The Annals of Probability
2017, Vol. 45, No. 4, 2353–2376
DOI: 10.1214/16-AOP1114

LIMITS OF LOCAL ALGORITHMS OVER SPARSE RANDOM GRAPHS[1]

BY DAVID GAMARNIK[2] AND MADHU SUDAN

Massachusetts Institute of Technology and Harvard University

Local algorithms on graphs are algorithms that run in parallel on the nodes of a graph to compute some global structural feature of the graph. Such algorithms use only local information available at nodes to determine local aspects of the global structure, while also potentially using some randomness. Recent research has shown that such algorithms show significant promise in computing structures like large independent sets in graphs locally. Indeed the promise led to a conjecture by Hatami, Lovász and Szegedy [*Geom. Funct. Anal.* **24** (2014) 269–296] that local algorithms defined specifically as so-called i.i.d. factors may be able to find approximately largest independent sets in random d-regular graphs. In this paper, we refute this conjecture and show that every independent set produced by local algorithms is multiplicative factor $1/2+1/(2\sqrt{2})$ smaller than the largest, asymptotically as $d \to \infty$.

Our result is based on an important clustering phenomena predicted first in the literature on spin glasses, and recently proved rigorously for a variety of constraint satisfaction problems on random graphs. Such properties suggest that the geometry of the solution space can be quite intricate. The specific clustering property that we prove and apply in this paper shows that typically every two large independent sets in a random graph either have a significant intersection, or have a very small intersection. As a result, large independent sets are clustered according to the proximity to each other. While the clustering property was postulated earlier as an obstruction for the success of local algorithms, our result is the first one where the clustering property is used to formally prove limits on local algorithms.

1. Introduction. Local algorithms are decentralized algorithms that run in parallel on nodes of a graph using only information available from local neighborhoods to compute some global function of data that is spread over the network. Local algorithms have been studied in the past in various communities. They arise as natural solution concepts in distributed computing (see, e.g., [25]). They also lead to efficient sub-linear algorithms—algorithms that run in time significantly less than the length of the input—and [19, 31, 32, 34] illustrate some of the progress in

Received February 2014; revised July 2015.

[1]Conference version of the paper appeared in Proceedings of 5th ACM Innovations in Theoretical Computer Science (ITCS) conference.

[2]Supported by NSF Grant CMMI-1031332.

MSC2010 subject classifications. Primary 05C80, 60C05; secondary 82-08.

Key words and phrases. Random graphs, algorithms.

this direction. Finally, local algorithms have also been proposed as natural heuristics for solving hard optimization problems with the popular Belief Propagation algorithm (see, for instance, [28, 36]) being one such example.

In this work, we study the performance of a natural class of local algorithms on *random regular graphs* and show limits on the performance of these algorithms. The motivation for our work comes from the notion of local algorithms based the concept of i.i.d. factors that has appeared recently in the context of the theory of graph limits, developed in several papers, including [5–8, 12, 20, 26]. In the realms of this theory, it was conjectured that every "reasonable" combinatorial optimization problem defined on a *random graphs* can be solved by means of some local algorithms. This conjecture for the first time was formally stated in Hatami, Lovász and Szegedy in [20], and thus, from now on, we will refer to it as Hatami–Lovász–Szegedy (or HLS) conjecture, though informally it was posed by Szegedy earlier, and was referenced in several papers, including Lyons and Nazarov [27], and Csóka and Lippner [11]. In the concrete context of the problem of finding largest independent sets in sparse random regular graphs, the conjecture is stated as follows. Let $\mathbb{T}_{d,r}$ be a rooted d-regular tree with depth r. Namely, every node including the root, has degree r, except for the leaves, and the distance from the root to every leaf is r. Consider a function $f_r : [0,1]^{\mathbb{T}_{d,r}} \to \{0,1\}$ which maps every such tree whose nodes are decorated with real values from $[0,1]$ to a "decision" encoded by 0 and 1. In light of the fact that in a random d-regular graph $\mathbb{G}_d(n)$ on n nodes the typical node has depth-r neighborhood isomorphic to $\mathbb{T}_{d,r}$, for any constant r, such a function f_r can be used to generate (random) subsets I of $\mathbb{G}_d(n)$ as follows: decorate nodes of $\mathbb{G}_d(n)$ using i.i.d. uniform random values from $[0,1]$ and apply function f_r in every node (ignoring the nodes whose r-neighborhood is not isomorphic to $\mathbb{T}_{d,r}$). The set of nodes for which f_r produces value 1 defines I, and is called "i.i.d. factor." It is clear that f_r essentially describes a local algorithm for producing sets I (sweeping the issue of the computability of f_r under rug). The HLS conjecture postulates the existence of a sequence of $f_r, r = 1, 2, \ldots$, such that the set I thus produced is an independent subset of $\mathbb{G}_d(n)$ and asymptotically achieves the size of a largest independent set in $\mathbb{G}_d(n)$, as first $n \to \infty$ and then $r \to \infty$. Namely, largest independent subsets of random regular graphs are i.i.d. factors. The precise connection between this conjecture and the theory of graph limits is beyond the scope of this paper. Instead we refer the reader to the relevant papers [12, 20]. The concept of i.i.d. factors appears also in one of the open problem by David Aldous [2] in the context of coding invariant processes on infinite trees.

It turns out that an analogue of the HLS conjecture is indeed valid for another important combinatorial optimization problem—finding a largest matching of a graph. Lyons and Nazarov [27] established this for the case of bi-partite locally $\mathbb{T}_{d,r}$-tree-like graphs, and Csóka and Lippner established this result for general locally $\mathbb{T}_{d,r}$-tree-like graphs. Further, one can modify the framework of i.i.d. factors by encapsulating non-$\mathbb{T}_{d,r}$ type neighborhoods, for example, by making f_r

depend not only on the realization of random uniform in $[0, 1]$ values, but also on the realization of the graph-theoretic neighborhoods around the nodes. Some probabilistic bound on a degree might be needed to make this definition rigorous (though we will not attempt this formalization in this paper). In this case, one can consider, for example, i.i.d. factors when neighborhoods are distributed as r generations of a branching process with Poisson distribution, and then ask which combinatorial optimization problems defined now on sparse Erdős–Rényi graphs $\mathbb{G}(n, d/n)$ can be solved as i.i.d. factors. Here, $\mathbb{G}(n, d/n)$ is a random graph on n nodes with each of the $\binom{n}{2}$ edges selected with probability d/n, independently for all edges, and $d > 0$ is a fixed constant. In this case, it is possible to show that when $d \leq e$, the maximum independent set problem on $\mathbb{G}(n, d/n)$ can be solved nearly optimally by the well-known Belief Propagation (BP) algorithm with a bounded number of iterations. Since the BP algorithm with a bounded number of iterations is a local algorithm, then the maximum independent set on $\mathbb{G}(n, d/n)$ with $d \leq e$ is an i.i.d. factor, in the extended framework defined above. (We should note that the original proof of Karp and Sipser [22] of the very similar result, relied on a different method.) Thus, the framework of local algorithms viewed as i.i.d. factors is rich enough to solve several interesting combinatorial optimization problems.

Nevertheless, in this paper we refute the HLS conjecture in the context of maximum independent set problem on random regular graphs $\mathbb{G}_d(n)$. Specifically, we show that for large enough d, with high probability (w.h.p.) as $n \to \infty$, every independent set producible as an i.i.d. factor is a multiplicative factor $\gamma < 1$ smaller than a largest independent subset of $\mathbb{G}_d(n)$. We establish that γ is asymptotically at most $\frac{1}{2} + \frac{1}{2\sqrt{2}}$. In the earlier conference version of the paper [17], we have conjectured that the result should hold for $\gamma = 1/2$. The value $\gamma = 1/2$ is the tightest possible, since applying algorithms described in [24] and [16], i.i.d. factors can be used to generate independent sets which are half of the optimal value. Recently, our conjecture was confirmed by Rahman and Virag [33] using a method similar to the one employed in this work (more on this below).

Our result is based on a powerful, though fairly easy to establish in our context, *clustering* or equivalently *shattering* property of some combinatorial optimization problems defined on random graphs. This property was first conjectured in the theory of spin glasses and later confirmed by rigorous methods. For the first time, this clustering property was discussed in terms of the so-called overlap structure of the solutions of the Sherrington–Kirkpatrick model [35]. Later, it featured in the context of the random K-SAT problem and was proved rigorously by Achlioptas, Coja-Oghlan and Ricci-Tersenghi [1], and by Mezard, Mora and Zecchina [29], independently. We do not define the random K-SAT problem here and instead refer the reader to the aforementioned papers. What these results state is that in certain regimes, the set of satisfying assignments w.h.p. can be clustered into groups such that two solutions within the same cluster agree on a certain minimum number of variables, while two solutions from different clusters have to disagree on a certain

minimum number of variables. In particular, one can identify a certain nonempty interval $[z_1, z_2] \subset [0, 1]$, which we call *overlap gap*, such that for all $z \in [z_1, z_2]$ *no* two solutions of the random K-SAT problem agree on z fraction of variables. One can further show that the onset of clustering property occurs when the density of clauses to variables becomes at least $2^K/K$, while at the same time the formula remains satisfiable w.h.p. when the density is below approximately $2^K \log 2$. Intriguingly, the known algorithms for finding solutions of random instances of K-SAT problem also stop working around the $2^K/K$ threshold. It was widely conjectured that the onset of this clustering phase is the main obstruction for finding such algorithms. In fact, Coja-Oghlan [9] showed that the BP algorithm, which was earlier conjectured to be a good contender for solving the random instances of K-SAT problems, also fails when the density of clauses to variables is at least $2^K \log K/K$, though Coja-Oghlan's approach does not explicitly rely on the clustering property, and one could argue that the connection between the clustering property and the failure of the BP algorithm is coincidental. In fact, in the follow-up paper by the authors [18], using the approach similar to the one developed in this paper, it is shown that not only the BP algorithm but a more advanced version of the message passing algorithm called Survey Propagation (SP) algorithm with a bounded number of iterations fails to find solutions of a very related random Not-All-Equal K-SAT (NAE-K-SAT) model, which also exhibits clustering property. This is particularly interesting since the SP algorithm was put forward by statistical physicists precisely as a technique to deal the clustering issue in random constraint satisfaction problems [30].

Closer to the topic of this paper, the clustering property was also recently established for independent sets in Erdős–Rényi graphs. In particular Coja-Oghlan and Efthymiou [10] established the following result. It is known that the largest independent subset of $\mathbb{G}(n, d/n)$ has size approximately $(2\log d/d)n$, when d is large (see the next section for precise details). The authors of [10] show that the set of independent sets of size at least approximately $(\log d/d)n$ (namely those within factor $1/2$ of the optimal), are also clustered. Namely, one can split them into groups such that intersection of two independent sets within a group has a large cardinality, while intersection of two independent sets from different groups has a small cardinality. One should note that the known best algorithms for producing large independent subsets of random graphs also stop short factor $1/2$ of the optimal, both in the case of sparse and in the dense random graph cases, as exhibited by the well-known Karp's open problem regarding independent subsets of $\mathbb{G}(n, 1/2)$ [23].

The result in [9] is almost what we need for our analysis with two exceptions. First, we need to establish this clustering property for random regular as opposed Erdős–Rényi graphs. Second, the result in [10] applies to *typical* independent sets and does not rule out the possibility that there are two independent sets with some "intermediate" intersection cardinality, though the number of such pairs is insignificant compared to the total number of independent sets. For our result, we need to

show that, without exception, every pair of "large" independent sets has either large or small intersection. We indeed establish this, but at the cost of loosing additional factor $1/(2\sqrt{2})$. In particular, we show that for large enough d, w.h.p. as $n \to \infty$, every two independent subsets of $\mathbb{G}_d(n)$ with cardinality asymptotically $(1+\beta)(\log d/d)n$, where $1 \geq \beta > \frac{1}{2} + \frac{1}{2\sqrt{2}}$ either have intersection size at least $(1+z)(\log d/d)n$ or at most $(1-z)(\log d/d)n$, for some $z < \beta$. The result is established using a straightforward first moment argument: we compute the expected number of pairs of independent sets with intersection lying in the interval $[(1-z)(\log d/d)n, (1+z)(\log d/d)n]$, and show that this expectation converges to zero exponentially fast.

With this result at hand, the refutation of the HLS conjecture is fairly simple to derive. We prove that if local algorithms can construct independent sets of size asymptotically $(1+\beta)(\log d/d)n$, then, by means of a simple coupling construction, we can construct two independent sets with intersection size z for *all* z in the interval $[(1+\beta)^2(\log d/d)^2 n, (1+\beta)(\log d/d)n]$, clearly violating the clustering property. The additional factor $1/(2\sqrt{2})$ is an artifact of the analysis, and hence it is natural to expect that our result holds for all $\beta \in (0, 1]$, as was later confirmed in [33]. Namely, no local algorithm is capable of producing independent sets with size larger than factor $1/2$ of the optimal, asymptotically in d. We note again that this coincides with the barrier for all known algorithms. It is noteworthy that our result is the first one where algorithmic hardness derivation relies directly on the geometry of the solution space, vis-à-vis the clustering phenomena, and thus the connection between algorithmic hardness and clustering property is not coincidental.

The approach undertaken in [33] is very similar with one undertaken in the present paper except for one important difference: the authors consider the overlap structure of many and not just two independent sets. They show that the multiset of overlap values satisfies a certain set of constraints, and then show that if independent sets with size $\gamma(\log d/d)n, \gamma > 1$ can be produced by i.i.d. factors these constraints are violated.

It thus appears that the approach of using clustering property as refutation technique for existence of local algorithms is very fruitful as is already demonstrated in [18] and [33] along with the present paper.

The remainder of the paper is structured as follows. We introduce some basic material and the HLS conjecture in the next section. In the same section, we state our main theorem—nonvalidity of the conjecture (Theorem 2.5). We also state two secondary theorems, the first describing the overlap structure of independent sets in random graphs (Theorem 2.6)—the main tool in the proof of our result, and the second describing overlaps that can be found if local algorithms work well (Theorem 2.7). We prove our main theorem easily from the two secondary theorems in Section 3. We prove Theorem 2.7 in Section 4. Sections 5 and 6 are devoted to proofs of the theorem regarding the overlap property, for the case of Erdős–Rényi and random regular graph, respectively. While technically we do not need

such a result for the Erdős–Rényi graph, it is very simple to derive and provides the road map for the case of the regular graphs (where the calculations are a bit more tedious). The Erdős–Rényi case might also be useful for further studies of i.i.d. factors on Erdős–Rényi graphs as opposed to random regular graphs, in the framework described above.

2. Preliminaries and main result. For convenience, we repeat here some of the notions and definitions already introduced in the first section.

Basic graph terminology. All graphs in this paper are understood to be simple undirected graphs. Given a graph $\mathbb{G}$ with node set $V(\mathbb{G})$ and edge set $E(\mathbb{G})$, a subset of nodes $I \subset V(\mathbb{G})$ is an independent set if $(u, v) \notin E(\mathbb{G})$ for all $u, v \in I$. A path between nodes u and v with length r is a sequence of nodes $u_1, \ldots, u_{r-1}$ such that $(u, u_1), (u_1, u_2), \ldots, (u_{r-1}, v) \in E(\mathbb{G})$. The distance between nodes u and v is the length of the shortest path between them. For every positive integer value r and every node $u \in V(\mathbb{G})$, $B_{\mathbb{G}}(u, r)$ denotes the depth-r neighborhood of u in $\mathbb{G}$. Namely, $B_{\mathbb{G}}(u, r)$ is the subgraph of $\mathbb{G}$ induced by nodes v with distance at most r from u. When $\mathbb{G}$ is clear from context, we drop the subscript. The degree of a vertex $u \in V(\mathbb{G})$ is the number of vertices v such that $(u, v) \in E(\mathbb{G})$. The degree of a graph $\mathbb{G}$ is the maximum degree of a vertex of $\mathbb{G}$. A graph $\mathbb{G}$ is d-regular if the degree of every node is d.

Random graph preliminaries. Given a positive real d, $\mathbb{G}(n, d/n)$ denotes the Erdős–Rényi graph on n nodes $\{1, 2, \ldots, n\} \triangleq [n]$, with edge probability d/n. Namely each of the $\binom{n}{2}$ edges of a complete graph on n nodes belongs to $E(\mathbb{G}(n, d/n))$ with probability d/n, independently for all edges. Given a positive integer d, $\mathbb{G}_d(n)$ denotes a graph chosen uniformly at random from the space of all d-regular graphs on n nodes. This definition is meaningful only when nd is an even number, which we assume from now on. Given a positive integer m, let $\mathcal{I}(n, d, m)$ denote the set of all independent sets in $\mathbb{G}(n, d/n)$ with cardinality m. $\mathcal{I}_d(n, m)$ stands for a similar set for the case of random regular graphs. Given integers $0 \le k \le m$, let $\mathcal{O}(n, d, m, k)$ denote the set of pairs $I, J \in \mathcal{I}(n, d, m)$ such that $|I \cap J| = k$. The definition of the set $\mathcal{O}_d(n, m, k)$ is similar. The sizes of the sets $\mathcal{O}(n, d, m, k)$ and $\mathcal{O}_d(n, m, k)$, and in particular whether these sets are empty or not, is one of our focuses.

Denote by $\alpha(n, d)$ the size of a largest in cardinality independent subset of $\mathbb{G}(n, d/n)$, normalized by n. Namely,

$$\alpha(n, d) = n^{-1} \max\{m : \mathcal{I}(n, d, m) \neq \varnothing\}.$$

$\alpha_d(n)$ stands for the similar quantity for random regular graphs. It is known that $\alpha(n, d)$ and $\alpha_d(n)$ have deterministic limits as $n \to \infty$.

THEOREM 2.1. *For every $d \in \mathbb{R}_+$ there exists $\alpha(d)$ such that w.h.p. as $n \to \infty$,*

$$\alpha(n, d) \to \alpha(d). \tag{2.1}$$

Similarly, for every positive integer d there exists α_d such that w.h.p. as $n \to \infty$

$$\alpha_d(n) \to \alpha_d. \tag{2.2}$$

Furthermore,

$$\alpha(d) = \frac{2\log d}{d}\big(1 - o_d(1)\big), \tag{2.3}$$

$$\alpha_d = \frac{2\log d}{d}\big(1 - o_d(1)\big), \tag{2.4}$$

as $d \to \infty$.

From here onward, the standard notation o_d and O_d stand for asymptotics when $d \to \infty$. Specifically, for any nonnegative function $f(d)$, $o_d(f)$ stands for any function $g(d)$ satisfying $\lim_{d\to\infty} |g(d)|/f(d) = 0$ and $O_d(f)$ stands for any function $g(d)$ satisfying $\limsup_d |g(d)|/f(d) < \infty$. The same notation with subscripts stand for asymptotics when $n \to \infty$.

The convergence (2.1) and (2.2) was established in Bayati et al. [3] with some generalizations established in Gamarnik [15]. The limits (2.3) and (2.4) follow from much older results by Frieze [13] for the case of Erdős–Rényi graphs and by Frieze and Łuczak [14] for the case of random regular graphs, which established these limits in the $\limsup_n$ and $\liminf_n$ sense. The fallout of these results is that graphs $\mathbb{G}(n, d/n)$ and $\mathbb{G}_d(n)$ have independent sets of size up to approximately $(2\log d/d)n$, when n and d are large, namely in the doubly asymptotic sense when we first take n to infinity and then d to infinity.

Local graph terminology. A *decision function* is a measurable function $f = f(u, \mathbb{G}, \mathbf{x})$ which returns a Boolean value $\{0, 1\}$, where $\mathbb{G}$ is a graph on vertex set $[n]$ for some positive integer n, $u \in [n]$ is a vertex and $\mathbf{x} \in [0, 1]^N$ is a sequence of real numbers for some $N \geq n$. A decision function f is said to compute an independent set if for every graph $\mathbb{G}$ and every sequence $\mathbf{x}$ and for every pair $(u, v) \in E(\mathbb{G})$ it is the case that either $f(u, \mathbb{G}, \mathbf{x}) = 0$ or $f(v, \mathbb{G}, \mathbf{x}) = 0$, or both. We refer to such an f as an independence function. For an independence function f, graph $\mathbb{G}$ on vertex set $[n]$ and $\mathbf{x} \in [0, 1]^N$ for $N \geq n$, we let $I_{\mathbb{G}}(f, \mathbf{x})$ denote the independent set of $\mathbb{G}$ returned by f, i.e., $I_{\mathbb{G}}(f, \mathbf{x}) = \{u \in [n] \mid f(u, \mathbb{G}, \mathbf{x}) = 1\}$. We will assume later that $\mathbf{x}$ is chosen randomly according to some probability distribution. In this case $I_{\mathbb{G}}(f, \mathbf{x})$ is a randomly chosen independent set in $\mathbb{G}$.

We now define the notion of a "local" decision function, that is, one whose actions depend only on the local structure of a graph and the local randomness.

The definition is a natural one, but we formalize it below for completeness. Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be graphs on vertex sets $[n_1]$ and $[n_2]$, respectively. Let $u_1 \in [n_1]$ and $u_2 \in [n_2]$. We say that $\pi : [n_1] \to [n_2]$ is an r-local isomorphism mapping u_1 to u_2 if π is a graph isomorphism from $B_{\mathbb{G}_1}(u_1, r)$ to $B_{\mathbb{G}_2}(u_2, r)$ [so in particular it is a bijection from $B_{\mathbb{G}_1}(u_1, r)$ to $B_{\mathbb{G}_2}(u_2, r)$, and further it preserves adjacency within $B_{\mathbb{G}_1}(u_1, r)$ and $B_{\mathbb{G}_2}(u_2, r)$]. For $\mathbb{G}_1, \mathbb{G}_2, u_1, u_2$ and an r-local isomorphism π, we say sequences $x^{(1)} \in [0,1]^{N_1}$ and $x^{(2)} \in [0,1]^{N_2}$ are r-locally equivalent if for every $v \in B_{\mathbb{G}_1}(u_1, r)$ we have $x_v^{(1)} = x_{\pi(v)}^{(2)}$. Finally, we say $f(u, \mathbb{G}, x)$ is an r-local function if for every pair of graphs $\mathbb{G}_1, \mathbb{G}_2$, for every pair of vertices $u_1 \in V(\mathbb{G}_1)$ and $u_2 \in V(\mathbb{G}_2)$, for every r-local isomorphism π mapping u_1 to u_2 and r-locally equivalent sequences $x^{(1)}$ and $x^{(2)}$ we have $f(u_1, \mathbb{G}_1, x^{(1)}) = f(u_2, \mathbb{G}_2, x^{(2)})$. We often use the notation f_r to denote an r-local function.

Let $n_{d,r} \triangleq 1 + d \cdot ((d-1)^r - 1)/(d-2)$ denote the number of vertices in a rooted tree of degree d and depth r. We let $\mathbb{T}_{d,r}$ denote a canonical rooted tree on vertex set $[n_{d,r}]$ with root being 1. For $n \geq n_{d,r}$, $\mathbf{x} \in [0,1]^n$ and an r-local function f_r, we let $f_r(\mathbf{x})$ denote the quantity $f_r(1, \mathbb{T}_{d,r}, \mathbf{x})$. Let $\mathbf{X}$ be chosen according to a uniform distribution on $[0,1]^n$. The subset of nodes $I_{\mathbb{G}_d(n)}(f_r, \mathbf{X})$ is called the *i.i.d. factor* produced by the r-local function f_r. As we will see below, the $\alpha(f_r) \triangleq \frac{1}{n} \cdot \mathbb{E}_{\mathbf{X}}[f_r(\mathbf{X})]$ accurately captures [to within an additive $o(1)$ factor] the density of an independent returned by an r-local independence function f_r on $\mathbb{G}_d(n)$. From this point on, we write $I(f_r, \mathbf{X})$ in place of $I_{\mathbb{G}_d(n)}(f_r, \mathbf{X})$ for simplicity.

First, we recall the following folklore proposition which we will also use often in this paper.

PROPOSITION 2.2. *As $n \to \infty$, with probability tending to* 1 *almost all local neighborhoods in $\mathbb{G}_d(n)$ are isomorphic to a tree $\mathbb{T}_{d,r}$. Formally, for every d, r and ε, and for all sufficiently large n,*

$$\mathbb{P}_{\mathbb{G}_d(n)}\big(\big|\{u \in [n] \mid B_{\mathbb{G}_d(n)}(u, r) \text{ not isomorphic to } \mathbb{T}_{d,r}\}\big| \geq \varepsilon n\big) \leq \varepsilon.$$

This immediately implies that the expected value of the independent set $I(f_r, \mathbf{X})$ produced by f_r is $\alpha(f_r)n + o(n)$. In fact, the following concentration result holds.

PROPOSITION 2.3. *As $n \to \infty$, with probability tending to* 1 *the independent set produced by an r-local function f on $\mathbb{G}_d(n)$ is of size $\alpha(f)n + o(n)$. Formally, for every d, r, ε and every r-local function f, for sufficiently large n,*

$$\mathbb{P}_{\mathbb{G}_d(n), \mathbf{X} \in [0,1]^N}\big(\big|\big|I(f_r, \mathbf{X})\big| - \alpha(f_r)n\big| \geq \varepsilon n\big) \leq \varepsilon.$$

PROOF. The proof follows from the fact that the variance of $|I_{\mathbb{G}_d(n), \mathbf{X}}|$ is $O(n)$ and its expectation is $\alpha(f_r)n + o(n)$, and so the concentration follows by Chebyshev's inequality. The bound on the variance in turn follows from the fact that

for every graph $\mathbb{G}$, there are at most $O(n)$ pairs of vertices u and v for which the events $f(u, \mathbb{G}, \mathbf{X})$ and $f(v, \mathbb{G}, \mathbf{X})$ are not independent for random $\mathbf{X}$. Details omitted. $\square$

The Hatami–Lovász–Szegedy conjecture and our result. We now turn to describing the Hatami–Lovász–Szegedy (HLS) conjecture and our result. Recall α_d defined by (2.2). The HLS conjecture can be stated as follows.

CONJECTURE 2.4. *There exists a sequence of r-local independence functions $f_r, r \geq 1$ such that almost surely $I(f_r, n)$ is an independent set in $\mathbb{G}_d(n)$ and $\alpha(f_r) \to \alpha_d$ as $r \to \infty$.*

Namely, the conjecture asserts the existence of a local algorithm (r-local independence function f_r) which is capable of producing independent sets in $\mathbb{G}_d(n)$ of cardinality close to the largest that exist. For such an algorithm to be efficient, the function $f_r(u, \mathbb{G}, \mathbf{x})$ should also be efficiently computable *uniformly*. Even setting this issue aside, we show that there is a limit on the power of local algorithms to find large independent sets in $\mathbb{G}_d(n)$ and in particular the HLS conjecture does not hold. Let $\hat{\alpha}_d = \sup_r \sup_{f_r} \alpha(f_r)$, where the second supremum is taken over all r-local independence functions f_r.

THEOREM 2.5 (Main). *For every $\varepsilon > 0$ and all sufficiently large d,*

$$\frac{\hat{\alpha}_d}{\alpha_d} \leq \frac{1}{2} + \frac{1}{2\sqrt{2}} + \varepsilon.$$

That is, w.h.p. for every $\varepsilon > 0$ and for all sufficiently large d, a largest independent set obtainable by r-local functions is at most $(\frac{1}{2} + \frac{1}{2\sqrt{2}} + \varepsilon)\alpha_d n$ for all r.

Thus for all large enough d, there is a multiplicative gap between $\hat{\alpha}_d$ and the independence ratio α_d. That being said, our result does not rule out that for small d, $\hat{\alpha}_d$ in fact equals α_d, thus leaving the HLS conjecture open in this regime.

The two main ingredients in our proof of Theorem 2.5 both deal with the *overlaps* between independent sets in random regular graphs. Informally, our first result on the size of the overlaps shows that in random graphs the overlaps are not of "intermediate" size—this is formalized in Theorem 2.6. We then show that we can apply any r-local function f_r twice, with coupled randomness, to produce two independent sets of intermediate overlap where the size of the overlap depends on the size of the independent sets found by f_r and the level of coupling. This is formalized in Theorem 2.7. Theorem 2.5 follows immediately by combining the two theorems (and appropriate setting of parameters).

Overlaps in random graphs. We now state our main theorem about the overlap of large independent sets. We interpret the statement after we make the formal statement.

THEOREM 2.6. *For $\beta \in (1/\sqrt{2}, 1)$ and $0 < z < \sqrt{2\beta^2 - 1} < \beta$ and d, let $s = (1+\beta)d^{-1}\log d$ and let $K(z)$ denote the set of integers between $\frac{(1-z)n\log d}{d}$ and $\frac{(1+z)n\log d}{d}$. Then, for all large enough d, we have*

$$\lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{k\in K(z)} \mathcal{O}(n, d, \lfloor sn \rfloor, k) \neq \varnothing \Big) = 0, \tag{2.5}$$

and

$$\lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{k\in K(z)} \mathcal{O}_d(n, \lfloor sn \rfloor, k) \neq \varnothing \Big) = 0. \tag{2.6}$$

In other words, both in the Erdős–Rényi and in the random regular graph models, when $\beta > 1/\sqrt{2}$, and d is large enough, with probability approaching unity as $n \to \infty$, one cannot find a pair of independent sets I and J with size $\lfloor ns \rfloor$, such that their overlap (intersection) has cardinality at least $\frac{n(1-z)\log d}{d}$ and at most $\frac{n(1+z)\log d}{d}$.

Note that for all $\beta > 1/\sqrt{2}$, there exists z satisfying $0 < z < \sqrt{2\beta^2 - 1}$ and so the theorem is not vacuous in this setting. Furthermore, as $\beta \to 1$, z can be chosen arbitrarily close to 1 making the forbidden overlap region extremely broad. That is, as the size of the independent sets in consideration approaches the maximum possible (namely as $\beta \uparrow 1$), and as $d \to \infty$, we can take $z \to 1$. In other words, with probability approaching one, two nearly largest independent sets either overlap almost entirely or almost do not have an intersection. This is the key result for establishing our hardness bounds for existence of local algorithms.

A slightly different version of the first of these results can be found as Lemma 12 in [10]. It is shown in [10] that if an independent set I with size nearly $(1+\beta)n\log d/d$ is chosen uniformly at random from the set of all independent sets of this size, then w.h.p. (with respect to the graph randomness and the choice of I), there exists an empty overlap region in the sense described above with respect to *all* other independent sets J of the same cardinality. In fact, this empty overlap region exists for every $\beta \in (0, 1)$, as opposed to just $1 > \beta > 1/2 + 1/(2\sqrt{2})$ as in our case. Unfortunately, this result cannot be used for our purposes, since this result does not rule out the existence of rare sets I for which no empty overlap region exists.

Overlapping from local algorithms. Fix an r-local independence function f_r. Given a vector $\mathbf{X} = (X_u, 1 \le u \le n)$ of variables $X_u \in [0, 1]$, recall that

$I_{\mathbb{G}}(f_r, \mathbf{X})$ denotes the independent set of $\mathbb{G}$ given by $u \in I_{\mathbb{G}}(f_r, \mathbf{X})$ if and only if $f_r(u, \mathbb{G}, \mathbf{X}) = 1$.

Recall that $\mathbf{X}$ is chosen according to the uniform distribution on $[0, 1]^n$. Namely, X_u are independent and uniformly distributed over $[0, 1]$. In what follows, we consider some joint distributions on pairs of vectors $(\mathbf{X}, \mathbf{Y})$ such that marginal distributions on the vector $\mathbf{X}$ and $\mathbf{Y}$ are uniform on $[0, 1]^n$, though $\mathbf{X}$ and $\mathbf{Y}$ are dependent on each other. The intuition behind the proof of Theorem 2.5 is as follows. Note that if $\mathbf{X} = \mathbf{Y}$ then $I_{\mathbb{G}}(f_r, \mathbf{X}) = I_{\mathbb{G}}(f_r, \mathbf{Y})$. As a result the overlap $I_{\mathbb{G}}(f_r, \mathbf{X}) \cap I_{\mathbb{G}}(f_r, \mathbf{Y})$ between $I_{\mathbb{G}}(f_r, \mathbf{X})$ and $I_{\mathbb{G}}(f_r, \mathbf{Y})$ is $\alpha(f_r)n + o(n)$ in expectation. On the other hand, if $\mathbf{X}$ and $\mathbf{Y}$ are independent, then the overlap between $I_{\mathbb{G}}(f_r, \mathbf{X})$ and $I_{\mathbb{G}}(f_r, \mathbf{Y})$ is $\alpha^2(f_r)n + o(n)$ in expectation, since the decision to pick a vertex u in I is independent for most vertices when $\mathbf{X}$ and $\mathbf{Y}$ are independent. [In particular, note that if the local neighborhood around u is a tree, which according to Proposition 2.2 happens with probability approaching unity, then the two decisions are independent, and $u \in I$ with probability $\alpha(f_r)$.] Our main theorem shows that by coupling the variables, the overlap can be arranged to be of any intermediate size, to within an additive $o(n)$ factor. In particular, if $\alpha(f_r)$ exceeds $\frac{1}{2} + \frac{1}{2\sqrt{2}}$ we will be able to show that the overlap can be arranged to be between the values $\frac{(1-z)n \log d}{d}$ and $\frac{(1+z)n \log d}{d}$, described in Theorem 2.6 which contradicts the statement of this theorem.

THEOREM 2.7. *Fix a positive integer d. For constant r, let $f_r(u, \mathbb{G}, \mathbf{x})$ be an r-local independence function and let $\alpha = \alpha(f_r)$. For every $\gamma \in [\alpha^2, \alpha]$ and $\varepsilon > 0$, and for every sufficiently large n, there exists a distribution on variables $(\mathbf{X}, \mathbf{Y}) \in [0, 1]^n \times [0, 1]^n$ such that*

$$\mathbb{P}_{\mathbb{G}_d(n),(\mathbf{X},\mathbf{Y})}\big(\big|I_{\mathbb{G}_d(n)}(f_r, \mathbf{X}) \cap I_{\mathbb{G}_d(n)}(f_r, \mathbf{Y})\big| \notin \big[(\gamma - \varepsilon)n, (\gamma + \varepsilon)n\big]\big) \le \varepsilon.$$

3. Proof of Theorem 2.5. We now show how Theorems 2.6 and 2.7 immediately imply Theorem 2.5.

PROOF OF THEOREM 2.5. Fix an r-local function f_r and let $\alpha = \alpha(f_r)$. Fix an arbitrary $0 < \eta < 1$. We will prove below that for sufficiently large d we have

$$\frac{\alpha}{\alpha_d} \le \frac{1 + 1/\sqrt{2} + 2\eta}{2 - \eta}. \tag{3.1}$$

The theorem will then follow by making η sufficiently small.

Let $\varepsilon = \frac{\eta \log d}{d}$ and let $\beta = \frac{1}{\sqrt{2}} + \eta$. Elementary algebra implies that for this choice of β we have $0 < \eta < \sqrt{2\beta^2 - 1} < \beta < 1$ for sufficiently large β. By Proposition 2.3, we have that w.h.p. an independent set returned by f_r on $\mathbb{G}_d(n)$ is of size at least $(\alpha - \varepsilon)n$. Furthermore, for every $\gamma \in [\alpha^2, \alpha]$, we have, by Theorem 2.7, that w.h.p. $\mathbb{G}_d(n)$ has two independent sets I and J, with

$$|I|, |J| \ge (\alpha - \varepsilon)n \quad \text{and} \quad |I \cap J| \in \big[(\gamma - \varepsilon)n, (\gamma + \varepsilon)n\big]. \tag{3.2}$$

Finally, by Theorem 2.1, we have that for sufficiently large d, $|I|, |J| \le (2d^{-1}\log d)(1+\eta)n \le 4d^{-1}\log dn$ and so for sufficiently large d, $\alpha^2 \le d^{-1}\log d$, allowing us to set $\gamma = d^{-1}/\log d$.

Now we apply Theorem 2.6 with $z = \eta$ and the given choice of β. Theorem 2.6 asserts that w.h.p. $\mathbb{G}_d(n)$ has no independent sets of size $\lfloor (1+\beta)d^{-1}\log dn \rfloor$ with intersection size in $[(1-\eta)d^{-1}\log dn, (1+\eta)d^{-1}\log dn]$. This implies that there are no independent sets of size *at least* $(1+\beta)d^{-1}\log dn$ with intersection size in $[(1-\eta)d^{-1}\log dn, (1+\eta)d^{-1}\log dn]$, since we can simply delete the appropriate number of nodes outside of the intersection and since $\eta < \beta$.

Since $|I \cap J| \in [(\gamma-\varepsilon)n, (\gamma+\varepsilon)n] = [(1-\eta)d^{-1}\log dn, (1+\eta)d^{-1}\log dn]$, we conclude that $\min\{|I|, |J|\} \le (1+\beta)d^{-1}\log dn$. Combining with equation (3.2) we get that $(\alpha-\varepsilon)n \le \min\{|I|, |J|\} \le (1+\beta)d^{-1}\log dn$ and so $\alpha \le (1+\beta)d^{-1}\log d + \varepsilon$, which by the given bound on β yields

$$\alpha \le (1 + 1/\sqrt{2} + 2\eta)d^{-1}\log d.$$

On the other hand, by the second part of Theorem 2.1, we also have $\alpha_d \ge (2-\eta)d^{-1}\log d$ for all sufficiently large d. This implies (3.1). □

4. Proof of Theorem 2.7. For parameter $p \in [0,1]$, we define the p-correlated distribution on vectors of random variables $(\mathbf{X}, \mathbf{Y})$ to be the following: Let $\mathbf{X}, \mathbf{Z}$ be independent uniform vectors over $[0,1]^n$. Now let $Y_u = X_u$ with probability p and $Y_u = Z_u$ with probability $1-p$ independently for every $u \in V(G)$.

Let $f(u, \mathbb{G}, \mathbf{x})$ and α be as in the theorem statement. Recall that $f(\mathbf{x}) = f(1, \mathbb{T}_{d,r}, \mathbf{x})$ is the decision of f on the canonical tree of degree d and depth r rooted at the vertex 1. Let $\gamma(p)$ be the probability that $f(\mathbf{X}) = 1$ and $f(\mathbf{Y}) = 1$, for p-correlated variables $(\mathbf{X}, \mathbf{Y})$. As with Proposition 2.3, we have the following.

LEMMA 4.1. *For every d, r, $\varepsilon > 0$ and r-local function f, for sufficiently large n, we have*

$$\mathbb{P}_{\mathbb{G}_d(n),(\mathbf{X},\mathbf{Y})}\big(\big|\big|I_{\mathbb{G}_d(n)}(f,\mathbf{X}) \cap I_{\mathbb{G}_d(n)}(f,\mathbf{Y})\big| - \gamma(p)\cdot n\big| \ge \varepsilon n\big) \le \varepsilon,$$

where $(\mathbf{X}, \mathbf{Y})$ are p-correlated distributions on $[0,1]^n$.

PROOF. By Proposition 2.2, we have that almost surely almost all local neighborhoods are trees and so for most vertices u the probability that u is chosen to be in the independent sets $I(f, \mathbf{X})$ and $I(f, \mathbf{Y})$ is $\gamma(p)$. By linearity of expectations, we get that $\mathbb{E}[|I(f,\mathbf{X}) \cap I(f,\mathbf{Y})|] = \gamma(p)\cdot n + o(n)$. Again observing that most local neighborhoods are disjoint we have that the variance of $|I(f,\mathbf{X}) \cap I(f,\mathbf{Y})|$ is $O(n)$. We conclude, by applying the Chebyshev bound, that $|I(f,\mathbf{X}) \cap I(f,\mathbf{Y})|$ is concentrated around the expectation and the lemma follows. □

We also note that for $p = 1$ and $p = 0$ the quantity $\gamma(p)$ follow immediately from their definition.

PROPOSITION 4.2. *$\gamma(1) = \alpha$ and $\gamma(0) = \alpha^2$.*

Now to prove Theorem 2.7 it suffices to prove that for every $\gamma \in [\alpha^2, \alpha]$ there exists a p such that $\gamma(p) = \gamma$. We show this next by showing that $\gamma(p)$ is continuous.

LEMMA 4.3. *For every r, $\gamma(p)$ is a continuous function of p.*

PROOF. Let $(W_u, u \in \mathbb{T}_{d,r})$ be random variables associated with nodes in $\mathbb{T}_{d,r}$, uniformly distributed over $[0, 1]$, which are independent for different u and also independent from X_u and Z_u. We use W_u as generators for the events $Y_u = X_u$ versus $Y_u = Z_u$. In particular, given p, set $Y_u = X_u$ if $W_u \le p$ and $Y_u = Z_u$ otherwise. This process is exactly the process of setting variables Y_u to X_u and Z_u with probabilities p and $1 - p$, respectively, independently for all nodes u. Now fix any $p_1 < p_2$, and let $\delta < (p_2 - p_1)/d^{r+1}$. We use the notation $f_r(X_u, Z_u, W_u, p)$ to denote the value of f_r when the seed variables realization is $(W_u, u \in \mathbb{T}_{d,r})$, and the threshold value p is used. Namely, $f_r(X_u, Z_u, W_u, p) = f_r(X_u \mathbf{1}\{W_u \le p\} + Z_u \mathbf{1}\{W_u > p\}, u \in \mathbb{T}_{d,r})$. Here, for ease of notation, the reference to the tree $\mathbb{T}_{d,r}$ is dropped. Utilizing this notation, we have

$$\gamma(p) = \mathbb{P}\big(f_r(X_u) = f_r(X_u, Z_u, W_u, p) = 1\big).$$

Therefore,

$$\begin{aligned}\gamma(p_2) - \gamma(p_1) &= \mathbb{P}\big(f_r(X_u) = f_r(X_u, Z_u, W_u, p_2) = 1\big) \\ &\quad - \mathbb{P}\big(f_r(X_u) = f_r(X_u, Z_u, W_u, p_1) = 1\big) \\ &= \mathbb{E}\big[f_r(X_u) f_r(X_u, Z_u, W_u, p_2) - f_r(X_u) f_r(X_u, Z_u, W_u, p_1)\big].\end{aligned}$$

Observe that the event $W_u \notin [p_1, p_2]$ for all $u \in \mathbb{T}_{d,r}$ implies $f_r(X_u, Z_u, W_u, p_1) = f_r(X_u, Z_u, W_u, p_2)$ for every realization of X_u and Z_u. Therefore, by the union bound and since $|\mathbb{T}_{d,r}| < d^{r+1}$, we have

$$\big|\gamma(p_2) - \gamma(p_1)\big| \le d^{r+1}(p_2 - p_1).$$

Since r is fixed, the continuity of $\gamma(p)$ is established. □

We are now ready to prove Theorem 2.7.

PROOF OF THEOREM 2.7. Given $\gamma \in [\alpha^2, \alpha]$ by Lemma 4.3, we have that there exists a p such that $\gamma = \gamma(p)$. For this choice of p, let $(\mathbf{X}, \mathbf{Y})$ be a pair of p-correlated distributions. Applying Lemma 4.1 to this choice of p, we get that with probability at least $1 - \varepsilon$ we have $|I_{\mathbb{G}_d(n)}(f, \mathbf{X}) \cap I_{\mathbb{G}_d(n)}(f, \mathbf{Y})| \in [(\gamma - \varepsilon)n, (\gamma + \varepsilon)n]$ as desired. □

5. Theorem 2.6: Case of the Erdős–Rényi graph $\mathbb{G}(n, d/n)$. In this section, we prove Theorem 2.6 for the case of the random Erdős–Rényi graph.

The proof is based on a simple moment argument. We first determine the expected number of pairs of independent sets with a prescribed overlap size and show that this expectation converges to zero as $n \to \infty$ and in fact converges to zero exponentially fast when the overlap size falls into the corresponding interval. The result then follows from Markov inequality.

Fix positive integers $k \le m \le n$. Recall that $\mathcal{O}(n, d, m, k)$ is the set of all pairs of independent sets of cardinality m with intersection size k in the random graph $\mathbb{G}(n, d/n)$. It is straightforward to see that

$$\mathbb{E}\big[\big|\mathcal{O}(n, d, m, k)\big|\big] = \frac{n!}{k!(m-k)!(m-k)!(n-2m+k)!}\left(1-\frac{d}{n}\right)^{\binom{2m-k}{2}-(m-k)^2}. \tag{5.1}$$

Let $m = \lfloor ns \rfloor$, where we remind that $s = (1+\beta)d^{-1}\log d$ is given by the statement of the theorem. Set $k = \lfloor nx \rfloor$ for any

$$x \in \left(\frac{(1-z)\log d}{d}, \frac{(1+z)\log d}{d}\right). \tag{5.2}$$

It suffices to show that there exists $\gamma > 0$ such that

$$\limsup_{n\to\infty} n^{-1}\log \mathbb{E}\big[\big|\mathcal{O}\big(n, d, \lfloor ns \rfloor, \lfloor nx \rfloor\big)\big|\big] \le -\gamma, \tag{5.3}$$

for all x in the interval (5.2), as then we can use a union bound on the integer choices

$$k \in \left(n\frac{(1-z)\log d}{d}, n\frac{(1+z)\log d}{d}\right).$$

From this point on, we ignore $\lfloor\cdot\rfloor$ notation for the ease of exposition. It should be clear that this does not affect the argument. From (5.1), after simplifying using Stirling's approximation $[a! \approx (a/e)^a]$ and the fact that $\ln(1-y) \approx -y$ as $y \to 0$, we have

$$\begin{aligned}\limsup_n n^{-1}\log \mathbb{E}\big[\big|\mathcal{O}\big(n, d, \lfloor ns \rfloor, \lfloor nx \rfloor\big)\big|\big] &= x\log x^{-1} + 2(s-x)\log(s-x)^{-1} + (1-2s+x)\log(1-2s+x)^{-1} \\ &\quad - d\left(\frac{(2s-x)^2}{2} - (s-x)^2\right).\end{aligned} \tag{5.4}$$

We further simplify this expression as

$$\begin{aligned}&x\log x^{-1} + 2(s-x)\log(s-x)^{-1} \\ &\quad + (1-2s+x)\log(1-2s+x)^{-1} - ds^2 + dx^2/2.\end{aligned}$$

We have from (5.2) that for large enough d

$$x^{-1} \le d.$$

Also, for large enough d, since $z < \beta$, then

$$(s-x)^{-1} \le \left(\frac{(1+\beta)\log d}{d} - \frac{(1+z)\log d}{d}\right)^{-1} \le d.$$

Finally, we use the following asymptotics valid as $d \to \infty$:

$$(1-2s+x)\log(1-2s+x)^{-1} = O\left(\frac{\log d}{d}\right), \tag{5.5}$$

which applies since $0 \le x \le s = O_d(\log d/d)$. Substituting the expression for $s = (1+\beta)d^{-1}\log d$, we obtain a bound

$$n^{-1}\log \mathbb{E}\big[|\mathcal{O}(ns, nx)|\big] \le x\log d + 2\left(\frac{(1+\beta)\log d}{d} - x\right)\log d + O_d(\log d/d) - d\left(\frac{(1+\beta)\log d}{d}\right)^2 + dx^2/2.$$

Writing $x = (1+\hat{z})\log d/d$, where according to (5.2) $\hat{z}$ varies in the interval $[-z, z]$, we can conveniently rewrite our bound as

$$\frac{\log^2 d}{d}\big(2(1+\beta) - (1+\beta)^2 - (1+\hat{z}) + (1+\hat{z})^2/2\big) + O_d(\log d/d).$$

Now we can force the expression to be negative for large enough d, provided that

$$2(1+\beta) - (1+\beta)^2 - (1+\hat{z}) + (1+\hat{z})^2/2 < 0,$$

which is equivalent to $|\hat{z}| < \sqrt{2\beta^2 - 1}$ which in turn follows from the conditions on z in the hypothesis of the theorem statement.

This completes the proof of (2.5), and thus the proof of the theorem for the case of Erdős–Rényi graph.

6. Theorem 2.6: Case of the random regular graph $\mathbb{G}_d(n)$. We now turn to the case of random regular graphs $\mathbb{G}_d(n)$. We use a configuration model of $\mathbb{G}_d(n)$ [4, 21], which is obtained by replicating each of the n nodes of the graph d times, and then creating a random uniformly chosen matching connecting these dn nodes. Since nd is assumed to be even, such a matching exists. Then for every two nodes $u, v \in [n]$ an edge is created between u and v, if there exists at least one edge between any of the replicas of u and any of the replicas of v. This step of creating edges between nodes in $[n]$ from the matching on nd nodes we call projecting. It is known that, conditioned on the absence of loops and parallel edges, this gives a model of a random regular graph. It is also known that the probability of appearing of at least one loop or at least two parallel edges is bounded away from zero when d

is bounded. Since we are only concerned with statements taking place w.h.p., such a conditioning is irrelevant to us and thus we assume that $\mathbb{G}_d(n)$ is obtained simply by taking a random uniformly chosen matching and projecting. The configuration model is denoted by $\bar{\mathbb{G}}_d(n)$, with nodes denoted by (i,r) where $i=1,2,\dots,n$ and $r=1,\dots,d$. Namely, (i,r) is the rth replica of node i in the original graph. Given any set $A\subset[n]$, let $\bar{A}$ be the natural extension of A into the configuration model. Namely, $\bar{A}=\{(i,r): i\in A, r=1,\dots,d\}$.

Recall that $\mathcal{O}_d(n,m,k)$ stands for the set of pairs of independent sets I,J in $\mathbb{G}_d(n)$ such that $|I|=|J|=m$ and $|I\cap J|=k$. Note that there are possibly some edges between $\bar{I}\setminus\bar{J}$ and $\bar{J}\setminus\bar{I}$ resulting in edges between $I\setminus J$ and $J\setminus I$. Let $\mathcal{R}(m,k,l)\subset\mathcal{O}_d(n,m,k)$ be the set of pairs I,J such that the number of edges between $\bar{I}\setminus\bar{J}$ and $\bar{J}\setminus\bar{I}$ in the configuration graph model $\bar{\mathbb{G}}_d(n)$ is exactly l. Here, for the ease of notation we dropped the references to d and n. Observe that l is at most $d(m-k)$ and $\bigcup_{l=0}^{d(m-k)}\mathcal{R}(m,k,l)=\mathcal{O}_d(n,m,k)$. In what follows, we will bound the expected size of $\mathcal{R}(m,k,l)$ for every l, and thus the expected size of their union.

For $(I,J)\in\mathcal{R}(m,k,l)$, the number of edges between the set $I\cup J$ and its complement $[n]\setminus(I\cup J)$ is precisely $(2m-k)d-2l$, since $|I\cup J|=2m-k$. The same applies to the configuration model: the number of edges between $\bar{I}\cup\bar{J}$ and its complement $[nd]\setminus(\bar{I}\cup\bar{J})$ is precisely $(2m-k)d-2l$. The value of $\mathbb{E}[|\mathcal{R}(m,k,l)|]$ is then computed as follows. Let $R=2m-k$ and $l\le d(m-k)$.

LEMMA 6.1.

$$\mathbb{E}|\mathcal{R}(m,k,l)| = \binom{n}{k,m-k,m-k,n-R}\binom{md-kd}{l}^2\binom{nd-Rd}{Rd-2l}l!(Rd-2l)! \times\frac{(nd-2Rd+2l)!}{(nd/2-Rd+l)!2^{nd/2-Rd+l}}\frac{(nd/2)!2^{\frac{nd}{2}}}{(nd)!}.$$

PROOF. The proof is based on the fact that the number of matchings on a set of m nodes (for even m) is $\frac{m!}{(m/2)!2^{\frac{m}{2}}}$. So the term $\frac{(nd/2)!2^{\frac{nd}{2}}}{(nd)!}$ is precisely the inverse of the number of configuration graphs $\bar{\mathbb{G}}_d(n)$. The term $\binom{n}{k,m-k,m-k,n-R}$ is the number of ways of selecting a pair of sets I and J with cardinality m each and intersection size k. Finally, the remaining term

$$\binom{md-kd}{l}^2\binom{nd-Rd}{Rd-2l}l!(Rd-2l)!\frac{(nd-2Rd+2l)!}{(nd/2-Rd+l)!2^{nd/2-Rd+l}}$$

is the number of graphs $\mathbb{G}_d(n)$ such that for a given choice of sets I and J, both sets are independent sets, and the number of edges between $I\setminus J$ and $J\setminus I$ is l. Here, $\binom{md-kd}{l}^2$ represents the number of choices for end points of the l edges between

$I \setminus J$ and $J \setminus I$; $l!$ represents the number of matchings once these choices are made; $\binom{nd-Rd}{Rd-2l}$ represents the number of choices for the end points of edges connecting $I \cup J$ with its complement; $(Rd-2l)!$ represents the number of matchings once these choices are made; and finally

$$\frac{(nd-2Rd+2l)!}{(nd/2-Rd+l)!2^{nd/2-Rd+l}}$$

represents the number of matching choices between the remaining $nd-2Rd+2l$ nodes in the complement of $\bar{I} \cup \bar{J}$. $\square$

We write $k = xn, m = sn, l = dyn$, where $x \le s \le 1$. Then $R = (2s-x)n$ and $y \le s-x$. Our main goal is establishing the following analogue of (5.3).

LEMMA 6.2. *There exists $\gamma > 0$ such that*

$$\limsup_{n\to\infty} n^{-1} \log \mathbb{E}\big[\big|\mathcal{R}\big(\lfloor ns\rfloor, \lfloor nx\rfloor, \lfloor ny\rfloor\big)\big|\big] \le -\gamma, \tag{6.1}$$

for $s = (1+\beta)d^{-1}\log d$, for all x in the interval (5.2) and all $0 \le y \le s-x$.

The claim (2.6) of Theorem 2.5 follows from Lemma 6.2 by an argument similar to the one for the Erdős–Rényi graph. The rest of this section is devoted to proving Lemma 6.2.

Next, we simplify the last two terms in the expression given in Lemma 6.1 using Stirling's approximtion. We have by this lemma

$$\begin{aligned}
&\mathbb{E}\big[\big|\mathcal{R}(m,k,l)\big|\big] \\
&\quad= \binom{n}{k, m-k, m-k, n-R}\binom{md-kd}{l}^2\binom{nd-Rd}{Rd-2l} l!(Rd-2l)!\big(1+o(1)\big) \\
&\qquad\times \frac{(nd-2Rd+2l)^{\frac{(nd-2Rd+2l)}{2}}}{e^{\frac{(nd-2Rd+2l)}{2}}}\frac{e^{\frac{nd}{2}}}{(nd)^{\frac{nd}{2}}}\big(1+o(1)\big) \\
&\quad= \frac{n!}{k!((m-k)!)^2(n-R)!}\frac{((md-kd)!)^2}{(l!)^2((md-kd-l)!)^2}\frac{(nd-Rd)!}{(Rd-2l)!(nd-2Rd+2l)!} \\
&\qquad\times l!(Rd-2l)!(nd-2Rd+2l)^{\frac{(nd-2Rd+2l)}{2}}e^{Rd-l}(nd)^{-\frac{nd}{2}}\big(1+o(1)\big) \\
&\quad= \frac{n!}{k!((m-k)!)^2(n-R)!}\frac{((md-kd)!)^2}{l!((md-kd-l)!)^2}\frac{(nd-Rd)!}{(nd-2Rd+2l)!} \\
&\qquad\times (nd-2Rd+2l)^{\frac{(nd-2Rd+2l)}{2}}e^{Rd-l}(nd)^{-\frac{nd}{2}}\big(1+o(1)\big).
\end{aligned}$$

Next, we consider the logarithm of the expression above normalized by n and apply Stirling's approximation to the remaining terms. We obtain

$$\begin{aligned}
&n^{-1}\log\mathbb{E}\big[\big|\mathcal{R}(m,k,l)\big|\big]\\
&\quad=-x\log x-2(s-x)\log(s-x)-(1-2s+x)\log(1-2s+x)\\
&\qquad+2(sd-xd)\log(sd-xd)-2(sd-xd)-dy\log dy+dy\\
&\qquad-2(sd-xd-dy)\log(sd-xd-dy)+2(sd-xd-dy)\\
&\qquad+(d-2ds+dx)\log(d-2ds+dx)-(d-2ds+dx)\\
&\qquad-(d-4ds+2dx+2dy)\log(d-4ds+2dx+2dy)\\
&\qquad+(d-4ds+2dx+2dy)\\
&\qquad+\frac{1}{2}(d-4ds+2dx+2dy)\log(d-4ds+2dx+2y)\\
&\qquad+d(2s-x-y)-\frac{d}{2}\log d.
\end{aligned}$$

Grouping the terms not involving log, we see that they sum up to zero:

$$\begin{aligned}
&-2(sd-xd)+dy+2(sd-xd-dy)-(d-2ds+dx)\\
&\quad+(d-4ds+2dx+2dy)+d(2s-x-y)\\
&\qquad=0.
\end{aligned}$$

Consider terms of the form $\log(dA)=\log d+\log A$ and consider the multiplier corresponding to the $\log d$ term:

$$\begin{aligned}
&2(sd-xd)-dy-2(sd-xd-dy)+(d-2ds+dx)\\
&\qquad-\frac{1}{2}(d-4ds+2dx+2dy)-\frac{d}{2},
\end{aligned}$$

which again is found to be zero. The final expression we obtain is then

$$\begin{aligned}
&=-x\log x-2(s-x)\log(s-x)-(1-2s+x)\log(1-2s+x)\\
&\quad+2d(s-x)\log(s-x)-dy\log y\\
&\quad-2d(s-x-y)\log(s-x-y)\\
&\quad+d(1-2s+x)\log(1-2s+x)\\
&\quad-\frac{d}{2}(1-4s+2x+2y)\log(1-4s+2x+2y).
\end{aligned}\tag{6.2}$$

We claim that this expression is asymptotically upper bounded by the value similar to the one given in the right-hand side of (5.4) for the case of Erdős–Rényi graph.

Specifically, we claim that it is upper bounded by

$$-x\log x - 2(s-x)\log(s-x) - \frac{d}{2}(2s-x)^2 + d(s-x)^2 + o_d\big(\log^2 d/d\big). \tag{6.3}$$

The remainder of the proof is then the same as for the Erdős–Rényi case.

The rest of the section is devoted to deriving the upper bound (6.3). We recall that $s = (1+\beta)\log d/d$ and x lies in the interval (5.2). For convenience, we start with the term $(1-2s+x)\log(1-2s+x)$ in (6.2). Using the first-order Taylor approximation $\log(1-t) = -t + o(t)$, and the fact $s = O_d(\log d/d), x = O_d(\log d/d)$, we have

$$\begin{aligned}(1-2s+x)\log(1-2s+x) &= O_d(\log d/d)\\ &= o_d\big(\log^2 d/d\big).\end{aligned}$$

Our next goal is to maximize the expression in (6.2) with respect to y. We consider two cases. Specifically, we first consider the case

$$(\beta+z+1)^2\frac{\log^2 d}{d^2} \le y \le s-x, \tag{6.4}$$

and then consider the case

$$0 \le y \le (\beta+z+1)^2\frac{\log^2 d}{d^2}. \tag{6.5}$$

Assume first that (6.4) holds. Consider the terms containing y:

$$\begin{aligned}f(y) \triangleq{}& -dy\log y - 2d(s-x-y)\log(s-x-y)\\ &- \frac{d}{2}(1-4s+2x+2y)\log(1-4s+2x+2y).\end{aligned}$$

We claim that for all sufficiently large d, the derivative of f is negative in the range described by (6.4), and thus the largest value is obtained when y equals the left end of the range and, therefore, we can restrict ourselves to the second case (6.5). To establish the claim, we have

$$\begin{aligned}d^{-1}\dot f(y) &= -\log y - 1 + 2\log(s-x-y) + 2 - \log(1-4s+2x+2y) - 1\\ &= -\log y + 2\log(s-x-y) - \log(1-4s+2x+2y).\end{aligned}$$

Now by our assumption (6.4), we have $y \ge (\beta+z+1)^2 d^{-2}\log^2 d$ implying

$$-\log y \le -2\log(\beta+z+1) + 2\log d - 2\log\log d.$$

Also $4s - 2x - 2y \le 4s < 8\log d/d = O_d(\log d/d)$, implying that $\log(1-4s+2x+2y) = O_d(\log d/d)$. Finally, from (5.2) we have

$$\begin{aligned}s-x-y &\le s-x\\ &\le (\beta+z)\log d/d,\end{aligned}$$

implying that $\log(s-x-y) \le -\log d + \log\log d + \log(\beta+z)$. Combining, we obtain that

$$\begin{aligned} d^{-1}\dot{f}(y) &\le -2\log(\beta+z+1) + 2\log d - 2\log\log d - 2\log d + 2\log\log d \\ &\quad + 2\log(\beta+z) + O_d(\log d/d) \\ &= -2\log(\beta+z+1) + 2\log(\beta+z) + O_d(\log d/d). \end{aligned}$$

In particular, the derivative is negative for large enough d as claimed.

Thus, we now assume that the bound (6.5) holds. Next, we obtain asymptotic upper bounds on terms in (6.2). We begin with the term $d(1-2s+x)\log(1-2s+x)$. Using the approximation

$$(1-t)\log(1-t) = -t + t^2/2 + O(t^3),$$

we obtain

$$d(1-2s+x)\log(1-2s+x) = -d(2s-x) + \frac{d}{2}(2s-x)^2 + O_d\bigl(d(2s-x)^3\bigr).$$

Before we expand this term in terms of d, it will be convenient to obtain a similar expansion for the last term in (6.2)

$$\begin{aligned} &\frac{d}{2}(1-4s+2x+2y)\log(1-4s+2x+2y) \\ &\quad = -\frac{d}{2}(4s-2x-2y) + \frac{d}{4}(4s-2x-2y)^2 + O_d\bigl(d(4s+2x+2y)^3\bigr) \\ &\quad = -d(2s-x) + dy + d(2s-x)^2 - 2d(2s-x)y + dy^2 \\ &\qquad + O_d\bigl(d(4s+2x+2y)^3\bigr). \end{aligned}$$

Applying the upper bound (6.5), we have $O_d(d(2s-x)^3) = O_d(\log^3 d/d^2) = o_d(\log^2 d/d)$, $O_d(d(4s+2x+2y)^3) = o_d(\log^2 d/d)$, and $dy^2 = O_d(\log^4 d/d^3) = o_d(\log^2 d/d)$. Combining, we obtain

$$\begin{aligned} &d(1-2s+x)\log(1-2s+x) - \frac{d}{2}(1-4s+2x+2y)\log(1-4s+2x+2y) \\ &\quad = -\frac{d}{2}(2s-x)^2 - dy + 2d(2s-x)y + o_d\bigl(\log^2 d/d\bigr) \qquad (6.6) \\ &\quad = -\frac{d}{2}(2s-x)^2 - dy + o_d\bigl(\log^2 d/d\bigr), \end{aligned}$$

where again applying bound (6.5) on y we have used

$$2d(2s-x)y = O\left(d\frac{\log d}{d}\frac{\log^2 d}{d^2}\right) = o_d\bigl(\log^2 d/d\bigr).$$

We conclude that the last two terms in (6.2) sum to at most

$$-\frac{d}{2}(2s-x)^2 - dy + o_d\big(\log^2 d/d\big). \tag{6.7}$$

Next, it is convenient to analyze the following two terms of (6.2) together:

$$\begin{aligned}
&2d(s-x)\log(s-x) - 2d(s-x-y)\log(s-x-y)\\
&\quad= 2d(s-x)\log(s-x) - 2d(s-x)\log(s-x-y) + 2dy\log(s-x-y)\\
&\quad= 2d(s-x)\log(s-x) - 2d(s-x)\log(s-x)\\
&\qquad - 2d(s-x)\log\big(1-y(s-x)^{-1}\big)\\
&\qquad + 2dy\log(s-x) - 2dy\log\big(1-y(s-x)^{-1}\big)\\
&\quad= -2d(s-x)\log\big(1-y(s-x)^{-1}\big) + 2dy\log(s-x)\\
&\qquad - 2dy\log\big(1-y(s-x)^{-1}\big)\\
&\quad= 2d(s-x)y(s-x)^{-1} + O_d\big(dy^2(s-x)^{-1}\big)\\
&\qquad + 2dy\log(s-x) + 2dy^2(s-x)^{-1} + O_d\big(dy^3(s-x)^{-2}\big)\\
&\quad= 2dy + 2dy\log(s-x) + 2dy^2(s-x)^{-1} + O_d\big(dy^2(s-x)^{-1}\big)\\
&\qquad + O_d\big(dy^3(s-x)^{-2}\big)\\
&\quad= 2dy + 2dy\log(s-x) + o_d\big(\log^2 d/d\big),
\end{aligned}$$

where in the last step we have used the asymptotics $y = O_d(\log^2 d/d^2)$ implied by (6.5) to obtain $2dy^2(s-x)^{-1} = o_d(\log^2 d/d)$.

We now analyze the remaining terms in (6.2) involving y and optimize with respect to y. From (6.7), we have the term $-dy$. Combining with the asymptotics above and the remaining term $-dy\log y$ from (6.2), we obtain

$$\begin{aligned}
&2dy + 2dy\log(s-x) - dy - dy\log y\\
&\quad= dy + 2dy\log(s-x) - dy\log y.
\end{aligned} \tag{6.8}$$

We compute the maximum value of this quantity in the relevant range of y given by (6.5). The first derivative of this expression is

$$d + 2d\log(s-x) - d - d\log y = 2d\log(s-x) - d\log y$$

which is positive (infinite) at $y = 0$. At $y = (\beta+z+1)^2\log^2 d/d^2$, the first derivative is

$$\begin{aligned}
&2d\log(s-x) - 2d\log(\beta+z+1) - 2d\log\log d + 2d\log d\\
&\quad\le 2d\log(\beta+z) + 2d\log\log d - 2d\log d - 2d\log(\beta+z+1)
\end{aligned}$$

$$\begin{aligned} &- 2d\log\log d + 2d\log d \\ &= 2d\log(\beta + z) - 2d\log(\beta + z + 1) \\ &< 0, \end{aligned}$$

where the inequality relies on $x \geq (1-z)\log d/d$ implied by (5.2), which gives

$$s - x \leq (\beta + z)\log d/d.$$

The second derivative is $-d/y$ which is negative since $y \geq 0$. Thus, the function is strictly concave with positive and negative derivatives at the ends of the relevant interval (6.5). The maximum is then achieved at the unique point y^* where the derivative is zero, namely when $2d\log(s-x) - d\log y^* = 0$, giving

$$y^* = (s-x)^2.$$

Plugging this into the right-hand side of (6.8), we obtain

$$\begin{aligned} &dy^* + 2dy^*\log(s-x) - dy^*\log y^* \\ &\quad = d(s-x)^2 + 2d(s-x)^2\log(s-x) - d(s-x)^2\log(s-x)^2 \\ &\quad = d(s-x)^2. \end{aligned}$$

Summarizing, and using (6.7), we find that the expression in (6.2) is at most

$$= -x\log x - 2(s-x)\log(s-x) - \frac{d}{2}(2s-x)^2 + d(s-x)^2 + o_d\big(\log^2 d/d\big),$$

namely we obtain (6.3) as claimed. This completes the proof of Lemma 6.2 and of Theorem 2.6.

Acknowledgments. Several insightful conversations with Gabor Elek, Russel Lyons, Andre Csóka, Laci Lovász and Balázs Szegedy are gratefully acknowledged. The first author wishes to thank Microsoft Research New England for providing exciting and hospitable environment during his visit in 2012 when this work was conducted.

REFERENCES

[1] ACHLIOPTAS, D., COJA-OGHLAN, A. and RICCI-TERSENGHI, F. (2011). On the solution-space geometry of random constraint satisfaction problems. *Random Structures Algorithms* **38** 251–268. MR2663730

[2] ALDOUS, D. Some open problems. Available at http://www.stat.berkeley.edu/~aldous/Research/OP/index.html.

[3] BAYATI, M., GAMARNIK, D. and TETALI, P. (2013). Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. *Ann. Probab.* **41** 4080–4115. MR3161470

[4] BOLLOBÁS, B. (1985). *Random Graphs*. Academic Press, London. MR0809996

[5] BORGS, C., CHAYES, J. and GAMARNIK, D. (2013). Convergent sequences of sparse graphs: A large deviations approach. Preprint. Available at arXiv:1302.4615.

[6] BORGS, C., CHAYES, J., KAHN, J. and LOVÁSZ, L. (2013). Left and right convergence of graphs with bounded degree. *Random Structures Algorithms* **42** 1–28. MR2999210

[7] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626

[8] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382

[9] COJA-OGHLAN, A. (2011). On belief propagation guided decimation for random k-SAT. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms* 957–966. SIAM, Philadelphia, PA. MR2857177

[10] COJA-OGHLAN, A. and EFTHYMIOU, C. (2011). On independent sets in random graphs. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms* 136–144. SIAM, Philadelphia, PA. MR2857116

[11] CSOKA, E. and LIPPNER, G. (2012). Invariant random matchings in Cayley graphs. Preprint. Available at arXiv:1211.2374.

[12] ELEK, G. and LIPPNER, G. (2010). Borel oracles. An analytical approach to constant-time algorithms. *Proc. Amer. Math. Soc.* **138** 2939–2947. MR2644905

[13] FRIEZE, A. M. (1990). On the independence number of random graphs. *Discrete Math.* **81** 171–175. MR1054975

[14] FRIEZE, A. M. and ŁUCZAK, T. (1992). On the independence and chromatic numbers of random regular graphs. *J. Combin. Theory Ser. B* **54** 123–132. MR1142268

[15] GAMARNIK, D. (2014). Right-convergence of sparse random graphs. *Probab. Theory Related Fields* **160** 253–278. MR3256814

[16] GAMARNIK, D. and GOLDBERG, D. A. (2010). Randomized greedy algorithms for independent sets and matchings in regular graphs: Exact results and finite girth corrections. *Combin. Probab. Comput.* **19** 61–85. MR2575098

[17] GAMARNIK, D. and SUDAN, M. (2014). Limits of local algorithms over sparse random graphs [extended abstract]. In *ITCS'*14*—Proceedings of the* 2014 *Conference on Innovations in Theoretical Computer Science* 369–375. ACM, New York. MR3359490

[18] GAMARNIK, D. and SUDAN, M. (2014). Performance of the survey propagation-guided decimation algorithm for the random NAE-K-SAT problem. Preprint. Available at arXiv:1402.0052.

[19] HASSIDIM, A., KELNER, J. A., NGUYEN, H. N. and ONAK, K. (2009). Local graph partitions for approximation and testing. In 2009 50*th Annual IEEE Symposium on Foundations of Computer Science* (*FOCS* 2009) 22–31. IEEE Computer Soc., Los Alamitos, CA. MR2648385

[20] HATAMI, H., LOVÁSZ, L. and SZEGEDY, B. (2014). Limits of locally-globally convergent graph sequences. *Geom. Funct. Anal.* **24** 269–296. MR3177383

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

[22] KARP, R. and SIPSER, M. (1981). Maximum matchings in sparse random graphs. In 22*nd Annual Symposium on Foundations of Computer Science* 364–375. Nashville, TN.

[23] KARP, R. M. (1976). The probabilistic analysis of some combinatorial search algorithms. In *Algorithms and Complexity* (*Proc. Sympos., Carnegie-Mellon Univ., Pittsburgh, Pa.*, 1976) 1–19. Academic Press, New York. MR0445898

[24] LAUER, J. and WORMALD, N. (2007). Large independent sets in regular graphs of large girth. *J. Combin. Theory Ser. B* **97** 999–1009. MR2354714

[25] LINIAL, N. (1992). Locality in distributed graph algorithms. *SIAM J. Comput.* **21** 193–201. MR1148825

[26] LOVÁSZ, L. and SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085

[27] LYONS, R. and NAZAROV, F. (2011). Perfect matchings as IID factors on non-amenable groups. *European J. Combin.* **32** 1115–1125. MR2825538

[28] MÉZARD, M. and MONTANARI, A. (2009). *Information, Physics, and Computation*. Oxford Univ. Press, Oxford. MR2518205

[29] MÉZARD, M., MORA, T. and ZECCHINA, R. (2005). Clustering of solutions in the random satisfiability problem. *Phys. Rev. Lett.* **94** 197205.

[30] MÉZARD, M., PARISI, G. and ZECCHINA, R. (2002). Analytic and algorithmic solution of random satisfiability problems. *Science* **297** 812–815.

[31] NGUYEN, H. N. and ONAK, K. (2008). Constant-time approximation algorithms via local improvements. In *FOCS* 327–336. IEEE Computer Society, Los Alamitos, CA.

[32] PARNAS, M. and RON, D. (2007). Approximating the minimum vertex cover in sublinear time and a connection to distributed algorithms. *Theoret. Comput. Sci.* **381** 183–196. MR2347402

[33] RAHMAN, M. and VIRAG, B. (2014). Local algorithms for independent sets are half-optimal. Preprint. Available at arXiv:1402.0485.

[34] RUBINFELD, R., TAMIR, G., VARDI, S. and XIE, N. (2011). Fast local computation algorithms. In *ICS* (B. Chazelle, ed.) 223–238. Tsinghua Univ. Press, Beijing, China.

[35] TALAGRAND, M. (2011). *Mean Field Models for Spin Glasses. Volume I: Basic Examples*. Springer, Berlin. MR2731561

[36] WAINWRIGHT, M. J. and JORDAN, M. I. (2005). *A Variational Principle for Graphical Models*. MIT Press, Cambridge, MA.

SLOAN SCHOOL OF MANAGEMENT, E62-563
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
100 MAIN STREET
CAMBRIDGE, MASSACHUSETTS 02140
USA
E-MAIL: gamarnik@mit.edu

SCHOOL OF ENGINEERING AND APPLIED SCIENCES
HARVARD UNIVERSITY
MAXWELL DWORKIN 339
33 OXFORD STREET
CAMBRIDGE, MASSACHUSETTS 02138
USA
E-MAIL: madhu@cs.harvard.edu