K-SAT models, and is a nonincreasing function of c for the independent set model.

Also for every $c > 0$ there exists $p(c)$ such that

$$\lim_{N \to \infty} N^{-1} \log p\bigl(N, \lfloor cN \rfloor\bigr) = p(c) \tag{2}$$

for coloring, K-SAT and NAE-K-SAT models.

As a corollary, one obtains the following variant of the satisfiability conjecture.

COROLLARY 1. *For coloring, K-SAT and NAE-K-SAT models, there exists a critical value c_H^* such that $H(c) = c$ when $c < c_H^*$ and $H(c) < c$ when $c > c_H^*$. Similarly, there exists c_p^*, such that $p(c) = 0$ when $c < c_p^*$ and $p(c) < 0$ when $c > c_p^*$.*

Namely, there exists a threshold value c^* such that if $c < c^*$ there exists w.h.p. as $N \to \infty$ a *nearly* satisfiable assignment [assignment satisfying all but $o(N)$ clauses], and if $c > c^*$, then w.h.p. as $N \to \infty$, every assignment violates linearly in N many clauses. The interpretation for coloring is similar. The result above was established earlier by the second author for randomly generated linear programming problems, using the local weak convergence and martingale techniques [21]. It would be interesting to see if the same result is obtainable using the interpolation method.

Can one use Corollary 1 to prove the satisfiability conjecture in the precise sense? The answer would be affirmative, provided that a stronger version of Friedgut's result [19] on the sharp thresholds for satisfiability properties holds.

CONJECTURE 1. *For the coloring, K-SAT and NAE-K-SAT models there exists a sequence M_N^* such that for every $\varepsilon > 0$ there exists $\gamma = \gamma(\varepsilon)$ such that $\lim_{N\to\infty} p(N, \lfloor(1-\varepsilon)M_N^*\rfloor) = 1$ and $p(N, \lfloor(1+\varepsilon)M_N^*\rfloor) = O(\exp(-\gamma N))$, for all N.*

In contrast, Friedgut's sharp phase transition result [19] replaces the second part of this conjecture with (a weaker) statement $\lim_{N\to\infty} p(N, \lfloor(1+\varepsilon)M_N^*\rfloor) = 0$. Thus, we conjecture that beyond the phase transition region M_N^*, not only is the model *not satisfiable* w.h.p., but in fact the probability of satisfiability converges to zero exponentially fast. The import of this (admittedly bold) statement is as follows:

Conjecture 1 together with Theorem 1 implies the satisfiability conjecture. Indeed, it suffices to show that c_h^* is the satisfiability threshold. We already know that for every $\varepsilon > 0$, $p(N, \lfloor(1+\varepsilon)c_h^* N) \to 0$, since $H((1+\varepsilon)c_h^*) < (1+\varepsilon)c_h^*$. Now, for the other part it suffices to show that $\liminf_N M_N^*/N \to c_h^*$. Suppose not, namely there exists $\varepsilon > 0$ and a sequence N_k such that $(M_{N_k}^*/N_k) + \varepsilon < c_h^*$ for all k. Then $(M_{N_k}^*/N_k) + \varepsilon/2 < c_h^* - \varepsilon/2$, implying that

$$\frac{H(\mathbb{G}(N_k, \lfloor M_{N_k}^* + (\varepsilon/2)N_k\rfloor))}{M_{N_k}^* + (\varepsilon/2)N_k} \to 1, \tag{3}$$

w.h.p. by Corollary 1. On the other hand, since M_N^* grows at most linearly with N, we may say $M_{N_k}^* + (\varepsilon/2)N_k \ge (1+\varepsilon')M_{N_k}^*$, for some $\varepsilon' > 0$ for all k. By Conjecture 1, this implies that $p(N_k, \lfloor M_{N_k}^* + (\varepsilon/2)N_k\rfloor) \to 0$ exponentially fast in N_k. This in turn means that there exists a sufficiently small $\delta > 0$ such that the deletion of *every* δN_k edges (clauses) keeps the instance unsatisfiable w.h.p. Namely, $H(\mathbb{G}(N_k, \lfloor M_{N_k}^* + (\varepsilon/2)N_k\rfloor)) \le M_{N_k}^* + (\varepsilon/2)N_k - \delta N_k$, w.h.p. as $k \to \infty$, which contradicts (3).

Let us now state our results for the existence of the scaling limit for the log-partition functions.

THEOREM 2. *For every $c > 0$, $1 \le \lambda < \infty$, and for every one of the models described in Section 2, there exists (model dependent) $z(c)$ such that*

$$\lim_{N\to\infty} N^{-1} \log Z(\mathbb{G}(N, \lfloor cN \rfloor)) = z(c), \tag{4}$$

w.h.p., where $z(c)$ is a Lipschitz continuous function of c. Moreover, $z(c)$ is nondecreasing for MAX-CUT, coloring, K-SAT and NAE-K-SAT models, and is a nonincreasing function of c for the independent set model.

REMARK 3. The case $\lambda = 1$ is actually not interesting as it corresponds to no interactions between the nodes leading to $Z(\mathbb{G}) = \prod_{i\in[N]} \lambda^{\sum_{x\in[q^-]} H_i(x)}$. In this case the limit of $N^{-1} \log Z(\mathbb{G}(N, \lfloor cN \rfloor))$ exists trivially when node potentials H_i are i.i.d. For independent set, our proof holds for $\lambda < 1$ as well. But, unfortunately our proof does not seem to extend to the case $\lambda < 1$ in the other models. For the Ising model this corresponds to the ferromagnetic case and the existence of the limit was established in [13] using a local analysis technique. The usage of local techniques is also discussed in [14] and [15]. Finally, we remark that the proof assumes the finiteness of λ. In fact, Coja-Oghlan observed [9] that if the above theorem could suitably be extended (addressing the case of when the number of solutions might be zero), to include the case of $\lambda = \infty$, then the satisfiability conjecture would follow.

We now turn to our results on random regular graphs.

THEOREM 3. *For every $r \in \mathbb{Z}_+$, and for all of the models described in the previous section, there exists (model dependent) $H(r)$ such that*

$$\lim_{N\to\infty, N\in r^{-1}K\mathbb{Z}_+} N^{-1} H(\mathbb{G}(N, r)) = H(r) \qquad \textit{w.h.p.}$$

Note, that in the statement of the theorem we take limits along subsequence N such that NrK^{-1} is an integer, so that the resulting random hypergraph is well-defined. Unlike the case of Erdös–Rényi graph, we were unable to prove the existence of the large deviation rate

$$\lim_{N\to\infty, N\in r^{-1}K\mathbb{Z}_+} N^{-1} \log p(N, r)$$

for the coloring, K-SAT and NAE-K-SAT problems and leave those as open questions.

Finally, we state our results for the log-partition function limits for random regular graphs.

THEOREM 4. *For every $r \in \mathbb{Z}_+$, $1 \le \lambda < \infty$, and for every one of the six models described in the previous section, there exists (model dependent) $z(r)$ such that w.h.p., we have*

$$\lim_{N\to\infty} N^{-1} \log Z(\mathbb{G}(N, r)) = z(r). \tag{5}$$

4. Proofs: Optimization problems in Erdös–Rényi graphs. The following simple observation will be useful throughout the paper. Given two hypergraphs $\mathbb{G}_i = ([N], E_i), i = 1, 2$ on the same set of nodes $[N]$ for each one of the six models in Section 2,

$$\big|H(\mathbb{G}_1) - H(\mathbb{G}_2)\big| = L|E_1 \Delta E_2|, \tag{6}$$

where we can take $L = 1$ for all the models except Ising, and we can take $L = \beta$ for the Ising model. This follows from the fact that adding (deleting) an edge to (from) a graph changes the value of H by at most 1 for all models except for the Ising model, where the constant is β.

Our main technical result leading to the proof of Theorem 1 is as follows.

THEOREM 5. *For every $1 \le N_1, N_2 \le N - 1$ such that $N_1 + N_2 = N$, and all models*

$$\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN\rfloor)\big)\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N_1, \mathcal{M}_1)\big)\big] + \mathbb{E}\big[H\big(\mathbb{G}(N_2, \mathcal{M}_2)\big)\big], \tag{7}$$

where $\mathcal{M}_1 \stackrel{d}{=} \mathrm{Bi}(\lfloor cN\rfloor, N_1/N)$ and $\mathcal{M}_2 \triangleq \lfloor cN\rfloor - \mathcal{M}_1 \stackrel{d}{=} \mathrm{Bi}(\lfloor cN\rfloor, N_2/N)$.

Additionally, for the same choice of $\mathcal{M}_j$ as above and for coloring, K-SAT and NAE-K-SAT models,

$$p\big(N, \lfloor cN\rfloor\big) \ge \mathbb{P}\big(H\big(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)\big) = \mathcal{M}_1 + \mathcal{M}_2\big), \tag{8}$$

and $\mathbb{G}_1 \oplus \mathbb{G}_2$ denotes a disjoint union of graphs $\mathbb{G}_1, \mathbb{G}_2$.

REMARK 4. The event $H(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)) = \mathcal{M}_1 + \mathcal{M}_2$ considered above corresponds to the event that both random graphs are satisfiable (colorable) instances. The randomness of choices of edges within each graph is assumed to be independent, but the number of edges $\mathcal{M}_j$ are dependent since they sum to $\lfloor cN\rfloor$. Because of this coupling, it is not the case that

$$\begin{aligned}&\mathbb{P}\big(H\big(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)\big) = \mathcal{M}_1 + \mathcal{M}_2\big)\\&\quad= \mathbb{P}\big(H\big(\mathbb{G}(N_1, \mathcal{M}_1)\big) = \mathcal{M}_1\big)\mathbb{P}\big(H\big(\mathbb{G}(N_2, \mathcal{M}_2)\big) = \mathcal{M}_2\big).\end{aligned}$$

Let us first show that Theorem 5 implies Theorem 1.

PROOF OF THEOREM 1. Since $\mathcal{M}_j$ have binomial distribution, we have $\mathbb{E}[|\mathcal{M}_j - \lfloor cN_j\rfloor|] = O(\sqrt{N})$. This together with observation (6) and Theorem 5 implies

$$\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN\rfloor)\big)\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N_1, \lfloor cN_1\rfloor)\big)\big] + \mathbb{E}\big[H\big(\mathbb{G}(N_2, \lfloor cN_2\rfloor)\big)\big] - O(\sqrt{N}).$$

Namely the sequence $\mathbb{E}[H(\mathbb{G}(N, \lfloor cN\rfloor))]$ is "nearly" super-additive, short of the $O(\sqrt{N})$ correction term. Now we use Proposition 5 in Appendix B for the case

$\alpha = 1/2$ to conclude that the limit $\lim_{N\to\infty} N^{-1}\mathbb{E}[H(\mathbb{G}(N, \lfloor cN \rfloor))] \triangleq H(c)$ exists.

Showing that this also implies convergence of $H(\mathbb{G}(N, \lfloor cN \rfloor))/N$ to $H(c)$ w.h.p. can be done using standard concentration results [24], and we skip the details. It remains to show that $H(c)$ is a nondecreasing continuous function for MAX-CUT, coloring, K-SAT and NAE-K-SAT problems and is nonincreasing for the independent set problem. For the MAX-CUT, coloring, K-SAT and NAE-K-SAT problems, we have

$$\mathbb{E}\big[H\big(\mathbb{G}(N, M_1)\big)\big] \le \mathbb{E}\big[H\big(\mathbb{G}(N, M_2)\big)\big],$$

when $M_1 \le M_2$; adding hyperedges can only increase the objective value since the edge potentials are nonnegative. For the Independent set problem on the contrary

$$\mathbb{E}\big[H\big(\mathbb{G}(N, M_1)\big)\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N, M_2)\big)\big]$$

holds. The Lipschitz continuity follows from (6) which implies

$$\big|\mathbb{E}\big[H\big(\mathbb{G}(N, M_1)\big)\big] - \mathbb{E}\big[H\big(\mathbb{G}(N, M_2)\big)\big]\big| = L|M_1 - M_2|$$

with $L = \beta$ for the Ising model, and $L = 1$ for the remaining models. This concludes the proof of (1).

We now turn to the proof of (2) and use (8) for this goal. Our main goal is establishing the following superadditivity property:

PROPOSITION 1. *There exist* $0 < \alpha < 1$ *such that for all* N_1, N_2 *such that* $N = N_1 + N_2$

$$\log p\big(N, \lfloor cN \rfloor\big) \ge \log p\big(N_1, \lfloor cN_1 \rfloor\big) + \log p\big(N_2, \lfloor cN_2 \rfloor\big) - O\big(N^{\alpha}\big). \tag{9}$$

PROOF OF PROPOSITION 1. Fix any $1/2 < \nu < 1$. First we assume $N_1 \le N^{\nu}$. Let $\mathcal{M}_j$ be as in Theorem 5. We have

$$\begin{aligned}
&\mathbb{P}\big(H\big(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)\big) = \mathcal{M}_1 + \mathcal{M}_2\big)\\
&\qquad \ge p\big(N_2, \lfloor cN_2 \rfloor\big) p\big(N_1, \lfloor cN \rfloor - \lfloor cN_2 \rfloor\big)\mathbb{P}\big(\mathcal{M}_2 = \lfloor cN_2 \rfloor\big).
\end{aligned}$$

We have

$$\mathbb{P}\big(\mathcal{M}_2 = \lfloor cN_2 \rfloor\big) = \binom{\lfloor cN \rfloor}{\lfloor cN_2 \rfloor} (N_2/N)^{\lfloor cN_2 \rfloor} (N_1/N)^{\lfloor cN \rfloor - \lfloor cN_2 \rfloor}.$$

From our assumption $N_1 \le N^{\nu}$ it follows that

$$\begin{aligned}
(N_2/N)^{\lfloor cN_2 \rfloor} &\ge \big(1 - N^{\nu-1}\big)^{O(N)} = \exp\big(-O\big(N^{\nu}\big)\big),\\
(N_1/N)^{\lfloor cN \rfloor - \lfloor cN_2 \rfloor} &\ge (1/N)^{O(N_1)} \ge \exp\big(-O\big(N^{\nu} \log N\big)\big).
\end{aligned}$$

It then follows

$$\mathbb{P}\big(\mathcal{M}_2 = \lfloor cN_2 \rfloor\big) \ge \exp\big(-O\big(N^{\nu} \log N\big)\big).$$

Now we claim the following crude bound for every deterministic m and every one of the three models under the consideration.

$$p(N, m+1) \geq O(1/N)p(N, m).$$

Indeed, for the K-SAT model, conditional on the event that $H(\mathbb{G}(N, m)) = m$, the probability that $H(\mathbb{G}(N, m+1)) = m+1$ is at least $1 - 1/2^K$. We obtain thus a bound which is even stronger than claimed,

$$p(N, m+1) \geq \left(1 - 1/2^K\right)p(N, m) = O\left(p(N, m)\right).$$

The proof for the NAE-K-SAT is similar. For the coloring problem observe that this conditional probability is at least $(1 - 1/N)(2(N-1)/N^2) = O(1/N)$ since with probability $1 - 1/N$ the new edge chooses different nodes, and with probability at least $2(N-1)/N^2$ the new edge does not violate a given coloring (with equality achieved only when $q = 2$, and two coloring classes having cardinalities 1 and $N - 1$). The claim follows.

Now since $\lfloor cN \rfloor - \lfloor cN_2 \rfloor \leq \lfloor cN_1 \rfloor + 1$, the claim implies

$$p\left(N_1, \lfloor cN \rfloor - \lfloor cN_2 \rfloor\right) \geq O(1/N)p\left(N_1, \lfloor cN_1 \rfloor\right).$$

Combining our estimates we obtain

$$\begin{aligned}&\mathbb{P}\left(H\left(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)\right) = \mathcal{M}_1 + \mathcal{M}_2\right)\\&\qquad \geq p\left(N_1, \lfloor cN_1 \rfloor\right)p\left(N_2, \lfloor cN_2 \rfloor\right)O(1/N)\exp\left(-O\left(N^\nu \log N\right)\right).\end{aligned}$$

After taking logarithm of both sides we obtain (9) from (8).

The case $N_2 \leq N^\nu$ is considered similarly. We now turn to a more difficult case $N_j > N^\nu$, $j = 1, 2$.

First we state the following lemma (proved in Appendix A) for the three models of interest (coloring, K-SAT, NAE-K-SAT).

LEMMA 1. *The following holds for coloring, K-SAT, NAE-K-SAT models for all N, M, m and $0 < \delta < 1/2$:*

$$p(N, M+m) \geq \delta^m p(N, M) - (2\delta)^{M+1}\exp\left(H(\delta)N + o(N)\right), \tag{10}$$

where $H(\delta) = -\delta\log\delta - (1-\delta)\log(1-\delta)$ is the entropy function.

We now prove (2). Fix $h \in (1/2, \nu)$. We have from (8),

$$\begin{aligned}&p\left(N, \lfloor cN \rfloor\right)\\&\qquad \geq \mathbb{P}\left(H\left(\mathbb{G}(N_1, \mathcal{M}_1) \oplus \mathbb{G}(N_2, \mathcal{M}_2)\right) = \mathcal{M}_1 + \mathcal{M}_2\right)\\&\qquad \geq \sum_{cN_1 - N^h \leq m_1 \leq cN_1 + N^h, m_2 = \lfloor cN \rfloor - m_1} p(N_1, m_1)p(N_2, m_2)\mathbb{P}(\mathcal{M}_1 = m_1).\end{aligned}$$

Note that $cN_1 - N^h \le m_1 \le cN_1 + N^h$ implies $cN_2 - N^h - 1 \le m_2 \le cN_2 + N^h$. Applying Lemma 1 we further obtain for the relevant range of m_j that

$$\begin{aligned}
&p(N_j, m_j) \\
&\quad \ge \delta^{(m_j - \lfloor cN_j \rfloor)^+} p(N_j, \lfloor cN_j \rfloor) - (2\delta)^{cN_j} \exp\bigl(H(\delta)N_j + o(N_j)\bigr) \\
&\quad \ge \delta^{N^h+1} p(N_j, \lfloor cN_j \rfloor) - (2\delta)^{cN_j} \exp\bigl(H(\delta)N_j + o(N_j)\bigr) \\
&\quad \ge \delta^{N^h+1} p(N_j, \lfloor cN_j \rfloor)\left(1 - 2^{cN_j} \delta^{cN_j - N^h - 1}\left(1 - \frac{1}{q}\right)^{-cN_j} e^{H(\delta)N_j + o(N_j)}\right),
\end{aligned}$$

where we have used a simple bound $p(N_j, \lfloor cN_j \rfloor) \ge (1 - 1/q)^{cN_j}$. Now let us take δ so that

$$\beta(\delta) \triangleq \bigl(2\delta(1 - 1/q)\bigr)^{-c} \exp\bigl(H(\delta)\bigr) < 1. \tag{11}$$

Then using the assumptions $N_j \ge N^\nu$ and $h < \nu$ we obtain

$$p(N_j, m_j) \ge \delta^{N^h+1} p(N_j, \lfloor cN_j \rfloor)\bigl(1 - \bigl(\beta(\delta)\bigr)^{O(N^\nu)}\bigr).$$

Combining we obtain

$$\begin{aligned}
p(N, \lfloor cN \rfloor) &\ge \mathbb{P}\bigl(cN_1 - N^h \le \mathcal{M}_1 \le cN_1 + N^h\bigr) \\
&\qquad \times \prod_{j=1,2} \delta^{O(N^h)} p(N_j, \lfloor cN_j \rfloor)\bigl(1 - \bigl(\beta(\delta)\bigr)^{O(N^\nu)}\bigr).
\end{aligned}$$

This implies

$$\begin{aligned}
\log p(N, \lfloor cN \rfloor) &\ge \log \mathbb{P}\bigl(cN_1 - N^h \le \mathcal{M}_1 \le cN_1 + N^h\bigr) \\
&\qquad + N^h \log \delta + \sum_{j=1,2} \log p(N_j, \lfloor cN_j \rfloor) \\
&\qquad + \log\bigl(1 - \bigl(\beta(\delta)\bigr)^{O(N^\nu)}\bigr).
\end{aligned}$$

Since $\mathcal{M}_1 \overset{d}{=} \mathrm{Bi}(\lfloor cN \rfloor, N_1/N)$ and $h > 1/2$, then

$$\bigl|\log \mathbb{P}\bigl(cN_1 - N^h \le \mathcal{M}_1 \le cN_1 + N^h\bigr)\bigr| = o(1).$$

Since $\beta(\delta) < 1$, then

$$\log\bigl(1 - \bigl(\beta(\delta)\bigr)^{O(N^\nu)}\bigr) = O\bigl(\bigl(\beta(\delta)\bigr)^{O(N^\nu)}\bigr) = o(N^h),$$

where the last identity is of course a very crude estimate. Combining, we obtain

$$\log p(N, \lfloor cN \rfloor) \ge \sum_{j=1,2} \log p(N_j, \lfloor cN_j \rfloor) + O(N^h).$$

The claim of Proposition 1 is established. $\square$

Part (2) of Theorem 1 then follows from this proposition and Proposition 5 from Appendix B. $\square$

We now turn to the proof of Theorem 5 and, in particular, introduce the interpolation construction.

PROOF OF THEOREM 5. We begin by constructing a sequence of graphs interpolating between $\mathbb{G}(N, \lfloor cN \rfloor)$ and a disjoint union of $\mathbb{G}(N_1, \mathcal{M}_1)$ and $\mathbb{G}(N_2, \lfloor cN \rfloor - \mathcal{M}_1)$. Given N, N_1, N_2 s.t. $N_1 + N_2 = N$ and any $0 \le r \le \lfloor cN \rfloor$, let $\mathbb{G}(N, \lfloor cN \rfloor, r)$ be the random graph on nodes $[N]$ obtained as follows. It contains precisely $\lfloor cN \rfloor$ hyperedges. The first r hyperedges $e_1, \dots, e_r$ are selected u.a.r. from all the possible directed hyperedges [namely they are generated as hyperedges of $\mathbb{G}(N, \lfloor cN \rfloor)$]. The remaining $\lfloor cN \rfloor - r$ hyperedges $e_{r+1}, \dots, e_{\lfloor cN \rfloor}$ are generated as follows. For each $j = r+1, \dots, \lfloor cN \rfloor$, with probability N_1/N, e_j is generated independently u.a.r. from all the possible hyperedges on nodes $[N_1]$, and with probability N_2/N, it is generated u.a.r. from all the possible hyperedges on nodes $[N_2]$ (=$\{N_1+1, \dots, N\}$). The choice of node and edge potentials H_v, H_e is done exactly according to the corresponding model, as for the case of graphs $\mathbb{G}(N, \lfloor cN \rfloor)$. Observe that when $r = \lfloor cN \rfloor$, $\mathbb{G}(N, \lfloor cN \rfloor, r) = \mathbb{G}(N, \lfloor cN \rfloor)$, and when $r = 0$, $\mathbb{G}(N, \lfloor cN \rfloor, r)$ is a disjoint union of graphs $\mathbb{G}(N_1, \mathcal{M}_1), \mathbb{G}(N_2, \mathcal{M}_2)$, conditioned on $\mathcal{M}_1 + \mathcal{M}_2 = \lfloor cN \rfloor$, where $\mathcal{M}_j \stackrel{d}{=} \mathrm{Bi}(\lfloor cN \rfloor, N_j/N)$.

PROPOSITION 2. *For every* $r = 1, \dots, \lfloor cN \rfloor$,

$$\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big)\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big)\big].$$

Also for coloring, K-SAT and NAE-K-SAT models,

$$\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big) = \lfloor cN \rfloor\big) \ge \mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big) = \lfloor cN \rfloor\big).$$

Let us first show how Theorem 5 follows from this proposition. Observe that for a disjoint union of two deterministic graphs $\mathbb{G} = \mathbb{G}_1 + \mathbb{G}_2$, with $\mathbb{G} = (V, E), \mathbb{G}_1 = (V_1, E_1), \mathbb{G}_2 = (V_2, E_2)$, we always have $H(\mathbb{G}) = H(\mathbb{G}_1) + H(\mathbb{G}_2)$. Claim (7) then follows. Caim (8) follows immediately from the interpolation construction by comparing the cases $r = 0$ and $r = \lfloor cN \rfloor$. $\square$

PROOF OF PROPOSITION 2. Observe that $\mathbb{G}(N, \lfloor cN \rfloor, r-1)$ is obtained from $\mathbb{G}(N, \lfloor cN \rfloor, r)$ by deleting a hyperedge chosen u.a.r. independently from r hyperedges $e_1, \dots, e_r$ and adding a hyperedge either to nodes $[N_1]$ or to $[N_2]$ with probabilities N_1/N and N_2/N, respectively. Let $\mathbb{G}_0$ be the graph obtained after deleting but before adding a hyperedge. For the case of K-SAT and NAE-K-SAT (two models with random edge potentials), assume that $\mathbb{G}_0$ also encodes the underlying edge potentials of the instance. For the case of coloring, K-SAT, NAE-K-SAT, note that the maximum value that H can achieve for the graph $\mathbb{G}_0$ is $\lfloor cN \rfloor - 1$ since exactly one hyperedge was deleted. We will establish a stronger result: conditional on *any* realization of the graph $\mathbb{G}_0$ (and random potentials), we claim that

$$\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN\rfloor, r)\big)|\mathbb{G}_0\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN\rfloor, r-1)\big)|\mathbb{G}_0\big] \tag{12}$$

and

$$\begin{aligned}&\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN\rfloor, r)\big) = \lfloor cN\rfloor|\mathbb{G}_0\big)\\ &\qquad \ge \mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN\rfloor, r-1)\big) = \lfloor cN\rfloor|\mathbb{G}_0\big)\end{aligned} \tag{13}$$

for coloring, K-SAT, NAE-K-SAT. Proposition then follows immediately from these claims by averaging over $\mathbb{G}_0$. Observe that *conditional on any realization* $\mathbb{G}_0$, $\mathbb{G}(N, \lfloor cN\rfloor, r)$ is obtained from $\mathbb{G}_0$ by adding a hyperedge to $[N]$ u.a.r. That is the generation of this hyperedge is independent from the randomness of $\mathbb{G}_0$. Similarly, conditional on any realization $\mathbb{G}_0$, $\mathbb{G}(N, \lfloor cN\rfloor, r-1)$ is obtained from $\mathbb{G}_0$ by adding a hyperedge to $[N_1]$ or $[N_2]$ u.a.r. with probabilities N_1/N and N_2/N, respectively.

We now prove properties (12) and (13) for each of the six models.

Independent sets. Let $O^* \subset [N]$ be the set of nodes which belong to *every* largest independent set in $\mathbb{G}_0$. Namely if $I \subset [N]$ is an i.s. such that $|I| = H(\mathbb{G}_0)$, then $O^* \subset I$. We note that O^* can be empty. Then for every edge $e = (i, k)$, $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0) - 1$ if $i, k \in O^*$ and $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0)$ if either $i \notin O^*$ or $k \notin O^*$. Here $\mathbb{G}_0 + e$ denotes a graph obtained from $\mathbb{G}_0$ by adding e. When the edge e is generated u.a.r. from the all possible edges, we then obtain $\mathbb{E}[H(\mathbb{G}_0 + e)|\mathbb{G}_0] - H(\mathbb{G}_0) = -(\frac{|O^*|}{N})^2$. Therefore, $\mathbb{E}[H(\mathbb{G}(N, \lfloor cN\rfloor, r))|\mathbb{G}_0] - H(\mathbb{G}_0) = -(\frac{|O^*|}{N})^2$. By a similar argument

$$\begin{aligned}&\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN\rfloor, r-1)\big)|\mathbb{G}_0\big] - H(\mathbb{G}_0)\\ &\quad = -\frac{N_1}{N}\bigg(\frac{|O^* \cap [N_1]|}{N_1}\bigg)^2 - \frac{N_2}{N}\bigg(\frac{|O^* \cap [N_2]|}{N_2}\bigg)^2\\ &\quad \le -\bigg(\frac{N_1}{N}\frac{|O^* \cap [N_1]|}{N_1} + \frac{N_2}{N}\frac{|O^* \cap [N_2]|}{N_2}\bigg)^2\\ &\quad = -\bigg(\frac{|O^*|}{N}\bigg)^2 = \mathbb{E}\big[H\big(\mathbb{G}\big(N, \lfloor cN\rfloor, r\big)|\mathbb{G}_0\big] - H(\mathbb{G}_0),\end{aligned}$$

and (12) is established.

MAX-CUT. Given $\mathbb{G}_0$, let $\mathcal{C}^* \subset \{0, 1\}^{[N]}$ be the set of optimal solutions. Namely $H(x) = H(\mathbb{G}_0), \forall x \in \mathcal{C}^*$ and $H(x) < H(\mathbb{G}_0)$ otherwise. Introduce an equivalency relationship $\sim$ on $[N]$. Given $i, k \in [N]$, define $i \sim k$ if *for every* $x \in \mathcal{C}^*, x_i = x_k$. Namely, in every optimal cut, nodes i and k have the same value. Let $O_j^* \subset [N], 1 \le j \le J$, be the corresponding equivalency classes. Given any edge $e =$

(i,k), observe that $H(\mathbb{G}_0+e)=H(\mathbb{G}_0)$ if $i\sim k$ and $H(\mathbb{G}_0+e)=H(\mathbb{G}_0)+1$ otherwise. Thus

$$\mathbb{E}\big[H\big(\mathbb{G}\big(N,\lfloor cN\rfloor,r\big)|\mathbb{G}_0\big]-H(\mathbb{G}_0)=1-\sum_{1\le j\le J}\left(\frac{|O_j^*|}{N}\right)^2$$

and

$$\begin{aligned}&\mathbb{E}\big[H\big(\mathbb{G}\big(N,\lfloor cN\rfloor,r-1\big)|\mathbb{G}_0\big]-H(\mathbb{G}_0)\\&\quad=1-\frac{N_1}{N}\sum_{1\le j\le J}\left(\frac{|O_j^*\cap[N_1]|}{N_1}\right)^2-\frac{N_2}{N}\sum_{1\le j\le J}\left(\frac{|O_j^*\cap[N_2]|}{N_2}\right)^2.\end{aligned}$$

Using $\frac{N_1}{N}(\frac{|O_j^*\cap[N_1]|}{N_1})^2+\frac{N_2}{N}(\frac{|O_j^*\cap[N_2]|}{N_2})^2\ge(\frac{|O_j^*|}{N})^2$ we obtain (12).

Ising. The proof is similar to the MAX-CUT problem but is more involved due to the presence of the magnetic field B. The presence of the field means that we can no longer say that $H(\mathbb{G}_0+e)=H(\mathbb{G}_0)+\beta$ or $=H(\mathbb{G}_0)$. This issue is addressed by looking at suboptimal solutions and the implied sequence of equivalence classes. Thus let us define a sequence $H_0>H_1>H_2\cdots>H_M$ and an integer $M\ge 0$ as follows. Define $H_0=H(\mathbb{G}_0)$. Assuming H_{m-1} is defined, $m\ge 1$, let $H_m=\max H(x)$ over all solutions $x\in\{0,1\}^N$ such that $H(x)<H_{m-1}$. Namely, H_m is the next best solution after H_{m-1}. Define M to be the largest m such that $H_m\ge H(\mathbb{G}_0)-2\beta$. If this is not the case for all m, then we define $M\le 2^N$ to be the total number of possible values $H(x)$ (although typically the value of M will be much smaller). Let $\mathcal{C}_m=\{x:H(x)=H_m\}, 0\le m\le M$, be the set of solutions achieving value H_m. Observe that $\mathcal{C}_m$ are disjoint sets. For every $m\le M$ define an equivalency relationship as follows $i\overset{m}{\sim}k$ for $i,k\in[N]$ if and only if $x_i=x_k$ for all $x\in\mathcal{C}_0\cup\cdots\cup\mathcal{C}_m$. Namely, nodes i and k are m-equivalent if they take equal values in every solution achieving value at least H_m. Let $O_{j,m}^*$ be the corresponding equivalency classes for $1\le j\le J_m$. Note that the partition $O_{j,m+1}^*$ of the nodes $[N]$ is a refinement of the partition $O_{j,m}^*$.

LEMMA 2. *Given an edge $e=(i,k)$, the following holds*:

$$H(\mathbb{G}+e)=\begin{cases}H(\mathbb{G}_0)+\beta, & \textit{if } i\overset{0}{\nsim}k;\\ H_{m+1}+\beta, & \textit{if } i\overset{m}{\sim}k, \textit{ but } i\overset{m+1}{\nsim}k, \textit{for some } m\le M-1;\\ H(\mathbb{G}_0)-\beta, & \textit{if } i\overset{m}{\sim}k \textit{ for all } m\le M.\end{cases}$$

PROOF. The case $i\overset{0}{\nsim}k$ is straightforward. Suppose $i\overset{m}{\sim}k$, but $i\overset{m+1}{\nsim}k$ for some $m\le M-1$. For every $x\in\bigcup_{m'\le m}\mathcal{C}_{m'}$ we have for some $m'\le m$, $H_{\mathbb{G}_0+e}(x)=H_{m'}-\beta\le H_0-\beta$. Now since $i\overset{m+1}{\nsim}k$ there exists $x\in\mathcal{C}_{m+1}$ such that $x_i\ne x_k$, implying $H_{\mathbb{G}_0+e}(x) = H_{m+1} + \beta \ge H_0 - \beta$, where the inequality follows since $m+1 \le M$. Furthermore, for every $x \notin \bigcup_{m' \le m} \mathcal{C}_{m'}$ we have $H_{\mathbb{G}_0+e}(x) \le H_{\mathbb{G}_0}(x) + \beta \le H_{m+1} + \beta$. We conclude that $H_{m+1} + \beta$ is the optimal solution in this case.

On the other hand, if $i \overset{m}{\sim} k$ for all $m \le M$, then for all $x \in \bigcup_{m \le M} \mathcal{C}_m$, $H_{\mathbb{G}_0+e}(x) \le H(\mathbb{G}_0) - \beta$, with equality achieved for $x \in \mathcal{C}_0$. For all $x \notin \bigcup_{m \le M} \mathcal{C}_m$, we have $H_{\mathbb{G}_0+e}(x) \le H_{M+1} + \beta < H_0 - \beta$, and the assertion is established. Note that if $M = 2^N$, namely $M+1$ is not defined, then $\bigcup_{m \le M} \mathcal{C}_m$ is the entire space of solutions $\{0,1\}^N$, and the second part of the previous sentence is irrelevant. $\square$

We now return to the proof of the proposition. Recall that if an edge $e = (i,k)$ is added uniformly at random then $\mathbb{P}(i \overset{m}{\sim} k) = \sum_{1 \le j \le J_m} (\frac{|O^*_{j,m}|}{N})^2$. A similar assertion holds for the case e is added uniformly at random to parts $[N_l], l = 1, 2$, with probabilities N_l/N, respectively. We obtain that

$$\mathbb{P}(i \overset{0}{\nsim} k) = 1 - \sum_{1 \le j \le J_0} \left(\frac{|O^*_{j,0}|}{N}\right)^2,$$

$$\mathbb{P}(i \overset{m}{\sim} k, \text{but } i \overset{m+1}{\nsim} k) = \sum_{1 \le j \le J_m} \left(\frac{|O^*_{j,m}|}{N}\right)^2 - \sum_{1 \le j \le J_{m+1}} \left(\frac{|O^*_{j,m+1}|}{N}\right)^2,$$

$$\mathbb{P}(i \overset{m}{\sim} k, \forall m \le M) = \sum_{1 \le j \le J_M} \left(\frac{|O^*_{j,M}|}{N}\right)^2.$$

Applying Lemma 2 we obtain

$$\begin{aligned}
&\mathbb{E}[H(\mathbb{G}(N, \lfloor cN \rfloor, r)) | \mathbb{G}_0] \\
&\quad = (H + \beta)\left(1 - \sum_{1 \le j \le J_m} \left(\frac{|O^*_{j,0}|}{N}\right)^2\right) \\
&\qquad + \sum_{m=0}^{M-1} (H_{m+1} + \beta)\left(\sum_{1 \le j \le J_m} \left(\frac{|O^*_{j,m}|}{N}\right)^2 - \sum_{1 \le j \le J_{m+1}} \left(\frac{|O^*_{j,m+1}|}{N}\right)^2\right) \\
&\qquad + (H - \beta) \sum_{1 \le j \le J_M} \left(\frac{|O^*_{j,M}|}{N}\right)^2 \\
&\quad = H + \beta + \sum_{0 \le m \le M-1} (H_{m+1} - H_m) \sum_{1 \le j \le J_m} \left(\frac{|O^*_{j,m}|}{N}\right)^2 \\
&\qquad + (H - H_M - 2\beta) \sum_{1 \le j \le J_M} \left(\frac{|O^*_{j,M}|}{N}\right)^2.
\end{aligned}$$

By a similar argument and again using Lemma 2 we obtain

$$\begin{aligned}
&\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big)|\mathbb{G}_0\big] \\
&\quad = H + \beta + \sum_{0\le m\le M-1} (H_{m+1} - H_m) \sum_{1\le j\le J_m} \sum_{l=1,2} \frac{N_l}{N} \left(\frac{|O^*_{j,m} \cap [N_l]|}{N_l}\right)^2 \\
&\qquad + (H - H_M - 2\beta) \sum_{1\le j\le J_M} \sum_{l=1,2} \frac{N_l}{N} \left(\frac{N_l}{N} \frac{|O^*_{j,M} \cap [N_l]|}{N_l}\right)^2.
\end{aligned}$$

Recall, however, that $H_{m+1} - H_m < 0, m \le M-1$ and $H - H_M - 2\beta \le 0$. Again using the convexity of the $g(x) = x^2$ function, we obtain the claim.

Coloring. Let $\mathcal{C}^* \subset [q^-]^N$ be the set of optimal colorings. Namely $H(x) = H(\mathbb{G}_0), \forall x \in \mathcal{C}^*$. Introduce an equivalency relationship $\sim$ on the set of nodes as follows. Given $i, k \in [N]$, define $i \sim k$ if and only if $x_i = x_k$ for every $x \in \mathcal{C}^*$. Namely, in every optimal coloring assignments, i and k receive the same color. Then for every edge e, $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0)$ if $i \sim k$ and $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0) + 1$ otherwise. The remainder of the proof of (12) is similar to the one for MAX-CUT.

Now let us show (13). We fix graph $\mathbb{G}_0$. Notice that if $\mathbb{G}_0$ is not colorable, then both probabilities in (13) are zero, since adding edges cannot turn an uncolorable instance into the colorable one. Thus assume $\mathbb{G}_0$ is a colorable graph. Since it has $\lfloor cN \rfloor - 1$ edges it means $H(\mathbb{G}_0) = \lfloor cN \rfloor - 1$. Let $O^*_j \subset [N], 1 \le j \le J$ denote the $\sim$ equivalence classes, defined by $i \sim k$ if and only if in every proper coloring assignment i and k receive the same color. We obtain that

$$\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big) = \lfloor cN \rfloor|\mathbb{G}_0\big) = 1 - \sum_{1\le j\le J} \left(\frac{|O^*_j|}{N}\right)^2.$$

Similarly,

$$\begin{aligned}
&\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big) = \lfloor cN \rfloor|\mathbb{G}_0\big) \\
&\quad = 1 - \frac{N_1}{N} \sum_{1\le j\le J} \left(\frac{|O^*_j \cap [N_1]|}{N_1}\right)^2 - \frac{N_2}{N} \sum_{1\le j\le J} \left(\frac{|O^*_j \cap [N_2]|}{N_2}\right)^2.
\end{aligned}$$

Relation (13) then again follows from convexity.

K-SAT. Let $\mathcal{C}^* \subset \{0,1\}^N$ be the set of optimal assignments. Define a node i (variable x_i) to be *frozen* if either $x_i = 0, \forall x \in \mathcal{C}^*$ or $x_i = 1, \forall x \in \mathcal{C}^*$. Namely, in every optimal assignment the value of i is always the same. Let O^* be the set of frozen variables. Let $e = (i_1, \dots, i_K) \subset [N]$ be a hyperedge, and let $H_e : \{0,1\}^K \to \{0,1\}$ be the corresponding edge potential. Namely, for some $y_1, \dots, y_K \in \{0,1\}$, $H_e(x_{i_1}, \dots, x_{i_k}) = 0$ if $x_{i_1} = y_1, \dots, x_{i_K} = y_K$ and $H_e = 1$ otherwise. Consider adding e with H_e to the graph $\mathbb{G}_0$. Note that if $e \cap ([N] \setminus O^*) \ne \varnothing$, then $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0) + 1$, as in this case at least one variable in e is nonfrozen and can be adjusted to satisfy the clause. Otherwise, suppose $e \subset O^*$, and let $x^*_{i_1}, \dots, x^*_{i_K} \in \{0, 1\}$ be the corresponding frozen values of $i_1, \dots, i_K$. Then $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0)$ if $x^*_{i_1} = y_1, \dots, x^*_{i_K} = y_K$, and $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0) + 1$ otherwise. Moreover, for the random choice of H, the first event $H(\mathbb{G}_0 + e) = H(\mathbb{G}_0)$ occurs with probability $1/2^K$. We conclude that

$$\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big)|\mathbb{G}_0\big] - H(\mathbb{G}_0) = 1 - \frac{1}{2^K}\left(\frac{|O^*|}{N}\right)^K$$

and for every satisfiable instance $\mathbb{G}_0$ (namely $H(\mathbb{G}_0) = \lfloor cN \rfloor - 1$),

$$\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big) = \lfloor cN \rfloor|\mathbb{G}_0\big) = 1 - \frac{1}{2^K}\left(\frac{|O^*|}{N}\right)^K.$$

Similarly,

$$\begin{aligned}&\mathbb{E}\big[H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big)|\mathbb{G}_0, H_0\big] - H(\mathbb{G}_0)\\ &\qquad = 1 - \frac{1}{2^K}\frac{N_1}{N}\left(\frac{|O^* \cap [N_1]|}{N_1}\right)^K - \frac{1}{2^K}\frac{N_2}{N}\left(\frac{|O^* \cap [N_2]|}{N_2}\right)^K\end{aligned}$$

and for every satisfiable instance $\mathbb{G}_0$,

$$\begin{aligned}&\mathbb{P}\big(H\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big) = \lfloor cN \rfloor|\mathbb{G}_0\big)\\ &\qquad = 1 - \frac{1}{2^K}\frac{N_1}{N}\left(\frac{|O^* \cap [N_1]|}{N_1}\right)^K - \frac{1}{2^K}\frac{N_2}{N}\left(\frac{|O^* \cap [N_2]|}{N_2}\right)^K.\end{aligned}$$

Using the convexity of the function x^K on $x \in [0, \infty)$, we obtain the result.

NAE-K-SAT. The idea of the proof is similar and is based on the combination of the notions of frozen variables and equivalency classes. Two nodes (variables) i and k are defined to be equivalent $i \sim k$ if there do not exist two assignments x, x' such that $x_i = x'_i$, but $x_k \neq x'_k$, or vice verse, $x_i \neq x'_i$, but $x_k = x'_k$. Namely, either both nodes are frozen, or setting one of them determines the value for the other in every optimal assignment. Let O^*_j, $1 \le j \le J$, be the set of equivalence classes (the set of frozen variables is one of O^*_j). Let $e = (i_1, \dots, i_K) \subset [N]$ be a hyperedge added to $\mathbb{G}_0$, and let $H_e : \{0, 1\}^K \to \{0, 1\}$ be the corresponding edge potential. We claim that if $i_1, \dots, i_K$ are not all equivalent, then $H(\mathbb{G} + e) = H(\mathbb{G})$. Indeed, suppose without the loss of generality that $i_1 \not\sim i_2$ and x, x' are two optimal solutions such that $x_{i_1} = x'_{i_1}, x_{i_2} \neq x'_{i_2}$. From the definition of NAE-K-SAT model, it follows that at least one of the two solutions x and x' satisfies H_e as well, and the claim then follows. Thus, $H(\mathbb{G} + e) = H(\mathbb{G})$ only if $i_1, \dots, i_K$ all belong to the same equivalence class. Provided that this indeed occurs, it is easy to see that the probability that $H(\mathbb{G} + e) = H(\mathbb{G})$ is $2/2^K$. The remainder of the proof is similar to the one for the K-SAT model.

We have established (12) and (13). With this, the proof of Proposition 2 is complete. □

Finally we give a simple proof of Corollary 1.

PROOF OF COROLLARY 1. Define $c_H^* = \sup\{c \geq 0 : H(c) = c\}$. It suffices to show that $H(c) < c$ for all $c > c_H^*$. For every $\delta > 0$ we can find $c_0 \in (c, c+\delta)$ such that $H(c_0) < c_0$. By Lipshitz continuity result of Theorem 1 it follows that $H(c) \leq H(c_0) + (c - c_0) < c$ for all $c > c_0$, and the assertion is established. $\square$

5. Proofs: Log-partition function in Erdös–Rényi graphs. The following property serves as an analogue of (6). Given two hypergraphs $\mathbb{G}_i = ([N], E_i)$, $i = 1, 2$, on the same set of nodes $[N]$ for each one of the six models and each finite λ,

$$\bigl|\log Z(\mathbb{G}_1) - \log Z(\mathbb{G}_2)\bigr| = O\bigl(|E_1 \Delta E_2|\bigr). \tag{14}$$

This follows from the fact that adding (deleting) a hyperedge to (from) a graph results in multiplying or dividing the partition function by at most λ for all models except for the Ising and Independent set models. For the Ising model the corresponding value is λ^β. To obtain a similar estimate for the independent set, note that given a graph $\mathbb{G}$ and an edge $e = (u, v)$ which is not in $\mathbb{G}$, we have

$$Z(\mathbb{G}) = \sum_{e \subset I} \lambda^{|I|} + \sum_{e \not\subset I} \lambda^{|I|},$$

where in both sums we only sum over independent sets of $\mathbb{G}$. We claim that

$$\sum_{e \subset I} \lambda^{|I|} \leq \lambda \sum_{e \not\subset I} \lambda^{|I|}.$$

Indeed, for every independent set in $\mathbb{G}$ containing $e = (u, v)$, delete node u. We obtain a one-to-one mapping immediately leading to the inequality. Finally, we obtain

$$Z(\mathbb{G}) \geq Z(\mathbb{G} + e) = \sum_{e \not\subset I} \lambda^{|I|} \geq \frac{1}{1+\lambda} Z(\mathbb{G}),$$

where our claim was used in the second inequality. Assertion (14) then follows after taking logarithms.

The analogue of Theorem 5 is the following result.

THEOREM 6. *For every $1 \leq N_1, N_2 \leq N - 1$ such that $N_1 + N_2 = N$ and every $\lambda > 1$,*

$$\mathbb{E}\bigl[\log Z\bigl(\mathbb{G}(N, \lfloor cN \rfloor)\bigr)\bigr] \geq \mathbb{E}\bigl[\log Z\bigl(\mathbb{G}(N_1, \mathcal{M}_1)\bigr)\bigr] + \mathbb{E}\bigl[\log Z\bigl(\mathbb{G}(N_2, \mathcal{M}_2)\bigr)\bigr],$$

where $\mathcal{M}_1 \stackrel{d}{=} \mathrm{Bi}(\lfloor cN \rfloor, N_1/N)$ and $\mathcal{M}_2 \triangleq \lfloor cN \rfloor - \mathcal{M}_1 \stackrel{d}{=} \mathrm{Bi}(\lfloor cN \rfloor, N_1/N)$.

As before, we do not have independence of $\mathcal{M}_j$, $j = 1, 2$. Let us first show how this result implies Theorem 2.

PROOF OF THEOREM 2. Since $\mathcal{M}_j$ have binomial distribution, using observation (14) and Theorem 6, we obtain

$$\begin{aligned}&\mathbb{E}[\log Z(\mathbb{G}(N, \lfloor cN \rfloor))]\\&\qquad \geq \mathbb{E}[\log Z(\mathbb{G}(N_1, \lfloor cN_1 \rfloor))] + \mathbb{E}[Z(\mathbb{G}(N_2, \lfloor cN_2 \rfloor))] - O(\sqrt{N}).\end{aligned}$$

Now we use Proposition 5 in Appendix B for the case $\alpha = 1/2$ to conclude that the limit

$$\lim_{N\to\infty} N^{-1}\mathbb{E}[\log Z(\mathbb{G}(N, \lfloor cN \rfloor))] \triangleq z(c)$$

exists. Showing that this also implies the convergence of $N^{-1}\mathbb{E}[\log Z(\mathbb{G}(N, \lfloor cN \rfloor))]$ to $z(c)$ w.h.p. again is done using standard concentration results [24] by applying property (14), and we skip the details. The proof of continuity and monotonicity of $z(c)$ for relevant models is similar to the one of $H(c)$. $\square$

Thus it remains to prove Theorem 6.

PROOF OF THEOREM 6. We construct an interpolating sequence of graphs $\mathbb{G}(N, \lfloor cN \rfloor, r)$, $0 \leq r \leq \lfloor cN \rfloor$ exactly as in the previous subsection. We now establish the following analogue of Proposition 2. $\square$

PROPOSITION 3. *For every* $r = 1, \ldots, \lfloor cN \rfloor$,

$$\mathbb{E}[\log Z(\mathbb{G}(N, \lfloor cN \rfloor, r))] \geq \mathbb{E}[\log Z(\mathbb{G}(N, \lfloor cN \rfloor, r - 1))]. \tag{15}$$

Let us first show how Theorem 6 follows from this proposition. Observe that for disjoint union of two graphs $\mathbb{G} = \mathbb{G}_1 + \mathbb{G}_2$, with $\mathbb{G} = (V, E)$, $\mathbb{G}_1 = (V_1, E_1)$, $\mathbb{G}_2 = (V_2, E_2)$, we always have $\log Z(\mathbb{G}) = \log Z(\mathbb{G}_1) + \log Z(\mathbb{G}_2)$. Theorem 6 then follows from Proposition 3.

PROOF OF PROPOSITION 3. Recall that $\mathbb{G}(N, \lfloor cN \rfloor, r - 1)$ is obtained from $\mathbb{G}(N, \lfloor cN \rfloor, r)$ by deleting a hyperedge chosen u.a.r. independently from r hyperedges $e_1, \ldots, e_r$ and adding a hyperedge e either to nodes $[N_1]$ or to nodes $[N_2]$ with probabilities N_1/N and N_2/N, respectively. Let as before $\mathbb{G}_0$ be the graph obtained after deleting but before adding a hyperedge, and let $Z_0 = Z_0(\mathbb{G}_0)$ and $\mu_0 = \mu_{0,\mathbb{G}_0}$ be the corresponding partition function and the Gibbs measure, respectively. In the case of K-SAT and NAE-K-SAT models we assume that $\mathbb{G}_0$ encodes the realizations of the random potentials as well. We now show that conditional on any realization of the graph $\mathbb{G}_0$,

$$\mathbb{E}\big[\log Z\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big)|\mathbb{G}_0\big] \geq \mathbb{E}\big[\log Z\big(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\big)|\mathbb{G}_0\big]. \tag{16}$$

The proof of (16) is done on a case-by-case basis, and it is very similar to the proof of (12).

Independent sets. We have

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big)|\mathbb{G}_0\big] - \log Z_0 \\
&\quad = \mathbb{E}\bigg[\log \frac{Z(\mathbb{G}(N, \lfloor cN \rfloor, r))}{Z_0}\Big|\mathbb{G}_0\bigg] \\
&\quad = \mathbb{E}\bigg[\log \frac{\sum_I \lambda^{|I|} - \sum_I 1_{\{e \subset I\}} \lambda^{|I|}}{\sum_I \lambda^{|I|}}\Big|\mathbb{G}_0\bigg] \\
&\quad = \mathbb{E}\big[\log\big(1 - \mu_0(e \subset I_0)\big)|\mathbb{G}_0\big],
\end{aligned}$$

where the sums $\sum_I$ are over independent sets only, and I_0 denotes an independent set chosen randomly according to μ_0. Notice that since we are conditioning on graph $\mathbb{G}_0$, the only randomness underlying the expectation operator is the randomness of the hyperedge e and the randomness of set I_0. Note that $\mu_0(e \subset I_0) < 1$ since $\mu_0(e \not\subset I_0) \geq \mu_0(I_0 = \varnothing) > 0$. Using the expansion $\log(1-x) = -\sum_{m \geq 1} x^m/m$,

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N, \lfloor cN \rfloor, r)\big)|\mathbb{G}_0\big] - \log Z_0 \\
&\quad = -\mathbb{E}\bigg[\sum_{k=1}^{\infty} \frac{\mu_0(e \subset I_0)^k}{k}\Big|\mathbb{G}_0\bigg] \\
&\quad = -\sum_{k=1}^{\infty} \frac{1}{k} \mathbb{E}\bigg[\sum_{I^1, \dots, I^k} 1_{\{e \subset \bigcap_{j=1}^k I^j\}} \frac{\lambda^{\sum_{j=1}^k |I^j|}}{Z_0^k}\Big|\mathbb{G}_0\bigg] \\
&\quad = -\sum_{k=1}^{\infty} \frac{1}{k} \sum_{I^1, \dots, I^k} \frac{\lambda^{\sum_{j=1}^k |I^j|}}{Z_0^k} \mathbb{E}[1_{\{e \subset \bigcap_{j=1}^k I^j\}}|\mathbb{G}_0] \\
&\quad = -\sum_{k=1}^{\infty} \frac{1}{k} \sum_{I^1, \dots, I^k} \frac{\lambda^{\sum_{j=1}^k |I^j|}}{Z_0^k} \bigg(\frac{|\bigcap_{j=1}^k I^j|}{N}\bigg)^2,
\end{aligned}$$

where the sum $\sum_{I^1, \dots, I^k}$ is again over independent subsets $I^1, \dots, I^k$ of $\mathbb{G}_0$ only, and in the last equality we have used the fact that e is distributed u.a.r. Similar calculation for $\log Z(\mathbb{G}(N, \lfloor cN \rfloor, r-1))$ that is obtained by adding a hyperedge to nodes $[N_1]$ with probability N_1/N, or to nodes $[N_2]$ with probability N_2/N, gives

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N,\lfloor cN\rfloor,r-1)\big)|\mathbb{G}_0\big]-\log Z_0\\
&\quad=-\sum_{k=1}^{\infty}\frac{1}{k}\sum_{I^1,\dots,I^k}\frac{\lambda^{\sum_{j=1}^k|I^j|}}{Z_0^k}\bigg[\frac{N_1}{N}\bigg(\frac{|\bigcap_{j=1}^k I^j\cap[N_1]|}{N_1}\bigg)^2\\
&\qquad\qquad+\frac{N_2}{N}\bigg(\frac{|\bigcap_{j=1}^k I^j\cap[N_2]|}{N_2}\bigg)^2\bigg].
\end{aligned}$$

Again using the convexity of $f(x)=x^2$ we obtain

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N,\lfloor cN\rfloor,r)\big)|\mathbb{G}_0\big]-\log Z_0\\
&\qquad\geq \mathbb{E}\big[\log Z\big(\mathbb{G}(N,\lfloor cN\rfloor,r-1)\big)|\mathbb{G}_0\big]-\log Z_0,
\end{aligned}$$

and (16) is established.

MAX-CUT. Similarly to the independent set model, if $\mathbb{G}(N,\lfloor cN\rfloor,r)$ is obtained from $\mathbb{G}_0$ by adding an edge (i,j) where i,j are chosen uniformly at random, we have

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N,\lfloor cN\rfloor,r)\big)|\mathbb{G}_0\big]-\log Z_0\\
&\quad=\mathbb{E}\bigg[\log\frac{Z(\mathbb{G}(N,\lfloor cN\rfloor,r))}{Z_0}\Big|\mathbb{G}_0\bigg]\\
&\quad=\mathbb{E}\bigg[\log\frac{\sum_x 1_{\{x_i=x_j\}}\lambda^{H(x)}+\lambda\sum_x 1_{\{x_i\neq x_j\}}\lambda^{H(x)}}{\sum_x\lambda^{H(x)}}\Big|\mathbb{G}_0\bigg]\\
&\quad=\log\lambda+\mathbb{E}\bigg[\log\bigg(1-\bigg(1-\frac{1}{\lambda}\bigg)\mu_0(x_i=x_j)\bigg)\Big|\mathbb{G}_0\bigg].
\end{aligned}$$

Since $\lambda>1$ we have $0<(1-\lambda^{-1})\mu_0(x_i=x_j)<1$ (this is where the condition $\lambda>1$ is used), implying

$$\begin{aligned}
&\mathbb{E}\big[\log Z\big(\mathbb{G}(N,\lfloor cN\rfloor,r)\big)|\mathbb{G}_0\big]-\log Z_0-\log\lambda\\
&\quad=-\mathbb{E}\bigg[\sum_{k=1}^{\infty}\frac{(1-\lambda^{-1})^k\mu_0(x_i=x_j)^k}{k}\Big|\mathbb{G}_0\bigg]\\
&\quad=-\sum_{k=1}^{\infty}\frac{(1-\lambda^{-1})^k}{k}\mathbb{E}\bigg[\sum_{x_1,\dots,x_k}\frac{\lambda^{\sum_{\ell=1}^k H(x_\ell)}}{Z_0^k}1_{\{x_i^\ell=x_j^\ell,\forall\ell\}}\Big|\mathbb{G}_0\bigg]\\
&\quad=-\sum_{k=1}^{\infty}\frac{(1-\lambda^{-1})^k}{k}\sum_{x_1,\dots,x_k}\frac{\lambda^{\sum_{\ell=1}^k H(x_\ell)}}{Z_0^k}\mathbb{E}[1_{\{x_i^\ell=x_j^\ell,\forall\ell\}}|\mathbb{G}_0].
\end{aligned}$$

Now for every sequence of vectors $x_1,\dots,x_k$ introduce equivalency classes on $[N]$. Given $i,k\in[N]$, say $i\sim k$ if $x_i^\ell=x_k^\ell,\forall\ell=1,\dots,k$. Namely, in every one of the cuts defined by x_ℓ, $\ell = 1, \dots, k$, the nodes i and k belong to the same side of the cut. Let O_s, $1 \le s \le J$ be the corresponding equivalency classes. For an edge $e = (i, j)$ generated u.a.r., observe that $\mathbb{E}[1_{\{x_i^\ell = x_j^\ell \forall \ell\}} | \mathbb{G}_0] = \sum_{s=1}^J (\frac{|O_s|}{N})^2$. Thus

$$\mathbb{E}\bigl[\log Z\bigl(\mathbb{G}(N, \lfloor cN \rfloor, r)\bigr)|\mathbb{G}_0\bigr] - \log Z_0 - \log \lambda$$
$$= -\sum_{k=1}^{\infty} \frac{(1-\lambda^{-1})^k}{k} \sum_{x_1,\dots,x_k} \frac{\lambda^{\sum_{\ell=1}^k H(\ell)}}{Z_0^k} \sum_{s=1}^J \biggl(\frac{|O_s|}{N}\biggr)^2$$

and similarly,

$$\mathbb{E}\bigl[\log Z\bigl(\mathbb{G}(N, \lfloor cN \rfloor, r-1)\bigr)|\mathbb{G}_0\bigr] - \log Z_0 - \log \lambda$$
$$= -\sum_{k=1}^{\infty} \frac{(1-1/\lambda)^k}{k} \sum_{x_1,\dots,x_k} \frac{\lambda^{\sum_{\ell=1}^k H(\ell)}}{Z_0^k}$$
$$\times \sum_{s=1}^J \biggl(\frac{N_1}{N}\biggl(\frac{|O_s \cap [N_1]|}{N_1}\biggr)^2 + \frac{N_2}{N}\biggl(\frac{|O_s \cap [N_2]|}{N_2}\biggr)^2\biggr).$$

Using the convexity of the function $f(x) = x^2$, we obtain (16).

Ising, *coloring*, *K-SAT and NAE-K-SAT*. The proofs of the remaining cases are obtained similarly and are omitted. The condition $\lambda > 1$ is used to assert positivity of $1 - \lambda^{-1}$ in the logarithm expansion. □

6. Proofs: Random regular graphs.

For the proofs related to random regular graphs, we will need to work with random "nearly" regular graphs. For this purpose, given N, r and K such that Nr/K is an integer and given any positive integer $T \le Nr/K$, let $\mathbb{G}(N, r, T)$ denote the graph obtained by creating a size T partial matching on Nr nodes of the configuration model uniformly at random and then projecting. For example, if T was Nr/K, then we would have obtained the random regular graph $\mathbb{G}(N, r)$.

Our result leading to the proof of Theorem 3 is as follows.

THEOREM 7. *For every N_1, N_2 such that $N = N_1 + N_2$ and $N_1 r/K, N_2 r/K$ are integers*,

$$\mathbb{E}\bigl[H\bigl(\mathbb{G}(N, r)\bigr)\bigr] \ge \mathbb{E}\bigl[H\bigl(\mathbb{G}(N_1, r)\bigr)\bigr] + \mathbb{E}\bigl[H\bigl(\mathbb{G}(N_2, r)\bigr)\bigr] - O\bigl(N^{5/6}\bigr). \tag{17}$$

PROOF. Fix N_1, N_2 such that $N_1 + N_2 = N$ and $N_1 r/K$, $N_2 r/K$ are integers. Let us first prove Theorem 7 for the simpler case $\min_{j=1,2} N_j < 40N^{5/6}$. In this case starting from the graph $\mathbb{G}(N, r)$, we can obtain a disjoint union of graphs $\mathbb{G}(N_j, r)$ via at most $O(N^{5/6})$ hyperedge deletion and addition operations. Indeed, suppose without the loss of generality that $N_1 < 40N^{5/6}$. Delete all the hyperedges inside $[N_1]$ as well as all the hyperedges connecting two parts. Then generate a random graph $\mathbb{G}(N_1, r)$ from scratch. Finally, complete a so-obtained partial matching in the configuration model on $[N_2 r]$ and project. The total number of deleted and generated hyperedges is $O(N^{5/6})$, and indeed we obtain a disjoint union of graphs $\mathbb{G}(N_j, r)$, $j = 1, 2$. Since the hyperedge deletion and generation operation changes the value of H by at most $O(N^{5/6})$, then the proof of (17) follows.

Now, throughout the remainder of the section we assume $\min_{j=1,2} N_j \geq 40N^{5/6}$. Fix $T = Nr/K - \lfloor (1/K)N^{2/3} \rfloor$, and consider the graph $\mathbb{G}(N, r, T)$. Note that $Nr/K - T = O(N^{2/3})$.

We now describe an interpolation procedure which interpolates between $\mathbb{G}(N, r, T)$ and a union of certain two graphs on nodes $[N_1]$ and $[N_2]$, each of which will be "nearly" regular. For every integer partition $K = K_1 + K_2$ such that $K_1, K_2 \geq 1$ let $T_{K_1,K_2} \leq T$ be the (random) number of hyperedges which connect parts $[N_1]$ and $[N_2]$ in $\mathbb{G}(N, r, T)$ and such that each connecting hyperedge has exactly K_j nodes in part $[N_j r]$ in the configuration model. Let $T_0 = \sum_{K_1,K_2 \geq 1 : K_1+K_2=K} T_{K_1,K_2}$. Observe that $T_0 \leq \min_{j=1,2}(N_j r)$.

Define $\mathbb{G}(N, T, 0) = \mathbb{G}(N, r, T)$ and define $\mathbb{G}(N, T, t)$, $1 \leq t \leq T_{1,K-1}$, recursively as follows. Assuming $\mathbb{G}(N, T, t-1)$ is already defined, consider the graph $\mathbb{G}_0$ obtained from $\mathbb{G}(N, T, t-1)$ by deleting a hyperedge connecting $[N_1]$ and $[N_2]$ chosen uniformly at random from the collection of hyperedges which have exactly 1 node in part $[N_1 r]$ and $K - 1$ nodes in part $[N_2 r]$ [from the remaining $T_{1,K-1} - (t-1)$ such hyperedges]. Then we construct $\mathbb{G}(N, T, t)$ by adding a hyperedge to the resulting graph as follows: with probability $1/K$ a hyperedge is added to connect K isolated nodes chosen uniformly at random among the isolated nodes from the set $[N_1 r]$. With the remaining probability $(K-1)/K$ a hyperedge is added to connect K isolated nodes chosen uniformly at random among the isolated nodes from the set $[N_2 r]$. It is possible that at some point there are no K isolated nodes available in $[N_j r]$. In this case we say that the interpolation procedure fails. In fact we say that the interpolation procedure fails if in either of the two parts the number of isolated nodes is strictly less than K, even if the attempt was made to add a hyperedge to a part where there is no shortage of such nodes.

Thus we have defined an interpolation procedure for $t \leq T_{1,K-1}$. Assuming that the procedure did not fail for $t \leq T_{1,K-1}$, we now define it for $T_{1,K-1} + 1 \leq t \leq T_{2,K-2}$ analogously: we delete a randomly chosen hyperedge connecting two parts such that the hyperedge has 2 nodes in part $j = 1$, and $K - 2$ nodes in part $j = 2$. Then we add a hyperedge uniformly at random to part $j = 1, 2$ to connect K isolated nodes with probability $2/K$ and $(K-2)/K$, respectively. The failure of the interpolation is defined similarly as above. We continue this for all partitions (K_1, K_2) until $(K-1, 1)$, inclusive. For the (K_1, K_2) phase of the interpolation procedure the probabilities are K_1/K and K_2/K, respectively.

The interpolation procedure is particularly easy to understand in the special case $K = 2$. In this case $T_0 = T_{1,1}$, and there is only one phase in the interpolation procedure. In this phase every edge (which is simply a pair of nodes) connecting sets $[N_1 r]$ and $[N_2 r]$ (if any exists) is deleted and replaces by an edge connecting two isolated nodes in $[N_1 r]$ with probability $1/2$, or two isolated nodes in $[N_2 r]$ with probability $1/2$ as well. One might note the difference of probabilities $1/2$ and $1/2$ for the case of regular graphs vs. N_j/N, $j = 1, 2$, for the case of Erdös–Rényi graph. The difference stems from the regularity assumption of the graph $\mathbb{G}(N, r)$.

Let $\mathcal{I}_t$ be the event that the interpolation succeeds for the first t steps, and let $\mathcal{I} \triangleq \bigcap_{t \le T_0} \mathcal{I}_t$ denote the event that the interpolation procedure succeeds for all steps. For simplicity, even if the interpolation procedure fails in some step t', we still define $\mathbb{G}(N, T, t)$, $t' \le t \le T_0$ to be the same graph as the first graph at which the interpolation procedure fails, $\mathbb{G}(N, T, t) = \mathbb{G}(N, T, t')$. It will be also convenient to define $\mathbb{G}(N, T, t) = \mathbb{G}(N, T, T_0)$ for $T_0 \le t \le \min_{j=1,2}(N_j r)$, whether the interpolation procedure fails or not. This is done in order to avoid dealing with graphs observed at a random (T_0) time, as opposed to the deterministic time $\min_{j=1,2}(N_j r)$.

Provided that the interpolation procedure succeeds, the graph $\mathbb{G}(N, T, \min_{j=1,2} N_j r)$ is a disjoint union of two graphs on $[N_j]$, $j = 1, 2$, each "close" to being an r-regular random graph, in some appropriate sense to be made precise later.

Our next goal is establishing the following analogue of Proposition 2. As in previous sections, let $\mathbb{G}_0$ denote the graph obtained from $\mathbb{G}(N, T, t-1)$ after deleting a hyperedge connecting two parts, but before a hyperedge is added to one of the parts, namely, before creating $\mathbb{G}(N, T, t)$, conditioned on $t \le T_0$ and the event that the interpolation process succeeds till $t - \bigcap_{t' \le t} \mathcal{I}_{t'}$. If, on the other hand the interpolation procedure fails before t, let $\mathbb{G}_0$ be the graph obtained at the last successful interpolation step after the last hyperedge deletion. Let Δ_i denotes the degree of the node $i \in [N]$ in the graph $\mathbb{G}_0$, and let

$$Z_j(t) \triangleq \sum_{i \in [N_j]} (r - \Delta_i)$$

denote the number of isolated nodes in the jth part of the configuration model for $\mathbb{G}_0$ for $j = 1, 2$.

PROPOSITION 4. *For every $t \le \min_j N_j r$,*

$$\mathbb{E}\big[H\big(\mathbb{G}(N, T, t-1)\big)\big] \ge \mathbb{E}\big[H\big(\mathbb{G}(N, T, t)\big)\big] - O\left(\mathbb{E} \max_{j=1,2} \frac{1}{Z_j(t)}\right). \tag{18}$$

PROOF. The claim is trivial when $T_0 + 1 \le t$, since the graph remains the same. Notice also that

$$\mathbb{E}\big[H\big(\mathbb{G}(N, T, t-1)\big)|\mathcal{I}_{t-1}^c\big] = \mathbb{E}\big[H\big(\mathbb{G}(N, T, t)\big)|\mathcal{I}_{t-1}^c\big],$$

since the two graphs are identical, and thus the statement of the proposition holds.

Now we will condition on the event $\mathcal{I}_t$. We now establish a stronger result. Namely,

$$\mathbb{E}\big[H\big(\mathbb{G}(N,T,t-1)\big)|\mathbb{G}_0\big] \geq \mathbb{E}\big[H\big(\mathbb{G}(N,T,t)\big)|\mathbb{G}_0\big] - O\biggl(\max_{j=1,2}\frac{1}{Z_j(t)}\biggr). \tag{19}$$

Observe that conditioned on obtaining graph $\mathbb{G}_0$, the graph $\mathbb{G}(N,T,t-1)$ can be recovered from $\mathbb{G}_0$ in distributional sense by adding a hyperedge connecting K_1 isolated nodes from $[N_1 r]$ to K_2 isolated nodes from $[N_2 r]$, all chosen uniformly at random, and then projecting.

We now conduct model-dependent, case-by-case analysis.

Independent sets. In this case $K=2$, and the only possibility is $K_1=K_2=1$. As in the previous section, O^* again denote the set of nodes in $[N]$ which belong to *every* largest independent set in $\mathbb{G}_0$. Then in the case of creating graph $\mathbb{G}(N,T,t-1)$ from $\mathbb{G}_0$, the newly added edge e decreases H by one if both ends of e belong to O^*, and leaves it the same otherwise. The first event occurs with probability

$$\begin{aligned}
&\frac{\sum_{i_1\in O^*\cap[N_1],i_2\in O^*\cap[N_2]}(r-\Delta_{i_1})(r-\Delta_{i_2})}{\sum_{i_1\in[N_1],i_2\in[N_2]}(r-\Delta_{i_1})(r-\Delta_{i_2})}\\
&\qquad=\frac{\sum_{i\in O^*\cap[N_1]}(r-\Delta_i)}{\sum_{i\in[N_1]}(r-\Delta_i)}\frac{\sum_{i\in O^*\cap[N_2]}(r-\Delta_i)}{\sum_{i\in[N_2]}(r-\Delta_i)}.
\end{aligned}$$

We now analyze the case of creating $\mathbb{G}(N,T,t)$. Conditioning on the event that e was added to part $[N_j r]$, the value of H decreases by one if and only if both ends of e fall into $O^*\cap[N_j]$. This occurs with probability

$$\begin{aligned}
&\frac{(\sum_{i\in[O^*\cap N_j]}(r-\Delta_i))^2-\sum_{i\in O^*\cap[N_j]}(r-\Delta_i)}{(\sum_{i\in[N_j]}(r-\Delta_i))^2-\sum_{i\in[N_j]}(r-\Delta_i)}\\
&\qquad=\frac{(\sum_{i\in[O^*\cap N_j]}(r-\Delta_i))^2}{(\sum_{i\in[N_j]}(r-\Delta_i))^2}-O\biggl(\frac{1}{\sum_{i\in[N_j]}(r-\Delta_i)}\biggr).
\end{aligned}$$

Therefore, the value of H decreases by one with probability

$$\frac{1}{2}\sum_{j=1,2}\frac{(\sum_{i\in[O^*\cap N_j]}(r-\Delta_i))^2}{(\sum_{i\in[N_j]}(r-\Delta_i))^2}-O\biggl(\max_{j=1,2}\frac{1}{Z_j(t)}\biggr)$$

and stays the same with the remaining probability. Using the inequality $(1/2)(x^2+y^2)\geq xy$, we obtain (19).

MAX-CUT, *Ising*, *coloring*. As in the proof of Theorem 1, we introduce equivalence classes $O_j^*\subset[N]$, $1\leq j\leq J$, on the graph $\mathbb{G}_0$. The rest of the proof is almost identical to the one for the independent set model, and we skip the details.

Notice that in all of these cases we have $K = 2$, and the interpolation phase has only one stage corresponding to $(K_1, K_2) = (1, 1)$.

K-SAT. This is the first model for which $K > 2$. We fix $K_1, K_2 \geq 1$ such that $K_1 + K_2 = K$ and further condition on the event that the graph $\mathbb{G}_0$ was created in stage (K_1, K_2). As in the previous section, let O^* be the set of frozen variables in all optimal assignments of $\mathbb{G}_0$. Reasoning as in the previous section, when we reconstruct graph $\mathbb{G}(N, T, t-1)$ in the distributional sense by adding a random hyperedge connecting K_1 nodes in $[N_1 r]$ with K_2 nodes in $[N_2 r]$, the probability that the value of H remains the same (does not increase by one) is precisely

$$\frac{1}{2^K}\left[\frac{\sum_{i\in O^*\cap[N_1]}(r-\Delta_i)}{\sum_{i\in[N_1]}(r-\Delta_i)}\right]^{K_1}\left[\frac{\sum_{i\in O^*\cap[N_2]}(r-\Delta_i)}{\sum_{i\in[N_2]}(r-\Delta_i)}\right]^{K_2}. \tag{20}$$

Similarly, creating $\mathbb{G}(N, T, t)$ from $\mathbb{G}_0$ keeps the value of H the same with probability

$$\begin{aligned}&\frac{1}{2^K}\frac{K_1}{K}\left[\frac{\sum_{i\in O^*\cap[N_1]}(r-\Delta_i)}{\sum_{i\in[N_1]}(r-\Delta_i)}\right]^{K}+\frac{1}{2^K}\frac{K_2}{K}\left[\frac{\sum_{i\in O^*\cap[N_2]}(r-\Delta_i)}{\sum_{i\in[N_2]}(r-\Delta_i)}\right]^{K}\\&\quad-O\left(\max_{j=1,2}\frac{1}{\sum_{i\in[N_j]}(r-\Delta_i)}\right).\end{aligned} \tag{21}$$

Applying Young's inequality, namely that $ab \leq pa^{1/p} + qb^{1/q}$ for every $a, b \geq 0$, $p + q = 1$, $p, q > 0$, with the choice $p = K_1/K, q = K_2/K$,

$$a = \left[\frac{\sum_{i\in O^*\cap[N_1]}(r-\Delta_i)}{\sum_{i\in[N_1]}(r-\Delta_i)}\right]^{K_1},$$

$$b = \left[\frac{\sum_{i\in O^*\cap[N_2]}(r-\Delta_i)}{\sum_{i\in[N_2]}(r-\Delta_i)}\right]^{K_2},$$

and canceling $1/2^K$ on both sides, we obtain the result.

NAE-K-SAT. The proof is similar to the one for K-SAT and for NAE-K-SAT for the $\mathbb{G}(N, \lfloor cN \rfloor)$ model. This completes the proof of the proposition. □

Our next step is to control the error term in (18).

LEMMA 3. *The interpolation procedure succeeds* (*event $\mathcal{I}$ holds*) *with probability at least* $1 - O(N\exp(-N^{\delta}))$ *for some* $\delta > 0$. *Additionally*,

$$\mathbb{E}\left[\sum_{1\leq t\leq T_0}\max_{j=1,2}\frac{1}{Z_j(t)}\right] = O(N^{2/5}). \tag{22}$$

PROOF. Since $\mathbb{G}_0$ is obtained after deleting one hyperedge connecting two parts, but before adding a new hyperedge, then $Z_j(t) \geq 1$. A crude bound on the required expression is then $\mathbb{E}[T_0] = O(\min N_j)$. We have $\mathbb{E}[Z_j(0)] = (N_j/N)N^{2/3} = N_j/N^{1/3} \geq 40N^{1/2}$ since the initial number of isolated nodes was $Nr/K - T = N^{2/3}$ and $\min_j N_j \geq 40N^{5/6}$. Moreover, using a crude concentration bound $\mathbb{P}(Z_j(0) < (1/2)(N_j/N^{1/3}) + K) = O(\exp(-N^{\delta_1}))$ for some $\delta_1 > 0$. Observe that $Z_j(t+1) - Z_j(t) = 0$ with probability one if the interpolation procedure failed for some $t' \leq t$. Otherwise, if t corresponds to phase (K_1, K_2), then $Z_j(t+1) - Z_j(t)$ takes values $-K_j + K$ with probability K_j/K and $-K_j$ with the remaining probability. This is because during the hyperedge deletion step, $Z_j(t)$ decreases by K_j, and during the hyperedge addition step, it increases by K or by zero with probabilities K_j/K and $1 - K_j/K$, respectively. In particular, $\mathbb{E}[Z_j(t+1) - Z_j(t)] = 0$. The decision of whether to put the hyperedge into part 1 or 2 is made independently. Since $t \leq T_0 \leq N_j$, we conclude that for each $t \leq T_0$ we have $\mathbb{P}(Z_j(0) - Z_j(t) > N_j^{3/5}) = O(\exp(-N^{\delta_2}))$ for some $\delta_2 > 0$. Here any choice of exponent strictly larger than $1/2$ applies, but for our purposes $3/5$ suffices. It follows that $Z_j(t) \geq (1/2)N_j/N^{1/3} + K - N_j^{3/5}$ for all t with probability $1 - O(N_j \exp(-N^{\delta})) = 1 - O(N \exp(-N^{\delta}))$ for $\delta = \min(\delta_1, \delta_2)$. The assumption $\min N_j \geq 40N^{5/6}$ implies that a weaker bound $\min N_j \geq 32^{1/2}N^{5/6}$, which translates into $(1/2)N_j/N^{1/3} - N_j^{3/5} \geq 0$. Thus $Z_j(t) \geq K$ for all t, with probability $1 - O(N \exp(-N^{\delta}))$, and therefore the interpolation procedure succeeds.

Now ignoring term K in the expression $(1/2)N_j/N^{1/3} + K - N_j^{3/5}$ and using $T_0 \leq \min_j(N_j r)$, we obtain that with probability $1 - O(N \exp(-N^{\delta}))$, the expression inside the expectation on the left-hand side of (22) is at most

$$\frac{N_j r}{(1/2)N_j N^{-1/3} - N_j^{3/5}} = \frac{N_j^{2/5} r}{(1/2)N_j^{2/5} N^{-1/3} - 1}.$$

The numerator is at most $N^{2/5}r$. Also the assumption $\min N_j \geq 40N^{5/6}$ implies that the denominator is at least 1. We conclude that the expression inside the expectation is at most $N^{2/5}r$ with probability at least $1 - O(N \exp(-N^{\delta}))$. Since we also have $T_0 \leq Nr$ w.p.1, then using a very crude estimate $O(N \exp(-N^{\delta})) = O(N^{-3/5})$, and $NN^{-3/5} = N^{2/5}$, we obtain the required result. □

As a corollary of Proposition 4 and Lemma 3 we obtain

COROLLARY 2.

$$\mathbb{E}[H(\mathbb{G}(N, T, 0))] \geq \mathbb{E}\Big[H\Big(\mathbb{G}\Big(N, T, \min_{j=1,2} N_j r\Big)\Big)\Big] - O(N^{2/5}).$$

Let us consider graph $\mathbb{G}(N, T, T_0)$. We further modify it by removing all hyperedges which connect two parts $[N_j]$ of the graph, if there are any such hyperedges left. Notice that if the event $\mathcal{I}$ occurs, namely the interpolation procedure succeeds, no further hyperedges need to be removed. The resulting graph is a disjoint union of graphs obtained on nodes $[N_1 r]$ and $[N_2 r]$ by adding a random size partial matching uniformly at random. The actual size of these two matchings depends on in the initial size of the partial matching within each part, and also on how many of T_0 hyperedges go into each part during the interpolation steps, and how many were removed in the final part (if any). We now obtain bounds on the sizes of these matchings.

Recall $\min_j N_j \geq 40N^{5/6}$. We showed in the proof of Lemma 3 that the interpolation procedure succeeds with probability $O(N\exp(-N^{\delta}))$ for some δ. This coupled with the fact that w.p.1, the number of hyperedges removed in the final stage is at most rN/K, gives us that the expected number of hyperedges removed in the final stage is at most $O(N^2\exp(-N^{\delta}))$ which (as a very crude estimate) is $O(N^{2/3})$. Moreover, since the initial number of isolated nodes was $N^{2/3}$, and during the interpolation procedure the total number of isolated nodes in the entire graph never increases, then the total number of isolated nodes before the final removal of hyperedges in $\mathbb{G}(N, T, T_0)$ is at most $N^{2/3}$. We conclude that the expected number of isolated nodes in the end of the interpolation procedure is $O(N^{2/3})$. Then we can complete uniform random partial matchings on $[N_j r]$ to a perfect uniform random matchings by adding at most that many hyperedges in expectation. The objective value of H changes by at most that much as well. The same applies to $\mathbb{G}(N, r, T)$—we can complete the configuration model corresponding to this graph to a perfect matching on Nr nodes by adding at most $N^{2/3}$ hyperedges since $Nr/K - T = O(N^{2/3})$. Coupled with Corollary 2 we then obtain

$$\mathbb{E}\big[H\big(\mathbb{G}(N, r)\big)\big] \geq \mathbb{E}\big[H\big(\mathbb{G}(N_1, r)\big)\big] + \mathbb{E}\big[H\big(\mathbb{G}(N_2, r)\big)\big] - O\big(N^{2/3}\big)$$

for the case $\min_j N_j \geq 40N^{5/6}$. This completes the proof of Theorem 7. $\square$

PROOF OF THEOREM 3. The existence of the limit

$$\lim_{N\to\infty, N\in r^{-1}K\mathbb{Z}_+} N^{-1}\mathbb{E}\big[H\big(\mathbb{G}(N, r)\big)\big] = H(r)$$

follows immediately from Theorem 7 and Proposition 5 from Appendix B. Then the convergence w.h.p.

$$\lim_{N\to\infty, N\in r^{-1}K\mathbb{Z}_+} N^{-1}H\big(\mathbb{G}(N, r)\big) = H(r)$$

follows once again using standard concentration results [24]. $\square$

The proof of Theorem 4 uses the same interpolation as the one above, and the proof itself mimics the one for Theorem 2. For this reason, we omit the details.

APPENDIX A: PROOF OF LEMMA 1

We first assume K-SAT or NAE-K-SAT models. Let us show for these models that there exists a constant $\omega \geq 1/2$ such that for *every* graph and potential realization $(\mathbb{G} = (V, E), H)$ such that the problem is satisfiable [namely $H(\mathbb{G}) = |E|$], if a randomly chosen hyperedge e is added with a potential chosen according to the model, then

$$\mathbb{P}\big(H(\mathbb{G} + e) = |E| + 1\big) \geq \omega.$$

In other words, if the current graph is satisfiable, the new graph obtained by adding a random hyperedge remains satisfiable with probability at least ω. Indeed, for example, for the case of K-SAT, if the instance is satisfiable and x is a satisfying assignment, the added edge remains consistent with x with probability at least $\omega \triangleq 1 - 1/2^K > 1/2$. For the case of NAE-K-SAT it is $\omega = 1 - 1/2^{K-1} \geq 1/2$. We obtain that for every positive M, m and recalling assumption $\delta < 1/2$,

$$p(N, M + m) \geq \omega^m p(N, M) \geq \delta^m p(N, M),$$

and the assertion is established.

The proof for the case of coloring is more involved. Given $0 < \delta < 1/2$ we call a graph $\mathbb{G}$ on N nodes δ-unusual if it is colorable, and in every coloring assignment there exists a color class with size at least $(1 - \delta)N$. Namely, for every color assignment x such that $H(x) = |E|$, there exists $k \in [q^-]$ such that the cardinality of the set $\{i \in [N] : x_i = k\}$ is at least $(1 - \delta)N$. We claim that

$$\mathbb{P}\big(\mathbb{G}(N, M) \text{ is } \delta\text{-unusual}\big) \leq (2\delta)^M \exp\big(H(\delta)N + o(N)\big). \tag{23}$$

The claim is shown using the first moment method—we will show that the expected number of graphs with such a property is at most the required quantity. Indeed, given a subset $C \subset [N]$ such that $|C| \geq (1 - \delta)N$, the probability that the graph $\mathbb{G}(N, M)$ has proper coloring with all nodes in C having the same color is at most $(1 - (1 - \delta)^2)^M < (2\delta)^M$, since we must have that no edge falls within the class C. There are at most $\binom{N}{\delta N} = \exp(H(\delta)N + o(N))$ choices for the subset C. The claim then follows.

Now observe that if a graph $\mathbb{G} = (V, E)$ is colorable but not δ-unusual, then adding a random edge e, we obtain $\mathbb{P}(H(\mathbb{G} + e) = |E| + 1) \geq \delta(1 - \delta)$. Namely, in this case the probability goes down by at most a constant factor. We obtain

$$\begin{aligned}
&p(N, M + 1)\\
&\quad \geq \mathbb{P}\big(H\big(\mathbb{G}(N, M + 1) = M + 1\big) | \mathbb{G}(N, M) \text{ colorable, not } \delta\text{-unusual}\big)\\
&\qquad \times \mathbb{P}\big(\mathbb{G}(N, M) \text{ colorable and not } \delta\text{-unusual}\big)\\
&\quad \geq \delta(1 - \delta)\mathbb{P}\big(\mathbb{G}(N, M) \text{ colorable and not } \delta\text{-unusual}\big)\\
&\quad \geq \delta(1 - \delta)\mathbb{P}\big(\mathbb{G}(N, M) \text{ colorable}\big) - \delta(1 - \delta)\mathbb{P}\big(\mathbb{G}(N, M)\ \delta\text{-unusual}\big)\\
&\quad \geq \delta(1 - \delta)p(N, M) - \delta\mathbb{P}\big(\mathbb{G}(N, M)\ \delta\text{-unusual}\big)\\
&\quad \geq \delta(1 - \delta)p(N, M) - \delta(2\delta)^M \exp\big(H(\delta)N + o(N)\big),
\end{aligned}$$

using the earlier established claim. Iterating this inequality, we obtain for every $m \geq 1$,

$$\begin{aligned} p(N, M+m) &\geq \delta^m (1-\delta)^m p(N, M) - \delta(2\delta)^M \exp\bigl(H(\delta)N + o(N)\bigr) \sum_{0 \leq j \leq m-1} \delta^m (1-\delta)^m \\ &\geq \delta^m \mathbb{P}\bigl(H\bigl(\mathbb{G}(N, M) = M\bigr)\bigr) - (2\delta)^{M+1} \exp\bigl(H(\delta)N + o(N)\bigr), \end{aligned}$$

where $\sum_{0 \leq j \leq m} \delta^m (1-\delta)^m \leq 1/(1-\delta) < 2$ is used in the last inequality. This completes the proof of the lemma.

APPENDIX B: MODIFIED SUPER-ADDITIVITY THEOREM

To keep the proof of our main results self-contained, we state and prove the following proposition, used in proving several of the theorems presented in the earlier sections. However, Béla Bollobás and Zoltan Füredi kindly pointed out to us that the following proposition is a special case of a more general and classical theorem of de Bruijn and Erdös (see Theorem 22 on page 161 in [12]), which uses a weaker assumption on the additive term in the near super-additivity hypothesis; also see [11] and the Bollobás–Riordan percolation book [7] for more recent applications of this useful tool.

PROPOSITION 5. *Given* $\alpha \in (0, 1)$, *suppose a nonnegative sequence* a_N, $N \geq 1$ *satisfies*

$$a_N \geq a_{N_1} + a_{N_2} - O\bigl(N^\alpha\bigr) \tag{24}$$

for every N_1, N_2 *s.t.* $N = N_1 + N_2$. *Then the limit* $\lim_{N\to\infty} \frac{a_N}{N}$ *exists.*

PROOF. It is convenient to define $a_N = a_{\lfloor N \rfloor}$ for every real, but not necessarily integer value $N \geq 1$. It is then straightforward to check that property (24) holds when extended to reals as well [thanks to the correction term $O(N^\alpha)$]. Let

$$a^* = \limsup_{N\to\infty} \frac{a_N}{N}.$$

Fix $\varepsilon > 0$ and find k such that $1/k < \varepsilon \leq 1/(k-1)$. Find find $N_0 = N_0(\varepsilon)$ such that $N_0^{-1} a_{N_0} \geq a^* - \varepsilon$, $k^\alpha N_0^{\alpha-1} < \varepsilon$. Clearly, such N_0 exists. Consider any $N \geq kN_0$. Find r such that $kN_0 2^r \leq N \leq kN_0 2^{r+1}$. Applying (24) iteratively with $N_1 = N_2 = N/2$ we obtain

$$\begin{aligned} a_N &\geq 2^r a_{N/2^r} - \sum_{0 \leq l \leq r-1} O\biggl(2^l \biggl(\frac{N}{2^l}\biggr)^\alpha\biggr) \\ &= 2^r a_{N/2^r} - O\bigl(2^{(1-\alpha)r} N^\alpha\bigr). \end{aligned}$$

Now let us find i such that $(k+i)N_0 \le N/2^r \le (k+i+1)N_0$. Note $i \le k$. Again using (24) successively with N_0 for N_1 and $N/2^r$, $(N/2^r)-N_0$, $(N/2^r)-2N_0, \ldots$ for N_2, we obtain

$$\begin{aligned} a_{N/2^r} &\ge (k+i)a_{N_0} - O\biggl(k\biggl(\frac{N}{2^r}\biggr)^{\alpha}\biggr) \\ &\ge (k+i)a_{N_0} - O\biggl(k\biggl(\frac{N}{2^r}\biggr)^{\alpha}\biggr). \end{aligned}$$

Combining, we obtain

$$\begin{aligned} a_N &\ge 2^r(k+i)a_{N_0} - O\bigl(2^{(1-\alpha)r}N^{\alpha}\bigr) - O\bigl(k2^{r(1-\alpha)}N^{\alpha}\bigr) \\ &= 2^r(k+i)a_{N_0} - O\bigl(k2^{r(1-\alpha)}N^{\alpha}\bigr). \end{aligned}$$

Then

$$\begin{aligned} \frac{a_N}{N} &\ge \frac{2^r(k+i)}{2^r(k+i+1)}\frac{a_{N_0}}{N_0} - O\bigl(k2^{r(1-\alpha)}N^{\alpha-1}\bigr) \\ &\ge \biggl(1 - \frac{1}{(k+i+1)}\biggr)\bigl(a^* - \varepsilon\bigr) - O\bigl(k2^{r(1-\alpha)}N^{\alpha-1}\bigr) \\ &\ge (1-\varepsilon)\bigl(a^* - \varepsilon\bigr) - O\bigl(k2^{r(1-\alpha)}N^{\alpha-1}\bigr), \end{aligned}$$

where $1/k < \varepsilon$ is used in the last inequality. Now

$$k2^{r(1-\alpha)}N^{\alpha-1} \le k2^{r(1-\alpha)}\bigl(k2^rN_0\bigr)^{\alpha-1} = k^{\alpha}N_0^{\alpha-1} < \varepsilon,$$

again by the choice of N_0. We have obtained

$$\frac{a_N}{N} \ge (1-\varepsilon)\bigl(a^* - \varepsilon\bigr) - \varepsilon$$

for all $N \ge N_0 k$. Since ε was arbitrary the proof is complete. $\square$

Acknowledgments. The authors are grateful for the insightful discussions with Silvio Franz, Andrea Montanari, Lenka Zdeborová, Florant Krzakala, Jeff Kahn and James Martin. The authors thank Zoltan Füredi and Béla Bollobás for bringing the deBruijn–Erdös theorem and other relevant literature, mentioned in Appendix B, to the authors' attention. The authors are especially thankful to anonymous referees for helpful technical comments and notation suggestions. Authors also thank Microsoft Research New England for the hospitality and the inspiring atmosphere, in which this work began.

REFERENCES

[1] ALDOUS, D. Open problems. Preprint. Available at: http://www.stat.berkeley.edu/~aldous/Research/OP/sparse_graph.html.

[2] ALDOUS, D. Some open problems. Preprint. Available at: http://stat-www.berkeley.edu/users/aldous/Research/problems.ps.

[3] ALDOUS, D. and STEELE, J. M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures* (H. Kesten, ed.) 1–72. Springer, Berlin.

[4] ALON, N. and SPENCER, J. H. (1992). *The Probabilistic Method*. Wiley, New York. MR1140703

[5] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929

[6] BOLLOBÁS, B. (2001). *Random Graphs*. Cambridge Univ. Press, Cambridge.

[7] BOLLOBÁS, B. and RIORDAN, O. (2006). *Percolation*. Cambridge Univ. Press, New York. MR2283880

[8] BOLLOBÁS, B. and RIORDAN, O. (2011). Sparse graphs: Metrics and random models. *Random Structures Algorithms* **39** 1–38. MR2839983

[9] COJA-OGHLAN, A. (2012). Personal communication.

[10] COPPERSMITH, D., GAMARNIK, D., HAJIAGHAYI, M. T. and SORKIN, G. B. (2004). Random MAX SAT, random MAX CUT, and their phase transitions. *Random Structures Algorithms* **24** 502–545. MR2060633

[11] DE BRUIJN, N. G. and ERDÖS, P. (1951). Some linear and some quadratic recursion formulas. I. *Indag. Math.* (*N.S.*) **13** 374–382. MR0047161

[12] DE BRUIJN, N. G. and ERDÖS, P. (1952). Some linear and some quadratic recursion formulas. II. *Indag. Math.* (*N.S.*) **14** 152–163. MR0047162

[13] DEMBO, A. and MONTANARI, A. (2010). Ising models on locally tree-like graphs. *Ann. Appl. Probab.* **20** 565–592.

[14] DEMBO, A. and MONTANARI, A. (2010). Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24** 137–211. MR2643563

[15] DEMBO, A., MONTANARI, A. and SUN, N. (2011). Factor models on locally tree-like graphs. Available at arXiv:1110.4821.

[16] FRANZ, S. and LEONE, M. (2003). Replica bounds for optimization problems and diluted spin systems. *J. Stat. Phys.* **111** 535–564. MR1972121

[17] FRANZ, S. and LEONE, M. (2003). Replica bounds for optimization problems and diluted spin systems. *J. Phys. A Math. Gen.* **36** 10967–10985.

[18] FRANZ, S. and MONTANARI, A. (2009). Personal communication.

[19] FRIEDGUT, E. (1999). Sharp thresholds of graph properties, and the k-sat problem. *J. Amer. Math. Soc.* **12** 1017–1054. With an appendix by Jean Bourgain. MR1678031

[20] GALLAGER, R. G. (1963). *Low-Density Parity-Check Codes*. MIT Press, Cambridge, MA.

[21] GAMARNIK, D. (2004). Linear phase transition in random linear constraint satisfaction problems. *Probab. Theory Related Fields* **129** 410–440. MR2128240

[22] GAMARNIK, D., NOWICKI, T. and SWIRSZCZ, G. (2006). Maximum weight independent sets and matchings in sparse random graphs. Exact results using the local weak convergence method. *Random Structures Algorithms* **28** 76–106. MR2187483

[23] GUERRA, F. and TONINELLI, F. L. (2002). The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** 71–79. MR1930572

[24] JANSON, S., ŁUCZAK, T. and RUCINSKI, A. (2000). *Random Graphs*. Wiley, New York. MR1782847

[25] JANSON, S. and THOMASON, A. (2008). Dismantling sparse random graphs. *Combin. Probab. Comput.* **17** 259–264. MR2396351

[26] MONTANARI, A. (2005). Tight bounds for LDPC and LDGM codes under MAP decoding. *IEEE Trans. Inform. Theory* **51** 3221–3246. MR2239148

[27] PANCHENKO, D. and TALAGRAND, M. (2004). Bounds for diluted mean-fields spin glass models. *Probab. Theory Related Fields* **130** 319–336. MR2095932

[28] WORMALD, N. C. (1999). Models of random regular graphs. In *Surveys in Combinatorics, 1999 (Canterbury)*. *London Mathematical Society Lecture Note Series* **267** 239–298. Cambridge Univ. Press, Cambridge. MR1725006

M. BAYATI
GRADUATE SCHOOL OF BUSINESS
STANFORD UNIVERSITY
655 KNIGHT WAY
STANFORD, CALIFORNIA 94305
USA
E-MAIL: bayati@stanford.edu

D. GAMARNIK
MIT SLOAN SCHOOL OF MANAGEMENT
100 MAIN STREET
CAMBRIDGE, MASSACHUSETTS 02139
USA
E-MAIL: gamarnik@mit.edu

P. TETALI
SCHOOL OF MATHEMATICS
AND SCHOOL OF COMPUTER SCIENCE
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332
USA
E-MAIL: tetali@math.gatech.edu