Derandomized Graph Products

Noga Alon * Uriel Feige† Avi Wigderson ‡ David Zuckerman§

Abstract

Berman and Schnitger [10] gave a randomized reduction from approximating MAX-SNP problems [24] within constant factors arbitrarily close to 1 to approximating clique within a factor of n^ϵ (for some ϵ). This reduction was further studied by Blum [11], who gave it the name *randomized graph products*. We show that this reduction can be made deterministic (derandomized), using random walks on expander graphs [1]. The main technical contribution of this paper is in *lower bounding* the probability that all steps of a random walk stay within a specified set of vertices of a graph. (Previous work was mainly concerned with *upper bounding* this probability.) This lower bound extends also to the case that different sets of vertices are specified for different time steps of the walk.

1 Introduction

We present lower bounds on the probability that all steps of a random walk stay within a specified set of vertices of a graph. We then apply these lower bounds to amplify unapproximability results about certain NP-hard optimization problems. Our work was motivated by the problem of approximating the size of the maximum clique in graphs. This motivating problem, which serves also as an example of how our lower bounds can be applied, is described in section 1.1.

The constructions of this paper can also be used in order to amplify other unapproximability results. In Section 3 we use them in showing that it is NP-hard to approximate

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel and AT & T Bell Labs, Murray Hill, NJ, 07974, USA. Email: noga@math.tau.ac.il. Research supported in part by a USA-Israeli BSF grant.

†Department of Applied Math and computer Science, the Weizmann Institute, Rehovot, Israel. feige@wisdom.weizmann.ac.il. Part of this work was done while visiting IBM T.J. Watson Research Center at Yorktown Heights.

‡Department of Computer Science, Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel.

§MIT Lab for Computer Science, 545 Technology Square, Cambridge, MA 02139, diz@theory.lcs.mit.edu. This research was supported by an NSF Postdoctoral Fellowship.

the size of the maximum independent set within a factor of Δ^ϵ, in graphs of degree at most Δ. In [4], and in the final version of [5], they are used in showing that it is NP-hard to approximate to within a factor of n^ϵ the maximum number of simultaneously satisfiable equations in a system of n linear equations over the rationals, and that it is NP-hard to approximate to within a factor of q^ϵ the maximum number of simultaneously satisfiable equations in a system of linear equations over $GF(q)$.

1.1 The motivating example

The problem of computing the size of the maximum clique in a graph was one of the first problems shown to be NP-complete [19].

Definition 1 *Let G be a graph on n vertices. A* clique *is a set of vertices, any two of which are connected by an edge. The size of the maximum clique in G is denoted by $\omega(G)$.*

It is of interest to see how well $\omega(G)$ can be approximated in polynomial time.

Definition 2 *Let $g(n)$ and $f(n)$ be functions. We say that $g(n)$* approximates *$f(n)$ (from below) within a factor of $c(n) > 1$ if, for every n, $f(n)/c(n) \leq g(n) \leq f(n)$.*

Definition 3 *Algorithm A* approximates clique within a factor of $c(n)$ *if, for every graph G on n vertices, $A(G)$ approximates $\omega(G)$ within a factor of $c(n)$.*

Recently, using the theory of *interactive proofs*, it has been established that approximating $\omega(G)$ within a factor of n^ϵ (for some ϵ) is NP-complete [13, 7, 6]. It is of great interest to find a "graph theoretic" proof that clique is hard to approximate, without the use of interactive proofs. We make a modest step in this direction, by showing that (a small) part of the interactive proofs machinary can be replaced by a graph theoretic tool. Our proof is based on an approach of Berman and Schnitger [10], and essentially is a derandomization of a randomized reduction that they construct.

The basic graph theoretic tool that we use is that of graph products.

Definition 4 *Let $G = (V, E)$ be a graph. For a positive integer k, the k-*fold graph product *of G with itself, denoted by $G^k = (V^k, E_k)$, is the graph on $V^k = V \times V \times \ldots \times V$ such that $(u_1 u_2 \ldots u_k, w_1 w_2 \ldots w_k) \in E_k$ if and only if the set $\{u_1, u_2 \ldots, u_k, w_1, w_2, \ldots, w_k\}$ is a clique in G (note that the u_i and w_j do not have to be distinct).*

If G has n vertices, then G^k has n^k vertices. Recall that $\omega(G)$ denotes the size of the maximum clique in G. It is easy to verify that $\omega(G^k) = (\omega(G))^k$.

Graph products are used in order to "amplify gaps" in hardness results for approximation algorithms ([15], and see also [20]). If no polynomial time algorithm approximates clique

within a factor of c_1 (for some arbitrary $c_1 > 1$), then for any c_2 ($c_2 > c_1$), no polynomial time algorithm approximates clique within a factor of c_2. (The gap is amplified from c_1 to any arbitrary constant c_2.) The proof goes as follows: Assume that polynomial time algorithm B approximates clique within a factor of c_2. Derive a contradiction by designing a polynomial time algorithm A that approximates clique within a factor of c_1. Let $k = \lceil \frac{\log c_2}{\log c_1} \rceil$. Then A outputs $(B(G^k))^{1/k}$.

Can the graph product argument be pushed to derive even stronger consequences? Close inspection of the graph product argument reveals two limitations that it has:

1. If k above is nonconstant, then the number of vertices in G^k is superpolynomial.
2. For a given value of k, the number of vertices in G^k is $N = n^k$, and the ratio of approximation that is reached is $(c_1)^k$. Thus even if k is nonconstant, the largest ratio of approximation that we can exclude never reaches N^ϵ.

A way of overcoming both limitations was suggested by Berman and Schnitger [10]. The following is implicit in [10], and appears explicitly in [11] under the name of *randomized graph products*.

Let $0 < a < b < 1$ be two constants, let $\mathcal{G}_a$ be a family of graphs that satisfy $\omega(G) < an$, let $\mathcal{G}_b$ be a family of graphs that satisfy $\omega(G) > bn$, and let $\mathcal{G} = \mathcal{G}_a \bigcup \mathcal{G}_b$.

Assumption 1 *There exist constants $0 < a < b < 1$ such that it is NP-hard to decide for $G \in \mathcal{G}$ whether $G \in \mathcal{G}_a$ or $G \in \mathcal{G}_b$.*

Under Assumption 1 (which is a theorem by now [6]), we want to prove that it is NP-hard to approximate clique within n^ϵ, for some ϵ that depends on a and b. We sketch the randomized reduction of [10, 11].

On input $G \in \mathcal{G}$, we want to determine whether $G \in \mathcal{G}_a$ or $G \in \mathcal{G}_b$. Let $k = \Theta(\log n)$. Consider G^k. It has n^k vertices. If $G \in \mathcal{G}_a$ then $\omega(G^k) < (an)^k$. If $G \in \mathcal{G}_b$ then $\omega(G^k) > (bn)^k$. Sample at random $N = \Omega((1/a)^k)$ vertices from G^k (note that N is polynomial in n), and construct the vertex induced subgraph, which we denote by RG^k (R for randomized).

If the sampling procedure is truly random, then for each clique of size c in G^k, the expected number of sample points that belong to the clique is $(c/n^k)(1/a)^k$. We make the simplifying assumption that this is indeed the case, and that it holds simultaneously for all cliques in G^k (the largest clique is likely to be slightly larger, but this doesn't significantly affect the argument). Thus if $G \in \mathcal{G}_a$, then $\omega(RG^k) \simeq ((an)^k/n^k)(1/a)^k = 1$, and if $G \in \mathcal{G}_b$, then $\omega(RG^k) \simeq ((bn)^k/n^k)(1/a)^k = (b/a)^k$. Thus it suffices to approximate clique to within a factor of $(b/a)^k$ in order to distinguish between graphs in $\mathcal{G}_a$ and graphs in $\mathcal{G}_b$. Recall that the total number of vertices in RG^k is $N \simeq (1/a)^k$. Thus relative to N, the required ratio of approximation is N^ϵ, for $\epsilon \simeq 1 - \frac{\log b}{\log a}$.

The above gap amplification procedure is randomized. The purpose of this paper is to provide a deterministic gap amplification technique that has the same effect as the randomized graph products. Since our technique is a derandomization of the Berman and Schnitger technique, we call it *derandomized graph products*.

Remark: The concept of randomized graph products and its relation to clique approximation follows from the work of Berman and Schnitger [10], and is presented in [11]. At that time, Assumption 1 above was not known to hold. Instead, Berman and Schnitger considered MAX-SNP, a class of approximation problems defined in [24]. It is a simple matter to show that the problem presented in Assumption 1 is MAX-SNP-hard. That is, if for any $0 < a < b < 1$ a polynomial time algorithm could distinguish between the classes $\mathcal{G}_a$ and $\mathcal{G}_b$, then every problem in MAX-SNP would be approximable within factors arbitrarily close to 1. Berman and Schnitger concluded that it is MAX-SNP-hard under randomized reductions to approximate clique within a factor of n^ϵ.

To understand the part of the interactive proofs machinary that can be replaced by graph products, consider the state of affairs concerning clique approximation (following the paper [6]).

1. It is NP-hard to approximate MAX-SNP-hard problems within a factor of $1 + \delta$ (for some $\delta > 0$ that depends on the nature of the particular MAX-SNP-hard problem). Hence Assumption 1 above holds.

2. Using error reduction techniques for interactive proof systems (specifically suggested in [25]), one can show that approximating clique within a factor of n^ϵ is NP-hard.

3. By Assumption 1 and using randomized graph products, one can show that if clique can be approximated within a factor of n^ϵ (for some small enough ϵ), then any NP-statement can be decided in *random* polynomial time.

Derandomized graph products have no effect on Item 1 above, whose proof still requires a reduction from interactive proofs (though we hope that this situation will change in the future). However, we replace Item 2 (which is based on interactive proofs) by a graph theoretic tool (derandomized graph products) that is the deterministic analog of Item 3. (See remark at the end of Section 2 for a quantatative comparison between the estimates provided by randomized and derandomized graph products.)

1.2 Derandomization

The tool we use in order to derandomize the randomized graph products is that of random walks on constant degree expander graphs. This tool was first developed by [1]. Let U be a universe of n items from which we want to sample k items. Arrange the n items

as the vertices of a special type of constant degree graph (expander) H. The graph H is constructed in such a way that random walks on H have "nice" properties. In order to sample k points, start at a random vertex of H, and take a random walk of $k-1$ steps. The k vertices that are visited comprise the sample.

In terms of random bits used, a truly random sample of k points requires $k \log n$ random bits, whereas the random walk based sample requires only $\log n + O(k)$ random bits. Thus if $k = O(\log n)$, there are only polynomially many possible sets that arise in the process of random walk based sampling, and one may enumerate all possible sample sets in deterministic polynomial time. Each such sample set becomes a vertex in a new graph DG^k, and two vertices of this graph are connected by an edge if the $2k$ vertices from which they are composed form a clique in G. The main question is how well does random walk based sampling model truly random sampling. The answer depends on the property that one wants to consider.

For our goal of derandomizing graph products, we need to consider two properties. The first is to upper bound the probability that all k pseudorandom sample points fall within a prespecified set. This property is important for ensuring that if $G \in \mathcal{G}_a$ (that is, G has no large clique), then DG^k does not have a large clique. A sufficiently strong upper bound for our purpose was already proved in [1]. The second property that we need is to lower bound the probability that all k pseudorandom sample points do fall inside a prespecified set. This property insures that if $G \in \mathcal{G}_b$ (G has a large clique), then DG^k has a large clique. To the best of our knowledge, this question was not explicitly addressed before. It has been proved that the fraction of pseudorandom sample points that fall within a set is expected to be roughly proportional to its size [12, 17], but even for fairly large sets, this does not exclude the possibility that for every k size pseudo random sample, at least one point lies outside the set.

The main new technical lemma that we need is to lower bound the probability that a random walk (on a special type of expander graph) stays inside a prespecified set of nodes. This is not difficult to show, using the results of [3] and [18]. It turns out that the proof of this property does not rely at all on the expansion property of the graph, but rather on how "different" it is from a bipartite graph. Observe that for bipartite graphs, there is a set that contains at least half the vertices, such that any walk of length 1 has at least one of the two vertices that it visits outside the prespecified set. Thus it is apparent that having H significantly different from bipartite is a necessary condition for the above property to hold. Similarly, it is essential that H does not have "large" independent sets. A typical measure for these properties is the value of the smallest (most negative) eigenvalue of the adjacency matrix of H (see for example Corollary 2 in [18]), and indeed our lower bound is expressed in terms of this eigenvalue.

1.3 A stronger lemma

For the purpose of derandomized graph products, it suffices to consider a fixed set of size bn, and lower bound the probability that a random walk stays inside the set. However, in other contexts, it may be desirable to consider a sequence of k different sets, each of size bn, and to lower bound the probability that a random walk has its ith step in the ith set, for all $i \leq k$ simultaneously. The related question of upper bounding this probability is central to randomness-efficient error-reduction procedures for interactive proofs [8]. In anticipation of future applications, we provide a lower bound for this case. This lower bound uses in an essential way the expansion properties of the underlying graph H, and its proof is different from that of the lower bound discussed above.

2 Derandomized graph products

Our graph H is based on an explicit construction of a constant degree expander graph. It is simplest to assume that H is a non-bipartite d-regular Ramanujan graph as in [21], [22], where $d > \frac{16}{(b-a)^2}$. (If n is such that no respective H graph exists, then G can be slightly modified by adding dummy vertices until a desirable value of n is reached.)

We construct the graph DG^k (D stands for "derandomized") in the following way. We consider all possible random walks of length $k-1$ on H, where $k = \Theta(\log n)$ will be determined later. When at vertex v, the walk moves along one of the edges incident with v to the vertex at the other end of the edge. Note that there are nd^{k-1} such walks. Each walk corresponds to a single vertex of DG^k. Two vertices of DG^k are connected by an edge if the $2k$ vertices (not all of which have to be distinct) from which they are composed form a clique in G.

Recall that $\omega(G)$ denotes the size of the maximum clique in G. We want to bound $\omega(DG^k)$, the size of the maximum clique in the derandomized graph product of G. We will prove both an upper bound and a lower bound on $\omega(DG^k)$. These bounds are expressed in terms of the parameters $\omega(G)$, k, and d. Additional parameters of importance are the eigenvalues $\lambda_0 \geq \lambda_1 \geq \ldots \geq \lambda_{n-1}$ of the matrix A, where A is the transition matrix of the random walk on the graph H. (Entry A_{ij} specifies the probability of having the walk at vertex i move to vertex j. The matrix A is symmetric and has only real eigenvalues.) Observe that $\lambda_0 = 1$, and denote $\max(\lambda_1, |\lambda_{n-1}|)$ by λ. For nonbipartite Ramanujan graphs, $\lambda \leq 2\sqrt{d-1}/d$ (and this is known to be asymptotically the smallest possible value of λ for large n and small d, see [2]).

Theorem 1 *For any graph G and any odd k, the size of the maximum clique in the derandomized graph product DG^k is related to the size of the maximum clique in G in the following way:*

$$\omega(G)d^{k-1}(\frac{\omega(G)}{n}+\lambda_{n-1}(1-\frac{\omega(G)}{n}))^{k-1} \leq \omega(DG^k) \leq \omega(G)d^{k-1}(\frac{\omega(G)}{n}+\lambda_1(1-\frac{\omega(G)}{n}))^{k-1}.$$

In order to prove the theorem we need two lemmas. The first one is Lemma 3 of [18] which we state without proof.

Lemma 1 *([18]) Let H be a d-regular graph on n vertices, and let A and $1=\lambda_0 \geq \cdots \geq \lambda_{n-1}$ be as above. For a set of vertices W of H, let $\mu = |W|/n$ denote the density of W, and let γ denote the largest eigenvalue of $\frac{1}{d}M$, where M is the adjacency matrix of the induced subgraph of H on W. Then*

$$\gamma \leq \mu + \lambda_1(1-\mu).$$

The second lemma is similar to Lemma 2.3 of [3], and its proof presented below follows the one in [3].

Lemma 2 *Let $H, d, n, A, \lambda_0, \cdots, \lambda_{n-1}, W$ and μ be as in the previous lemma, and let d_W be the average degree of the induced subgraph of H on W. Then*

$$\frac{d_W}{d} \geq \mu + \lambda_{n-1}(1-\mu).$$

Proof. Let $v_0, v_1, \cdots, v_{n-1}$ be an orthonormal set of eigenvectors of A, where v_i is an eigenvector corresponding to the eigenvalue λ_i. Thus v_0 is the constant vector each coordinate of which is $1/\sqrt{n}$. Let u be the characteristic vector of W, and put $w = |W|$. Clearly if $u = \sum_{i=0}^{n-1} c_i v_i$ then c_0 is the scalar product of u and v_0 and hence $c_0^2 = w^2/n$. It follows that $u^t dAu$ (which is the sum of degrees in the induced subgraph of H on W) is

$$d\sum_{i=0}^{n-1} c_i^2\lambda_i = (w^2/n)d + d\sum_{i=1}^{n-1} c_i^2\lambda_i \geq (w^2/n)d + d(w-w^2/n)\lambda_{n-1}.$$

(Here we used the fact that the square of the L_2-norm of u is w and hence so is the sum of the c_i^2). Since $\mu = w/n$ this gives that for the average degree d_W

$$wd_W \geq w\mu d + dw(1-\mu)\lambda_{n-1},$$

implying the required bound $d_W/d \geq \mu + (1-\mu)\lambda_{n-1}$. □

We can now prove the key proposition needed to prove Theorem 1. (Our contribution to the proof of the proposition is confined to the proof of the lower bound. The upper bound is Corollary 1 in [18], and we give its proof only for the sake of completeness.)

Proposition 2 *Let $H, d, n, A, \lambda_0, \cdots, \lambda_{n-1}$ be as above, let W be a set of w vertices in H and put $\mu = w/n$. Let $P = P(W,k)$ be the total number of walks of k vertices that stay in W. Assume (for the lower bound only) that k is odd and that $\mu + \lambda_{n-1}(1-\mu) \geq 0$. Then*

$$wd^{k-1}(\mu + \lambda_{n-1}(1-\mu))^{k-1} \leq P \leq wd^{k-1}(\mu + \lambda_1(1-\mu))^{k-1}.$$

Proof. Define $L = \frac{1}{d}M$ where M is the adjacency matrix of the induced subgraph of H on W. Let $\gamma_1 \geq \gamma_2 \ldots \geq \gamma_w$ be the eigenvalues of L and let $u_1, \ldots, u_w$ be the corresponding eigenvectors. Let $p(W,k) = P(W,k)/nd^{k-1}$ denote the probability that a random walk of length $k-1$ (i.e., a walk of k vertices) stays in W. Let u be the all 1 vector of length w and observe that $p(W,k) = \frac{1}{n}u^t L^{k-1}u$. Therefore, if $u = \sum_{i=1}^{w} c_i u_i$ then

$$p(W,k) = \frac{1}{n}\sum_{i=1}^{w} c_i^2 \gamma_i^{k-1}. \tag{1}$$

Since $\sum_{i=1}^{w} c_i^2 = w$ (as this is the square of the L_2-norm of u), and, by Lemma 1 and the Perron-Frobenius Theorem (cf., e.g., [23]), each γ_i is in absolute value at most $\gamma_1 \leq \mu + \lambda_1(1-\mu)$, the required upper bound follows from (1) (for any k). To get the lower bound, observe that

$$\frac{n}{w}p(W,k) = \frac{1}{w}\sum_{i=1}^{w} c_i^2 \gamma_i^{k-1}.$$

Now consider the function $f(x) = x^{k-1}$, where k is odd. This is a convex function (its second derivative is nonnegative), and hence for any nonnegative $\alpha_1, ..., \alpha_m$ with $\sum_{i=1}^{m} \alpha_i = 1$, and for any $x_1, ..., x_m$, Jensen's inequality implies that $\sum_{i=1}^{m} \alpha_i f(x_i) \geq f(\sum_{i=1}^{m} \alpha_i x_i)$. Observing that $\frac{1}{w}\sum_{i=1}^{w} c_i^2 = 1$ we obtain

$$\frac{1}{w}\sum_{i=1}^{w} c_i^2 \gamma_i^{k-1} \geq (\frac{1}{w}\sum_{i=1}^{w} c_i^2 \gamma_i)^{k-1}$$

Observe that $\frac{1}{w}\sum c_i^2 \gamma_i$ is $\frac{n}{w}p(W,2)$ which is simply, as can be easily checked, d_W/d, where d_W denotes the average degree in the induced subgraph of H on W. Recall that by Lemma 2, $d_W/d \geq \mu + \lambda_{n-1}(1-\mu)$ and hence,

$$(\frac{1}{w}\sum_{i=1}^{w} c_i^2 \gamma_i)^{k-1} \geq (\mu + \lambda_{n-1}(1-\mu))^{k-1},$$

giving the required lower bound. □

Proof of Theorem 1. Let W be a fixed clique in G, and let $\mu = |W|/n$ denote its density. The vertices of DG^k are constructed by taking random walks of length k on H. We want to

count how many random walks fall entirely within the vertex set W. A lower and an upper estimate for this number follows from Proposition 2. Since a set of walks forms a clique in DG^k if and only if they all stay in the same clique of G the assertion of the theorem follows. □

For our application (seperating between $\mathcal{G}_a$ and $\mathcal{G}_b$), it suffices to assume that $\lambda < \frac{b-a}{2}$. This governs the choice of d in the Ramanujan graph H.

The number of vertices in DG^k is $N = nd^{k-1}$. If $G \in \mathcal{G}_a$, then $\omega(DG^k) \leq (an)d^{k-1}(a+\lambda)^{k-1}$. If $G \in \mathcal{G}_b$, then $\omega(DG^k) \geq (bn)d^{k-1}(b-\lambda)^{k-1}$. By making k sufficiently large (but still logarithmic in n), it follows that clique cannot be approximated to within N^ϵ, for ϵ close to $\log \frac{b-\lambda}{a+\lambda} / \log d$.

Remarks:

1. Qualitatively, derandomized graph products achieve the same effect as randomized graph products do. Quantatively, there is a difference. The value of ϵ achieved by the randomized version is better than the value achieved by the derandomized version.

   As a typical example, assume that $a << 1$ and $b = 2a$. Then randomized graph products give a value of $\epsilon \simeq 1 - \log 2a / \log a = 1/\log(1/a)$. For derandomized graph products based on Ramanujan graphs, recall that $d \simeq 4/\lambda^2$, and hence $\epsilon \simeq \log \frac{b-\lambda}{a+\lambda} / 2\log(2/\lambda)$. For, say, $\lambda \simeq a/\log(1/a)$, this gives $\epsilon \simeq 1/2\log(1/a)$.

2. Our proof goes through even if H is nonsimple (it has self loops and parallel edges). This gives greater flexibility in the design of H. In particular, one may start with a constant degree bipartite expander, such as the one constructed in [14], add self loops to each vertex (to destroy bipartiteness and make $|\lambda_{n-1}|$ bounded away from λ_0), and take the product of the resulting graph with itself sufficiently many times (until a desired value of λ_0/λ is reached).

3. Note that the lower bound in Theorem 1 depends on λ_{n-1}, and not on λ_1. Thus if all eigenvalues of H are nonnegative (this can be obtained, for example, by considering the square H'^2 of a given expander H'), then the lower bound is at least $\omega(G)d^{k-1}(\frac{\omega(G)}{n})^{k-1}$. This results in a slightly larger value of ϵ.

3 Independent sets in graphs of bounded degree

An independent set in a graph G is a set of vertices such that no two vertices in the set are connected by an edge. Let $\alpha(G)$ denote the size of the maximum independent set in the graph G. Any independent set in a graph G is a clique in $\bar{G}$, the complement of G. It follows that the same unapproximability results that hold for $\omega(G)$ hold also for $\alpha(G)$.

An interesting question is how well can $\alpha(G)$ be approximated in graphs of maximum degree at most Δ. Berman and Furer [9] design a polynomial time algorithm that approximates $\alpha(G)$ to within a ration of $\Delta/5 + c$, where $0 < c < 1$ is some universal constant. Halldorsson and Radhakrishnan [16] analyse the performance of the greedy algorithm on this problem, and discuss several extensions. As to hardness results for this problem, it follows from [6] and [24] that even if $\Delta = 3$, it is NP-hard to approximate $\alpha(G)$ within a factor of $1 + \epsilon$, for some $\epsilon > 0$. Halldorsson conjectured (private communication) that this NP-hardness result can be extended to showing that it is NP-hard to approximate $\alpha(G)$ within a factor of Δ^ϵ, for some $\epsilon > 0$. We show that derandomized graph products imply the correctness of this conjecture.

Theorem 3 *For some $\epsilon > 0$ and every $\Delta \geq 3$, it is NP-hard to approximate $\alpha(G)$ within a factor of Δ^ϵ in graphs of maximum degree at most Δ.*

Proof. Let $0 < a < b < 1$ be two constants, let $\mathcal{G}_a$ be a family of graphs that satisfy $\alpha(G) < an$, let $\mathcal{G}_b$ be a family of graphs that satisfy $\alpha(G) > bn$, and let $\mathcal{G} \subset \mathcal{G}_a \bigcup \mathcal{G}_b$. By [6] and [24], there exist constants $0 < a < b < 1$ such that it is NP-hard to decide for $G \in \mathcal{G}$ whether $G \in \mathcal{G}_a$ or $G \in \mathcal{G}_b$, even if $\mathcal{G}$ contains only graphs of maximum degree 3.

When considering derandomized graph products, we now define, for a graph $G \in \mathcal{G}$, the k-fold (modified) derandomized graph product $\tilde{DG}^k$. This modified graph product is constructed as described in Section 2, but with the changes required in order to handle independent sets rather than cliques. That is, $\tilde{DG}^k$ in our use in this section is just the complement of the k-fold derandomized graph product of the complement of G, that is, the graph $\overline{D\bar{G}^k}$, using the notation of Section 2. From Theorem 1 we have that

$$\alpha(G)d^{k-1}(\frac{\alpha(G)}{n} - \lambda)^{k-1} \leq \alpha(\tilde{DG}^k) \leq \alpha(G)d^{k-1}(\frac{\alpha(G)}{n} + \lambda)^{k-1}.$$

Assume that d, the degree of the expander graph, is sufficiently large so that $3\lambda < b - a$. Hence the gap between the cases that $G \in \mathcal{G}_a$ and $G \in \mathcal{G}_b$ is amplified from b/a to roughly $(\frac{b-\lambda}{a+\lambda})^k > (1 + \lambda)^k$.

We now analyse Δ, the maximum degree in $\tilde{DG}^k$. Consider any vertex $v \in \tilde{DG}^k$. It is composed of k vertices $v_1, v_2, \ldots, v_k$ of G. Consider now any other vertex $u \in \tilde{DG}^k$, composed of vertices $u_1, u_2, \ldots, u_k$ of G. In order to have an edge between v and u in $\tilde{DG}^k$, there must be two indices i and j, such that (v_i, u_j) is an edge in G. There are k^2 possible ways of selecting the indices i and j. Once we select i, this fixes v_i, and there remain at most 3 ways of selecting u_j (by the degree bound on G). Thereafter, there remain d^{k-1} ways of selecting the remaining vertices u_ℓ, $\ell \neq j$. Hence the degree of v can be at most $3k^2d^{k-1}$.

To express the gap in sizes of the maximum independent set in terms of Δ, let k be an integer such that $3k^2d^{k-1} \leq \Delta \leq 3(k+1)^2d^k$. Then in graphs of maximum degree Δ, it is NP-hard to approximate $\alpha(G)$ within a factor of $(1+\lambda)^k$. Recall that for Ramanujan graphs we have $d \simeq 4/\lambda^2$, and observe that $3(k+1)^2$ is a low order term relative to d^k. By selecting ϵ such that $(4/\lambda^2)^\epsilon \simeq 1+\lambda$ we obtain that it is NP-hard to approximate $\alpha(G)$ within a factor of Δ^ϵ. □

4 Staying inside changing sets

In this section we lower bound the probability that a random walk of length k stays inside a sequence of k different sets $W_1, W_2, \ldots, W_k$, each of size μn. Again we use linear algebra, but now we need a bound on λ, the second largest eigenvalue in absolute value of the transition matrix A. Our proof extends the techniques in [1].

We will also need the following notation:

P is the vector space R^n.
X is the subspace of multiples of $1_n = (1/n, 1/n, \ldots, 1/n)$.
Y is the subspace orthogonal to X.
For $p = (p_1, p_2, \ldots, p_n) \in P$
  $|p| = \sum_{i=1}^n |p_i|$, the L_1-norm.
  $\|p\| = \sqrt{\sum_{i=1}^n p_i^2}$, the L_2-norm.
$N_i : P \to P$ is the linear transformation

$$N_i(e_j) = \begin{cases} e_j & \text{if } j \in W_i \\ 0 & \text{otherwise,} \end{cases}$$

where e_j is the basis vector with a 1 in the jth place and 0's elsewhere.

Thus, if the vector $p = (p_1, \ldots, p_n)$ represents the probabilities of being at the different vertices, i.e. $p_i = \Pr[\text{particle is at vertex } i]$, then the ith component of $N_1 p$ is the probability of being at vertex i and of being in W_1. Thus $|N_1 p|$, the sum of the components, is the probability of being in W_1. Extending this, we see that the probability that the ith vertex of a random walk taking $k-1$ steps from a uniform starting vertex always lies in W_i for all i is $|(N_kA)(N_{k-1}A)\cdots(N_2A)N_1 1_n|$. Let $v_1 = N_1 1_n$ and $v_i = N_i A v_{i-1}$ for $i > 1$.

Lemma 3 *If $\mu \geq 6\lambda$, then $|v_{i+1}| > (\mu - 2\lambda)|v_i|$.*

Proof. Let $v_i = x_i + y_i$, where $x_i \in X$, $y_i \in Y$. All coordinates of v_i are nonnegative (they represent probabilities), and the sum of coordinates of y_i is 0 (by orthogonality to X). It follows that $|v_i| = |x_i| = \sqrt{n}||x_i||$, and the lemma is equivalent to $||x_{i+1}|| > (\mu - 2\lambda)||x_i||$.

We prove this by induction, including in our inductive statement that $||y_i|| \leq t||x_i||$. We will choose t later.

We claim that

$$||x_{i+1}|| \geq \mu||x_i|| - t\lambda\sqrt{\mu(1-\mu)}||x_i|| \tag{2}$$

$$||y_{i+1}|| \leq t\lambda||x_i|| + \sqrt{\mu(1-\mu)}||x_i|| \tag{3}$$

To see these, write $v_{i+1} = N_{i+1}Ax_i + N_{i+1}Ay_i$. Since $|N_{i+1}x_i| = \mu|x_i|$,

$$N_{i+1}Ax_i = N_{i+1}x_i = \mu x_i + z_i,$$

where $z_i \in Y$ and

$$||z_i||^2 = ||N_{i+1}x_i||^2 - ||\mu x_i||^2 = \mu(1-\mu)||x_i||^2.$$

Further note that $||Ay_i|| \leq \lambda||y_i||$ and $Ay_i \in Y$.

Now we claim that for $p = (p_1, \ldots, p_n) \in Y$, the component π of $N_{i+1}p$ in X has magnitude at most $\sqrt{\mu(1-\mu)}||p||$. To see this, note that $\pi = \sum_{j \in W_{i+1}} p_j = -\sum_{j \notin W_{i+1}} p_j$ (since $p \in Y$). Thus for a fixed value of $||p||^2 = \sum_j p_j^2$, we maximize $|\pi|$ by setting $p_j = p_k$ for j, k either both in W_{i+1} or both not in W_{i+1}. Doing the algebra yields the claim.

Letting w_i be the component of $N_{i+1}Ay_i$ in X, we have

$$x_{i+1} = \mu x_i + w_i,$$

so

$$||x_{i+1}|| \geq \mu||x_i|| - ||w_i|| \geq \mu||x_i|| - \lambda\sqrt{\mu(1-\mu)}||y_i||,$$

giving (2) after using $||y_i|| \leq t||x_i||$. Also

$$||y_{i+1}|| \leq ||Ay_i|| + ||z_i||,$$

which yields (3).

Once we have these equations, we choose t so as to satisfy the inductive assumption $||y_i|| \leq t||x_i||$. Observe that $||y_1|| = \sqrt{(1-\mu)/\mu}||x_1||$, and that for any value of t strictly larger than $\sqrt{(1-\mu)/\mu}$, equations (2) and (3) imply that $||y_{i+1}|| \leq t||x_{i+1}||$, if λ is sufficiently small. A simple calculation suffices to verify that if $\lambda \leq \mu/6$ then we may choose $t = 2\sqrt{(1-\mu)/\mu}$. The lemma then follows from equation (2). Also, if we assume $\lambda \leq \mu^2/2$, then by taking $t = 1/\sqrt{\mu(1-\mu)}$ we can conclude that $|v_{i+1}| \geq (\mu - \lambda)|v_i|$. □

This yields:

Theorem 4 *The probability that a random walk for $k-1$ steps from a uniformly random starting vertex stays inside $W_1, W_2, \ldots, W_k$ (each of density $\mu \geq 6\lambda$) is at least $\mu(\mu - 2\lambda)^{k-1}$.*

It is worth noting that the same computation also shows that the probability that a random walk for $k-1$ steps from a uniformly random starting vertex stays inside $W_1, W_2, \ldots, W_k$ (each of density $\mu \geq 6\lambda$) is at most $\mu(\mu + 2\lambda)^{k-1}$. This strengthens the estimate in [8].

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, "Deterministic Simulation in Logspace", *Proc. of the 19^{th} STOC 1987, pp. 132-140.*

[2] N. Alon, "Eigenvalues and expanders", *Combinatorica 6(1986), 83-96.*

[3] N. Alon, F. R. K. Chung, "Explicit construction of linear sized tolerant networks", *Discrete Math. 72(1988), 15-19*; (Proc. of the First Japan Conference on Graph Theory and Applications, Hakone, Japan, 1986.)

[4] E. Amaldi, V. Kann, "The complexity and approximability of finding maximum feasible subsystems of linear relations", *Theoretical Computer Science*, to appear.

[5] S. Arora, L. Babai, J. Stern, and Z. Sweedyk, "The Hardness of Approximate Optima in Lattices, Codes and Linear Equations", *Proc. of 34th FOCS, 1993, pp., 724-733.* Final version submitted to JCSS.

[6] S. Arora, C. Lund, R. Motwani, M. Sudan, M. Szegedy, "Proof Verification and Hardness of Approximation Problems", *Proc. of 33^{rd} FOCS, 1992, pp. 14-23.*

[7] S. Arora, S. Safra, "Probabilistic Checking of Proofs; A New Characterization of NP", *Proc. of 33^{rd} FOCS, 1992, pp. 2-13.*

[8] M. Bellare, O. Goldreich, S. Goldwasser, "Randomness in Interactive Proofs", *Proc. of 31^{st} FOCS, 1990, pp. 563-572.*

[9] P. Berman, M. Furer, "Approximating maximum independent set in bounded degree graphs", *Proc. of 5th ACM-SIAM Symposium on Discrete Algorithms, 1994.*

[10] P. Berman, G. Schnitger, "On the Complexity of Approximating the Independent Set Problem", *Information and Computation, Vol. 96, pp. 77-94, 1992.*

[11] A. Blum. *Ph.D. dissertation, MIT.*

[12] A. Cohen, A. Wigderson, "Dispersers, deterministic amplification, and weak random sources", *Proc. of* 30^{th} *FOCS, 1989, pp. 14-19.*

[13] U. Feige, S. Goldwasser, L. Lovász, S. Safra, M. Szegedy, "Approximating Clique is Almost NP-complete", *Proc. of* 32^{nd} *FOCS, 1991, pp. 2-12.*

[14] O. Gabber, Z. Galil, "Explicit Constructions of Linear Sized Superconcentrators", *JCSS, 22(3), pp. 407-420, 1981.*

[15] M. Garey, D. Johnson, "Computers and Intractability: A Guide to the Theory of NP-Completeness", *Freeman, 1979.*

[16] M. Halldorsson, J. Radhakrishnan, "Greed is good: approximating independent sets in sparse and bounded-degree graphs", *Proc. of 26th ACM Symposium on the Theory of Computing, 439–448, 1994.*

[17] R. Impagliazzo, D. Zuckerman, "How to Recycle Random Bits", *Proc. of* 30^{th} *FOCS, 1989, pp. 248-253.*

[18] N. Kahale, "Better Expansion for Ramanujan Graphs", *Proc.* 33^{rd} *FOCS, 1992, pp. 398-404.*

[19] R. Karp, "Reducibility Among Combinatorial Problems", *In R. Miller and J. Thatcher, editors, Complexity of Computer Computations, pp. 85-103, Plenum Press, 1972.*

[20] N. Linial, U. Vazirani, "Graph Products and Chromatic Numbers", *Proc.* 30^{th} *FOCS, 1989, pp. 124-128.*

[21] A. Lubotsky, R. Phillips, P. Sarnak, "Ramanujan Graphs", *Combinatorica, Vol. 8(3), 1988, pp. 261-277.*

[22] G. A. Margulis, "Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators", *Problemy Peredachi Informatsii 24(1988), 51-60 (in Russian).* English translation in Problems of Information Transmission 24(1988), 39-46.

[23] M. Marcus, H. Minc, "A Survey of Matrix Theory and Matrix Inequalities", *Allyn and Bacon, Inc., Boston, 1964.*

[24] C. Papadimitriou, M.Yannakakis, "Optimization, Approximation, and Complexity Classes", *JCSS, Vol. 43, pp. 425-440, 1991.*

[25] D. Zuckerman, "Simulating BPP Using a General Weak Random Source", *Proc.* 32^{nd} *FOCS, 1991, pp. 79-89.*