Combinatorial Approach to the Interpolation Method and Scaling Limits in Sparse Random Graphs

Mohsen Bayati * David Gamarnik † Prasad Tetali ‡

Abstract

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 [Alda] as one of his six favorite open problems. It was also mentioned as an open problem in several other places: Conjecture 2.20 in [Wor99], [BRar],[JT08] and [AS03]).

Our approach is based on extending and simplifying the interpolation method of Guerra and Toninelli [FF02] and Franz and Leone [FL03]. 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 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 [Wor99], [GNS06], [BRar], as well as [JT08] and [AS03] that $|\mathcal{I}_N|/N$ converges in probability as $N \to \infty$. Additionally, this problem was listed by D. Aldous as one of his six favorite open problems [Alda]. (For a new collection of Aldous' favorite open problems see [Aldb]). 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.

*Stanford University; e-mail: bayati@stanford.edu.

†MIT; e-mail: gamarnik@mit.edu. Research supported in part by the NSF grants CMMI-1031332.

‡Georgia Tech; e-mail: tetali@math.gatech.edu. Research supported in part by the NSF grants DMS-0701043 and CCR-0910584.

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 at least $1 - 2^{-K}$ fraction of the clauses can be satisfied with high probability (w.h.p.). It was conjectured in [CGHS04] 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 [FF02] and further developed by Franz and Leone [FL03], Franz, Leone and Toninelli [FLT03], Panchenko and Talagrand [PT04], and Montanari [Mon05] 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 [FL03], 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 also later resolved using the same method.

2 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 model, 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 non-binary 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 independent set problem, thus resolving an open problem earlier stated.

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 [FLT03] for the K-SAT and Viana-Bray models with even number of variables per clause, and Montanari [Mon05] 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 [FLT03] 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 [Fri99] states that for each of these models there exists a (model dependent) critical value c^* such that for every $\epsilon > 0$, when the number of edges (or clauses for a SAT-type problem) is at most $(c^* - \epsilon)N$, the model is colorable (satisfiable) w.h.p. and when it is at least $(c^* + \epsilon)N$, it is not colorable (not satisfiable) w.h.p. as $N \to \infty$. Friedgut [Fri99] 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 $(c^*_N - \epsilon)N$ and is not colorable (satisfiable) w.h.p. when the number of edges (clauses) is at least $(c^*_N + \epsilon)N$. It is also reasonable to conjecture, and indeed was shown for 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 establish the large deviations principle 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 $\epsilon > 0$, $p(c^*_N + \epsilon, N)$ converges to zero exponentially fast, where c^*_N is the same sequence as in Friedgut's theorem.

References

[Alda] D. Aldous, *Some open problems*, http://stat-www.berkeley.edu/users/aldous/ Research/problems.ps.

[Aldb] ______, *Some open problems*, http://www.stat.berkeley.edu/∼aldous/Research/OP/index.html.

[AS03] D. Aldous and J. M. Steele, *The objective method: Probabilistic combinatorial optimization and local weak convergence*, Discrete Combinatorial Probability, H. Kesten Ed., Springer-Verlag, 2003.

[BRar] B. Bollobás and O. Riordan, *Sparse graphs: metrics and random models*, Random Structures and Algorithms (To appear).

[CGHS04] D. Coppersmith, D. Gamarnik, M. Hajiaghayi, and G. Sorkin, *Random MAXSAT, random MAXCUT, and their phase transitions*, Random Structures and Algorithms **24** (2004), no. 4, 502–545.

[FF02] F.Guerra and F.L.Toninelli, *The thermodynamic limit in mean field spin glass models*, Commun. Math. Phys. **230** (2002), 71–79.

[FL03] S. Franz and M. Leone, *Replica bounds for optimization problems and diluted spin systems*, Journal of Statistical Physics **111** (2003), no. 3/4, 535–564.

[FLT03] S. Franz, M. Leone, and F. L. Toninelli, *Replica bounds for diluted non-Poissonian spin systems*, J. Phys. A: Math. Gen. **36** (2003), 10967 – 10985.

[Fri99] E. Friedgut, *Sharp thresholds of graph proprties, and the k-SAT problem*, J. Amer. Math. Soc. **4** (1999), 1017–1054.

[GNS06] D. Gamarnik, T. Nowicki, and G. Swirscsz, *Maximum weight independent sets and matchings in sparse random graphs. Exact results using the local weak convergence method*, Random Structures and Algorithms **28** (2006), no. 1, 76–106.

[JT08] S. Janson and A. Thomason, *Dismantling sparse random graphs*, Combinatorics, Probability and Computing **17** (2008), 259 – 264.

[Mon05] A. Montanari, *Tight bounds for LDPC and LDGM codes under map decoding*, IEEE Transactions on Information Theory **51** (2005), no. 9, 3221–3246.

[PT04] D. Panchenko and M. Talagrand, *Bounds for diluted mean-fields spin glass models*, Probability Theory and Related Fields **130** (2004), 312–336.

[Wor99] N. Wormald, *Models of random regular graphs*, in J.D. Lamb and D.A. Preece, editors, Surveys in Combinatorics, volume 267 ed., pp. 239–298, Cambridge University Press, 1999.