Average-case complexity of Maximum Weight Independent Set with random weights in bounded degree graphs

David Gamarnik [*] David Goldberg [†] Theophane Weber [‡]

August 3, 2009

Abstract

Finding the largest independent set in a graph is a notoriously difficult NP-complete combinatorial optimization problem. Unlike other NP-complete problems, it does not admit a constant factor approximation algorithm for general graphs. Furthermore, even for graphs with largest degree 3, no polynomial time approximation algorithm exists with a 1.0071-factor approximation guarantee.

We consider the problem of finding maximum weight independent set in bounded degree graph, when the node weights are generated i.i.d. from a common distribution. For instance, we construct a PTAS (Polynomial-Time Approximation Scheme) for the case of exponentially distributed weights and arbitrary graphs with degree at most 3. We generalize the analysis to phase-type distributions (dense in the space of all distributions), and provide partial converse results, showing that even under a random cost assumption, it can be NP-hard to compute the MWIS of a graph with sufficiently large degree. We also show how the method can be used to derive lower bounds for the expected size of maximum independent set in large random regular graphs.

Our algorithm, the cavity expansion, is based on combination of several ideas, including recent deterministic approximation algorithms for counting on graphs and local weak convergence/correlation decay methods.

Extended abstract

The problem of finding the largest independent set of a graph is a well-known NP-complete problem. Moreover, unlike some other NP-complete problems, it does not admit a constant factor approximation algorithm for general graphs: Hastad [Has99] showed that for every $0 < \delta < 1$, no $n^{1-\delta}$ approximation algorithm can exist for this problem unless $P = NP$, where n is the number of nodes. Even for the class of graphs with largest degree at most 3, no factor 1.0071-approximation algorithm can exist, under the same complexity-theoretic assumption (see Berman and Karpinski [BK99]). Similar results (with different constants) were established for the cases of graphs with larger maximum degree . Thus, the problem does not admit any polynomial time approximation scheme (PTAS), even in the least non-trivial class of degree-3 graphs.

[*]Operations Research Center and Sloan School of Management, MIT, Cambridge, MA, 02139, e-mail: `gamarnik@mit.edu`

[†]Operations Research Center and Sloan School of Management, MIT, Cambridge, MA, 02139, e-mail: `dag3141@mit.edu`

[‡]Operations Research Center, MIT, Cambridge, MA, 02139, e-mail: `theo_w@mit.edu`

The problem has also been studied in several average case settings. Karp and Sipser [KS81] constructed a simple algorithm which finds an asymptotically largest independent set in Erdős-Renyie random graphs with average degree $c \leq e$ (in fact the maximum matching problem is considered instead, but the approach works for independent set problem as well in this regime). The largest independent set is known to be order $(2+o(1))\frac{\log c}{c}n)$ in average degree c Erdős-Renyie graphs with n nodes [JŁR00]. Similarly, in dense random graphs with n nodes, where each edge is selected with probability $1/2$, the largest independent set is $2\log n + o(\log n)$, though the largest independent set produced by any known polynomial time algorithm is only $(1+o(1))\log n$. More recently, a different average case model was considered in Gamarnik et al. [GNS06]: the nodes of an Erdős-Renyie graph are equipped with random weights with exponential distribution. The limiting expression for the maximum weight independent set was found in the regime $c \leq 2e$. Similar results were established for r-regular graphs with girth diverging to infinity for the cases $r = 3, 4$.

In this paper we consider the following natural mixture of the worst case/average case assumptions. We consider an arbitrary graph with largest degree at most Δ, where the nodes are eqWe consider an arbitrary graph with largest degree at most 3, where the nodes are equipped with random weights, generated i.i.d. from an exponential distribution with parameter 1. The decision problem is to find the largest weighted independent set. Surprisingly, we discover that this is a tractable problem - we construct a randomized PTAS, even though the unit weight version of this problem (maximum cardinality independent set) does not admit any PTAS, as previously mentioned. We extend this result to more general graphs but for distributions which are mixtures of exponential distributions.

Our algorithm, which we call *Cavity Expansion* is new and draws upon several recent ideas. The first such idea is the local weak convergence method. This is a method which takes advantage of the observation that randomness sometimes induces a long-range independence (correlation decay) in the underlying decision problem. For example, as we show in the present paper, random weights with exponential distribution imply that whether node i belongs to the largest weighted independent set has asymptotically no correlation with whether a node j belongs to the largest weighted independent set, when the distance between i and j is large.

The notion of bonus was heavily used recently in the statistical physics literature under the name *cavity* [MP03] (see [RBMM04] for the independent sets setting). The local-weak convergence/cavity method thus was used extensively but only in the setting of random graphs which have a locally-tree like structure. Thus in order to go beyond the locally-tree like restriction on graphs, another idea is needed. Such an idea was proposed recently by Weitz [Wei06] and extended in Gamarnik and Katz [GK07a],[GK07b], Bayati et al [BGK$^+$07], Jung and Shah [JS06] in the context of graph counting problems. Weitz showed that a problem of counting on a graph and a more general problem of computing the partition function of Gibbs measures, can be reduced to the problem of counting on a related self-avoiding exponential size tree. Then if the correlation decay property can be established on this self-avoiding tree instead, then the tree can be truncated at small depth to obtain approximate inference. Intuitively, this implies that the decision of whether a node should be included in the maximum weighted independent set or not does not depend on the structure of the graph or the weights past a certain distance.

In this paper we combine the correlation decay/local weak convergence approach of [GNS06] and the self-avoiding tree approach of Weitz [Wei06] to obtain the stated result. Our approach does not explicitly use the notion of self-avoiding tree, rather a simpler notion of recursive cavity approximation and expansions are used and the correlation decay property is established. The algorithm is in fact decentralized: only a local constant size neighborhood around each node is used to make a decision whether this should be a part of the independent set.

The exponential distribution is not the only distribution which can be analyzed within this framework. In reality, it serves as a base case to analyze phase-type distributions. For any phase-type distribution, we can characterize correlation decay and, by an analysis similar to the one used for the exponential case, we can identify sufficient correlation decay conditions relating the degree of the graph to the distribution of the weights. In particular we use a family of hyperexponential distribution to exhibit examples of correlation decay in larger degree graphs.

Finally, we can show that the setting with random weights hits a complexity-theoretic barrier just as the classical cardinality problem does. Specifically, we show that for graphs with sufficiently large degree the problem of finding the largest weighted independent set with i.i.d. exponentially distributed weights does not admit any PTAS. This negative result is proven by showing that for large degree graphs largest weighted independent sets are dominated by independents sets with cardinality close to largest possible. Since the latter does not admit a constant factor approximation up to $O(\Delta/2^{\sqrt{O(\Delta)}})$ multiplicative factor, the same will apply to the former case.

Our results further highlight interesting and intriguing connection between the field of complexity of algorithms for combinatorial optimization problems and statistical physics (cavity method, long-range independence). It would be interesting to see if random weights assumptions can be substituted with deterministic weights which have some random like properties, similarly as pseudo-random graphs. This would move our approach even closer to the worst-case combinatorial optimization setting.

References

[BGK+07] M. Bayati, D. Gamarnik, D. Katz, C. Nair, and P. Tetali, *Simple deterministic approximation algorithms for counting matchings*, Proc. 39th Ann. Symposium on the Theory of Computing (STOC), 2007.

[BK99] P. Berman and M. Karpinski, *On some tighter inapproximability results*, Lecture notes in computer science ed., Springer Berlin / Heidelberg, 1999.

[GK07a] D. Gamarnik and D. Katz, *Correlation decay and deterministic FPTAS for counting list-colorings of a graph*, Proceedings of 18th ACM-SIAM Symposium on Discrete Algorithms (SODA), 2007.

[GK07b] ———, *A deterministic approximation algorithm for computing a permanent of a 0,1 matrix*, Preprint on http://arxiv.org/abs/math.CO/0702039 (2007).

[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.

[Has99] J. Hastad, *Clique is hard to approximate within* $n^{1-\epsilon}$, Acta Math. **182** (1999), 105–142.

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, John Wiley and Sons, Inc., 2000.

[JS06] K. Jung and D. Shah, *On correctness of belief propagation algorithm*, Preprint (2006).

[KS81] R. Karp and M. Sipser, *Maximum matchings in sparse random graphs*, 22nd Annual Symposium on Foundations of Computer Science, 1981, pp. 364–375.

[MP03] M. Mezard and G. Parisi, *The cavity method at zero temperature*, Journal of Statistical Physics **111** (2003), no. 1-2, 1–34.

[RBMM04] O. Rivoire, G. Biroli, O. C. Martin, and M. Mezard, *Glass models on Bethe lattices*, Eur. Phys. J. B **37** (2004), 55–78.

[Wei06] D. Weitz, *Counting independent sets up to the tree threshold*, Proc. 38th Ann. Symposium on the Theory of Computing, 2006.