Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 132 (2006)

Tight Integrality Gaps for Lovasz-Schrijver LP Relaxations of Vertex Cover and Max Cut

Grant Schoenebeck* Luca Trevisan† Madhur Tulsiani‡

Abstract

We study linear programming relaxations of Vertex Cover and Max Cut arising from repeated applications of the "lift-and-project" method of Lovasz and Schrijver starting from the standard linear programming relaxation.

For Vertex Cover, Arora, Bollobas, Lovasz and Tourlakis prove that the integrality gap remains at least $2-\varepsilon$ after $\Omega_\varepsilon(\log n)$ rounds, where n is the number of vertices, and Tourlakis proves that integrality gap remains at least $1.5-\varepsilon$ after $\Omega((\log n)^2)$ rounds. Fernandez de la Vega and Kenyon prove that the integrality gap of Max Cut is at most $\frac{1}{2}+\varepsilon$ after any constant number of rounds. (Their result also applies to the more powerful Sherali-Adams method.)

We prove that the integrality gap of Vertex Cover remains at least $2-\varepsilon$ after $\Omega_\varepsilon(n)$ rounds, and that the integrality gap of Max Cut remains at most $1/2+\varepsilon$ after $\Omega_\varepsilon(n)$ rounds.

1 Introduction

Lovasz and Schrijver [LS91] describe a method, referred to as LS, to tighten a linear programming relaxation of a 0/1 integer program. The method adds auxiliary variables and valid inequalities, and it can be applied several times sequentially, yielding a sequence (a "hierarchy") of tighter and tighter relaxations. The method is interesting because it produces relaxations that are both tightly constrained and efficiently solvable.

For a linear programming relaxation K, denote by $N(K)$ the relaxation obtained by the application of the LS method, and by $N^k(K)$ the relaxation obtained by applying the LS method k times. Following standard terminology, we will refer to the process of applying the LS method k times as the application of k *rounds* of LS. If the original linear program K had n variables and poly(n) constraints, then it is possible to optimize over $N^k(K)$ in time $n^{O(k)}$, which is polynomial for constant k and sub-exponential for $k=o(n/\log n)$.

Lovasz and Schrijver [LS91] prove that $N^k(K)$ enforces all valid linear inequalities over k-tuples of variables and, in particular, $N^n(K)$ contains only convex combinations of valid integral solutions. In addition to enforcing "local" constraints, relaxations of the form $N^k(K)$ also introduce more global constraints via the auxiliary variables. If a linear programming relaxation maintains an integrality

*This material is based upon work supported under a National Science Foundation Graduate Research Fellowship.

†This material is based upon work supported by the National Science Foundation under grant CCF 0515231 and by the US-Israel Binational Science Foundation Grant 2002246.

‡This material is based upon work supported by the National Science Foundation under grant CCF 0515231.

ISSN 1433-8092

gap[1] g or larger even after $\Omega(n)$ applications of LS, then we can take this fact as evidence that linear programming methods are unlikely to yield an efficient approximation algorithm with performance ratio better than g.

The approximability of the Vertex Cover problem is one of the outstanding questions in the study of the complexity of approximation. A simple linear programming relaxation yields a 2-approximate algorithm, and no better polynomial time approximation is known. Dinur and Safra [DS05] prove that, unless P=NP, there is no polynomial-time 1.36 approximation algorithm for Vertex Cover, and Khot and Regev [KR03] prove that if there is a $2-\varepsilon$ approximate algorithm, $\varepsilon > 0$, then the Unique Games Conjecture fails.

Arora, Bollobas, Lovasz, and Tourlakis [ABLT06, Tou06] consider the question of proving integrality gaps for LS relaxations of Vertex Cover. They show that even after $\Omega_\varepsilon(\log n)$ rounds the integrality gap is at least $2-\varepsilon$ [ABLT06], and that even after $\Omega_\varepsilon((\log n)^2)$ rounds the integrality gap is at least $1.5-\varepsilon$ [Tou06].

For some optimization problems, such as Max Cut and Sparsest Cut, the best known approximation algorithms are based on semidefinite programming. For Max Cut, in particular, the Goemans-Williamson approximation algorithm based on semidefinite programming [GW95] achieves an approximation ratio of about .878, whereas no known, efficiently solvable, linear programming relaxation of Max Cut has an integrality gap better than 1/2.

Fernandez de la Vega and Kenyon [FdlVKM07] prove, for every $\varepsilon > 0$, that the integrality gap of Max Cut remains at most $\frac{1}{2}+\varepsilon$ even after a constant number of rounds of LS. Their result also applies to the more powerful Sherali-Adams method [SA90]. (The relaxation obtained after r rounds of Sherali-Adams is at least as tight as the relaxation obtained after r round of Lovasz-Schrijver.)

Our Result

We prove that after $\Omega_\varepsilon(n)$ rounds of LS the integrality gap of Vertex Cover remains at least $2-\varepsilon$.

Following a suggestion of James Lee, we then apply our methods to the Max Cut problem, and we show that after $\Omega_\varepsilon(n)$ rounds of LS the integrality gap of Max Cut remains at most $\frac{1}{2}+\varepsilon$.

The instances for which we prove the integrality gap results are (slight modifications of) sparse random graphs. In such graphs, the size of the minimum vertex cover is $\approx n$, where n is the number of vertices, while we show the existence of a fractional solution of cost $n \cdot (\frac{1}{2}+\varepsilon)$ that remains feasible even after $\Omega_\varepsilon(n)$ rounds. The size of a maximum cut is $\approx \frac{m}{2}$, where m is the number of edges, while we show the existence of a fractional solution of cost $m \cdot (1-\varepsilon)$ that also remains feasible after $\Omega_\varepsilon(n)$ rounds.

We use two properties of (modified) sparse random graphs. The first property is large girth; it suffices for our application that the girth be a large constant depending on ε. The second property is that for every set of $k = o(n)$ vertices, such vertices induce a subgraph containing at most $(1+o(1))k$ edges. The same properties are also used in [ABLT06, Tou06].

In order to prove that a certain fractional solution y is feasible for a relaxation $N^k(K)$, it is sufficient to construct a matrix Y such that certain vectors obtained from the rows and columns of Y are all

[1] The integrality gap of a relaxation for a minimization problem (respectively, a maximization problem) is the supremum (respectively, the infimum), over all instances, of the ratio of the true optimum divided by the optimum of the relaxation.

feasible solutions for $N^{k-1}(K)$. (By convention, $N^0(K) := K$.) This suggest an inductive approach, where we have a theorem that says that all solutions satisfying certain conditions are feasible from $N^k(K)$; to prove the theorem we take a solution y that satisfies the conditions for a certain value of k, and then we construct a matrix Y such that all the derived vectors satisfy the conditions of the theorem for $k-1$, and hence, by inductive hypothesis, are feasible from $N^{(k-1)}(K)$, thus showing that y is feasible for $N^k(K)$. We can also use the fact that the set $N^{k-1}(K)$ is convex; this means that, once we define the matrix Y, and we have to prove that the associated vectors are in $N^{k-1}(K)$, it suffices to express each such vector as a *convex combination* of vectors that satisfy the conditions of the theorem for $k-1$. (These ideas all appear in previous work on LS and LS_+ integrality gaps.)

Roughly speaking, in the work of Arora et al. [ABLT06] on Vertex Cover, the appropriate theorem refers to solutions where all vertices are assigned the value $1/2+\varepsilon$, except for a set of exceptional vertices that belong to a set of constant-diameter disks. Oversimplifying, to prove a lower bound of k rounds, one needs to consider solutions that have up to k disks, and for the argument to go through one needs the union of the disks to induce a forest, hence the lower bound is of the same order as the girth of the graph. Tourlakis [Tou06] does better by showing that, due to extra conditions in the theorem, the subgraph induced by k "disks" has diameter $O(\sqrt{k})$, and so it contains no cycle provided that the girth of the graph is sufficiently larger than $\sqrt{k}$. This yields an integrality gap result that holds for a number of rounds up to a constant times the square of the girth of the graph.[2]

The solutions in our approach have a similar form, but we also require the disks to be far away from each other. When we start from one such solution y, we construct a matrix Y, and consider the associated vectors, we find solutions where disks are closer to each other than allowed by the theorem, and we have to express such solutions as convex combinations of allowed solutions. Roughly speaking, we show that such a step is possible provided that the union of the "problematic" disks (those that are too close to each other) induces a very sparse graph. Due to our choice of random graph, this is true provided that there are at most $c_\varepsilon \cdot n$ disks, where c_ε is a constant that depends only on ε. We also show that, in order to prove an integrality gap for k rounds, it is sufficient to consider solutions with $O(k)$ disks, and so our integrality gap applies even after $\Omega_\varepsilon(n)$ rounds. Hence (again, roughly speaking) our improvement over previous work comes from the fact that it suffices that the union of the disks induce a sparse graph (something which is true for a sublinear number of disks) rather than induce a forest (a requirement that fails once we have a logarithmic or polylogarithmic number of disks). This oversimplified sketch ignores some important technical points: We will give a more precise overview in Section 4.

Linear versus Semidefinite Relaxations

Lovasz and Schrijver [LS91] also describe a method (referred to as LS_+) to turn a linear programming relaxation into a sequence of tighter and tighter *semidefinite programming* relaxations.

After applying one round of LS_+ to the basic linear programming relaxation of Max Cut one obtains the Goemans-Williamson relaxation, which yields a .878 approximation. In contrast, we show that even after $\Omega_\varepsilon(n)$ rounds of LS the integrality gap remains $\frac{1}{2}+\varepsilon$. This gives a very strong separation

[2] Arora et al. [ABLT06, Tou06] present their proofs in the language of a "prover-verifier" game, but they can be equivalently formulated as inductive arguments.

between the approximability of LS versus LS_+ for a natural problem.

For all known approximation algorithms based on semidefinite programming, the semidefinite relaxation (or an even stronger one) can be obtained by a constant number of applications of LS_+ to a basic linear programming relaxation. Proving integrality gaps for LS_+ relaxations is hence a strong form of unconditional inapproximability result (since one rules out a class of natural and powerful algorithms). Proving a strong LS_+ integrality gap result for Vertex Cover remains an open question. We summarize below some known negative results about LS_+.

Buresh-Oppenheim, Galesy, Hoory, Magen and Pitassi [BOGH+03], and Alekhnovich, Arora, Tourlakis [AAT05] prove $\Omega(n)$ LS_+ round lower bounds for proving the unsatisfiablity of random instances of 3SAT (and, in general, kSAT with $k \geq 3$) and $\Omega_\varepsilon(n)$ round lower bounds for achieving approximation factors better than $7/8-\varepsilon$ for Max 3SAT, better than $(1-\varepsilon)\ln n$ for Set Cover, and better than $k-1-\varepsilon$ for Hypergraph Vertex Cover in k-uniform hypergraphs. We prove [STT06] a $7/6-\varepsilon$ integrality gap for Vertex Cover after $\Omega_\varepsilon(n)$ applications of LS_+.

For one round of LS_+ (or, equivalently, for the function defined as number of vertices minus the Theta function) Goemans and Kleinberg [KG98] had earlier proved a $2-o(1)$ integrality gap result by using a different family of graphs. Charikar [Cha02] proves a $2-o(1)$ integrality gap result for a semidefinite programming relaxation of Vertex Cover that includes additional inequalities. Charikar's relaxation is no tighter than 3 rounds of LS_+, and is incomparable with the relaxation obtained after two rounds.

Extending the $2-\varepsilon$ gap to a super-constant number of applications of LS_+ (or showing it does not extend) remains an outstanding open problem. However, one will likely have to use substantially different techniques then the ones in this paper (where we rely on random graph). Indeed, it is known that the Theta function of a random sparse graph is very small, and so already one round of LS_+ provides a good approximation to Vertex Cover on random graphs.

2 The Lovasz-Schrijver Hierarchy

In this section we define the Lovasz-Schrijver operator N, that maps a linear programming relaxation K into a tighter one $N(K)$. It is simpler to describe the application of the operator to *convex cones*, as defined next. A convex cone is a set $K \subseteq \mathbb{R}^d$ such that for every $\mathbf{y}, \mathbf{z} \in K$ and for every non-negative $\alpha, \beta \geq 0$ we have $\alpha\mathbf{y} + \beta\mathbf{z} \in K$.

We will use the following notation: for a matrix M, we denote by $\mathbf{M}_i$ the i-th row of M.

If $K \subseteq \mathbb{R}^d$ is a convex cone, then we define $N(K) \subseteq \mathbb{R}^d$ as follows: a vector $\mathbf{y} = (y_0, \ldots, y_{d-1})$ belongs to $N(K)$ if and only if there is a matrix $Y \in \mathbb{R}^{d\times d}$ such that

- Y is symmetric
- For all $i = 0, \ldots, d-1$, $Y_{0,i} = Y_{i,i} = y_i$
- For all $i = 0, \ldots, d-1$, $\mathbf{Y}_i$ and $\mathbf{Y}_0 - \mathbf{Y}_i$ are in K.

In such a case, we say that Y is a *protection matrix* for Y.

We also use the notation $N^0(K) := K$ and $N^k(K) := N(N^{k-1}(K))$.

Let $G = (V, E)$ be a graph, and assume $V = \{1, \ldots, n\}$. The cone of the linear progamming relaxation of the vertex cover problem is the set of vectors $\mathbf{y} \in \mathbb{R}^{n+1}$ such that

$$\begin{aligned} y_i + y_j &\geq y_0 \quad \forall (i,j) \in E \\ 0 \leq y_i &\leq y_0 \quad \forall i \in V \\ y_0 &\geq 0 \end{aligned} \qquad (VC(G))$$

The relaxation of the Vertex Cover problem arising from k rounds of Lovasz Schrijver is the solution of

$$\begin{aligned} &\min \sum_{i=1}^{n} y_i \\ &\text{subject to} \quad (y_0, y_1, \ldots, y_n) \in N^k(VC(G)) \\ &\qquad\qquad\quad y_0 = 1 \end{aligned}$$

The *integrality gap* of this relaxation for graphs of n vertices is the largest ratio between the minimum vertex cover size of G and the optimum of the above program, over all graphs G with n vertices.

The linear programming relaxation for MAX-CUT is a set of constraint on n *vertex variables* and m *edge variables*. For a vector $\mathbf{u} \in \mathbb{R}^{n+m+1}$, let u_0 be the extra coordinate for homogenization,[3] $(u_1, \ldots, u_n)$ denote the vertex variables and $(u_{e_1}, \ldots, u_{e_m})$ denote the the edge-variables. Then the cone is the solution set of the constraints

$$\begin{aligned} u_e &\leq u_i + u_j & &\forall e = (i,j) \in E \\ u_e &\leq 2u_0 - (u_i + u_j) & &\forall e = (i,j) \in E \\ 0 \leq u_i &\leq u_0 & &\forall i \in V \\ 0 \leq u_e &\leq u_0 & &\forall e \in E \\ u_0 &\geq 0 & & \end{aligned} \qquad (MC(G))$$

The relaxation of the MAX-CUT arising from r rounds of Lovasz Schrijver is the solution of

$$\begin{aligned} &\max \sum_{e \in E}^{n} u_e \\ &\text{subject to} \quad (u_0, u_1, \ldots, u_n, u_{e_1}, \ldots, u_{e_m}) \in N^r(MC(G)) \\ &\qquad\qquad\quad u_0 = 1 \end{aligned}$$

3 Our Results

Define an $(\alpha, \delta, \gamma, \eta)$ graph G on n vertices as a graph with girth $\delta \log(n)$, and such that no vertex cover of size $(1-\alpha)n$ exists and each induced subgraph of G with $k \leq \gamma n$ vertices, has at most $(1+\eta)k$ edges.

[3]Homogenization is the process of expressing a linear programming relaxation as a convex cone rather than as a subset of $[0,1]^n$.

Lemma 1 *For every $0 < \alpha < 1/125$, $\eta > 0$, there exists a $d = d(\alpha) \in \mathbb{N}$, δ, $\gamma > 0$, and $N \in \mathbb{N}$ such that for $n \geq N$ there exists an $(\alpha, \delta, \gamma, \eta)$ graph with max cut less than $\frac{1}{2}|E|(1+\alpha)$ and maximum degree at most d on n vertices. Here $d(\alpha)$ is an explicit function that depends only on α.*

Lemma 2 *For every η, δ, $\gamma > 0$, $0 < \varepsilon < 1/20$, $d \in \mathbb{N}$ if G is an $(\alpha, \delta, \gamma, \eta)$ graph with maximum degree at most d on n vertices then $(1, 1/2+\varepsilon, \dots, 1/2+\varepsilon) \in N^{\Omega_{\varepsilon,\eta,\delta,\gamma,d}(n)}(VC(G))$ if $\eta \leq \eta(\varepsilon, d)$ where $\eta(\varepsilon, d)$ is an explicit function that depends only on ε and d.*

Lemma 3 *For every η, δ, $\gamma > 0$, $0 < \varepsilon < 1/20$, $d \in \mathbb{N}$ if G is an $(\alpha, \delta, \gamma, \eta)$ graph with maximum degree at most d on n vertices then the solution $\mathbf{y}$ defined as $y_0 := 1$, $y_i := 1/2+\varepsilon$ and $y_e := 1-2\varepsilon$ is in $N^{\Omega_{\varepsilon,\eta,\delta,\gamma,d}(n)}(MC(G))$ if $\eta \leq \eta(\varepsilon, d)$ where $\eta(\varepsilon, d)$ is an explicit function that depends only on ε and d.*

Theorem 4 *For all $0 < \zeta < 1/50$, there is a constant $c_\zeta > 0$ such that, for all sufficiently large n, the integrality gap for vertex cover after $c_\zeta n$ rounds is at least $2 - \zeta$.*

PROOF: Let $\alpha = \zeta/6$ and $\varepsilon = \zeta/6$. Let $d = d(\alpha)$ where $d(\alpha)$ is as in Lemma 1. Let $\eta = \eta(\varepsilon, d)$ where $\eta(\varepsilon, d)$ is as in Lemma 2. Then by Lemma 1, there exists a $\delta, \gamma > 0$, $N \in N$ such that such that for $n \geq N$ there exists an $(\alpha, \delta, \gamma, \eta)$ graph with maximum degree at most d on n vertices. By Lemma 2, the vector $(1, 1/2+\varepsilon, \dots, 1/2+\varepsilon) \in N^{\Omega_{\varepsilon,\eta,\delta,\gamma,d}(n)}(VC(G))$ because $\eta = \eta(\varepsilon, d)$. This exhibits an integrality gap of $\frac{1-\alpha}{1/2+\varepsilon} = \frac{1-\zeta/6}{1/2+\zeta/6} \geq 2 - \zeta$. $\square$

Similarly, we have

Theorem 5 *For all $0 < \zeta < 1/50$, there is a constant $c_\zeta > 0$ such that, for all sufficiently large n, the integrality gap for max cut after $c_\zeta n$ rounds is at most $\frac{1}{2} + \zeta$.*

Lemma 1 is very similar to results already known in the literature (for example [ABLT06]) and so we only prove the additional properties that we require in the appendix. Most of the rest of the paper is dedicated to a proof of Lemma 2. Lemma 3 will follow via a relative simple "reduction" to Lemma 2.

4 Overview of the Proof

If D is a random variable ranging over vertex covers, then the solution $\mathbf{y}_D$ where $y_0 = 1$ and $y_i = \mathbf{Pr}[i \in D]$ is a convex combination of integral solutions, and so it survives an arbitrary number of rounds of LS. The protection matrix for $\mathbf{y}_D$ is the matrix $Y = Y_D$ such that $Y_{i,j} = \mathbf{Pr}[i \in D \wedge j \in D]$.

In trying to show that a given vector $\mathbf{y}$ survives several rounds of LS, it is a good intuition to think of $\mathbf{y}$ as being derived from a probability distribution over vertex covers (even if $\mathbf{y}$ is not a convex combination of integral solutions, and cannot be derived in this way) and, in constructing the protection matrix Y, to think of Y as being derived from the said distribution as above.

Note that for the above matrix, the vectors $\mathbf{z} = \mathbf{Y}_i/y_i$ and $\mathbf{w} = (\mathbf{Y}_0 - \mathbf{Y}_i)/(1 - y_i)$ correspond to conditional distributions with $z_j = \mathbf{Pr}[j \in D | i \in D]$ and $w_j = \mathbf{Pr}[j \in D | i \notin D]$. To show that

$\mathbf{y} \in N^k(VC(G))$, we must show that $\mathbf{z}, \mathbf{w} \in N^{k-1}(VC(G))$ for the vectors $\mathbf{z}$ and $\mathbf{w}$ corresponding to every i. The kth row in the protection matrices may now be interpreted as the distribution obtained by further conditioning on k. Intuitively, more rounds of LS correspond to further conditioning on other vertices which do not already have probability 0 or 1 in these conditional distributions. We often refer to vertices having probability 0/1 as being *fixed* in the distribution.

Since only r vertices can be conditioned upon in r rounds, we only need to create solutions that look "locally" like distributions over vertex covers for small sized subgraphs. Also, because the given graph has large girth, subgraphs of size $O(\log n)$ are trees. We thus start by expressing the vector $\mathbf{y} = (1, 1/2+\varepsilon, \ldots, 1/2+\varepsilon)$ as a probability distribution over vertex covers for a tree. This distribution we define has the property that conditioning on a vertex i only affects the vertices upto a constant distance ℓ from i. In fact, the effect of conditioning decreases exponentially with the distance from i and we explicitly truncate it at distance $\ell = O(\frac{1}{\varepsilon}\log(\frac{1}{\varepsilon}))$. The conditional distribution is referred to as a *splash* around i as it creates "ripples" (change in probabilities) which decrease with distance from i. Fernandez de la Vega and Kenyon [FdlVKM07, Section 5] describe essentially the same distribution of vertex covers over trees in their paper, suggesting its usefulness for proving integrality gaps for the vertex cover problem.

We start with the vector $(1, 1/2+\varepsilon, \ldots, 1/2+\varepsilon)$ for the given graph G. After one round of LS, each row i of the protection matrix is defined by changing only weights of vertices within distance a distance ℓ of vertex i according to a splash. Since it affects only a small subgraph, which is a tree rooted at i, the solution "looks" locally like a valid conditional distribution.

Now consider trying to extend this strategy to a second round. Say we want to show that the ith row of the protection matrix above survives another round. We thus need to create another protection matrix for this row. Each row of this new matrix corresponds to conditioning on some other vertex j. If i and j are at distance greater than 2ℓ, the weights (probabilities) of vertices within a distance ℓ from j are still $1/2+\varepsilon$. The conditional distribution can then be created by replacing these values according to a splash around j and leaving the weights of the other vertices as unchanged. If the distance between i and j is less than 2ℓ and k is a vertex within distance ℓ of either i or j, we modify the weight of k according to the probability that *both* i and j are in the vertex cover.

It would become, unfortunately, very complex to proceed for a large number of rounds with this kind of analysis, and it would appear that the girth of the graph would be a natural limit for the number of rounds for which we can extend this line of argument. (See indeed [ABLT06, Tou06].)

We note however that certain cases are simpler to handle. Suppose that we are given a vector $\mathbf{y}$ that is $1/2+\varepsilon$ everywhere except in a number of balls, all at distance at least 5ℓ from each other, in which the values of $\mathbf{y}$ are set according to splashes. Then the above ideas can be used to define a valid protection matrix. Unfortunately, this does not seem to help us in setting up an inductive argument, because the structure of the vector that we start from is not preserved in the rows of the protection matrix: we may end up with splash areas that are too close to each other, or with the more special structures that we get by conditioning on a vertex less than distance 2ℓ from the root of a splash.

Our idea, then, is to take such more complex vectors and express them as convex combinations of vectors that are $1/2+\varepsilon$ everywhere except in splash areas that are at distance at least 5ℓ from each other. We will refer to such solutions as *canonical* solutions. Since we are trying to show that the complex vector belongs to some convex cone, it suffices to show that each one of these simpler vectors is in the cone. Now we are back to the same type of vectors that we started from, and we can set up an inductive argument.

Our inductive argument proceeds as follows: we start from a solution $\mathbf{y}$ in a "canonical" form, that is, such that all vertices have value $1/2 + \varepsilon$ except for the vertices belonging to at most k splashes; furthermore, the roots of any two splashes are at distance at least 5ℓ from each other. We need to construct a protection matrix Y for this vector. To define the jth row $\mathbf{Y}_j$ of the protection matrix we reason as follows: if j is far (distance $> 2\ell$) from the roots of all the splashes in $\mathbf{y}$, then $\mathbf{Y}_j$ looks like $\mathbf{y}$, plus a new splash around j. If j is at distance $\leq 2\ell$ from a splash (and, necessarily, far from all the others) rooted at a vertex r, then we replace the splash rooted at r with a new splash which corresponds to our original distribution over trees conditioned on both r and j.

If $\mathbf{Y}_j$ happens to be a vector in canonical form, we are done, otherwise we need to express it as a convex combination of vectors in canonical form. There are two ways in which $\mathbf{Y}_j$ can fail to be canonical: j may be at distance more than 2ℓ but less than 5ℓ from the closest splash; in this case the new splash we create around j is too close to an already existing one. The other possibility is that j is at distance less than 2ℓ from an existing splash, in which case $\mathbf{Y}_j$ contains a "doubly-conditioned" splash which is not an allowed structure in a canonical solution.

Our idea is then to define a set S of "problematic vertices," namely, the vertices in the two close splashes, in the first case, or the vertices in the doubly-conditioned splash, in the second case. Then we prove that[4] that the restriction of $\mathbf{Y}$ to small (sub-linear) subset S of vertices can be expressed as a distribution of valid integral vertex covers over S. We would then like to use this fact to express $\mathbf{y}$ itself as a convex combination of solutions that are integral over S and agreeing with $\mathbf{y}$ outside S; if we could achieve this goal, we would have expressed $\mathbf{y}$ as a convex combination of vectors where the "problematic" coordinates of $\mathbf{y}$ are fixed, and the other coordinate are as nice as they were in $\mathbf{y}$.

Unfortunately, some complications arise. In order to express $\mathbf{y}$ as a convex combination $\sum_a \lambda_a \mathbf{y}_a$ of vectors such that each $\mathbf{y}_a$ is fixed in S, it is necessary that each $\mathbf{y}_a$ contains a splash around each of the newly fixed variables. The new splashes may themselves be at distance less than 5ℓ from each other, making the $\mathbf{y}_a$ not canonical. To remedy this problem, we define S (the set of vertices that will be fixed in the $\mathbf{y}_a$) via the following process: we initialize S to the initial set of problematic vertices, then we add all vertices that are at distance less than ℓ from S and that can be connected via a path of length $\leq 5\ell$ that does not pass through S, and so on. At the end of this process, we express $\mathbf{y}$ restricted to S as a convex combination of integral covers, and we extend each of these integral covers over S to a fractional solution over all vertices (by putting splashes around the vertices of S) and so express $\mathbf{y}$ as a convex combination of solutions that, now, are canonical.

The argument works provided that S is of sublinear size. A careful accounting guarantees that, if we want to show that our solution survives k rounds, we only need to consider instances where S is of size $O(k)$. Intuitively, this is due to the fact that each time we make S larger we discover a short path of length $t \leq 5\ell$ in the graph, and we add to the subgraph induced by S $t-1$ new vertices and t new edges. The subgraph induced by S can only include at most $|S|(1+\eta)$ edges, for some very small η, so it cannot happen that S grows too much at each step, because it is not

[4] Assuming some added conditions on the fractional solution $\mathbf{y}$, called *saturation*.

possible to consistently add more edges than vertices to the subgraph induced by S without causing a contradiction to the sparsity condition.

Since this ensures that it takes $\Omega(n)$ rounds before the set of fixed vertices grows to size γn, we can survive $\Omega(n)$ rounds.

5 Distributions of Vertex Covers in Trees

As a first (and useful) idealized model, suppose that our graph is a rooted tree. Consider the following distribution over valid vertex covers:

- The root belongs to the cover with probability $1/2 + \varepsilon$
- For every other vertex i, we make (independently) the following choice: if the parent of i does not belong to the vertex cover, then i is in the cover with probability one; if the parent of i is in the cover, then with probability $2\varepsilon/(\frac{1}{2} + \varepsilon)$ we include i in the cover, and with probability $1 - 2\varepsilon/(\frac{1}{2} + \varepsilon)$ we do not include i in the cover.

(The distribution is sampled by considering vertices in the order of a BFS, so that we make a decision about a vertex only after having made a decision about the parent.)

This is an instantiation of the Ising Model, about which much is known, but we will need only very elementary facts about it. The proofs of these facts are contained in the appendix.

A first observation is that each vertex has probability $1/2 + \varepsilon$ of being in the cover and $1/2 - \varepsilon$ of not being in the cover. The second observation is that, if we condition on the event that, say, the root is in the cover, then this condition affects very heavily the vertices that are close to root, but this effect decreases exponentially with the distance. In particular, for each vertex whose distance from the root is about $4\varepsilon^{-1} \cdot (\log \varepsilon^{-1})$, the probability of the vertex being in the cover condition on the root being in the cover is between $1/2 + \varepsilon - \varepsilon^4$ and $1/2 + \varepsilon + \varepsilon^4$, and the same is true conditioning on the root not being in the cover.

This second observation will show that reasoning about this distribution is useful to deal with graphs that are only *locally* like trees, that is, graphs of large girth. Before discussing this application, we slightly change the distribution so that, after a certain distance from the root, there is no effect (rather than a small effect) if we condition on the root being or not being in the cover. Hence the effect of conditioning on the root is explicitly cut-off after a certain distance.

In particular, consider the following two distributions which sample from the vertex covers of a tree rooted at a vertex i. The conditioning on the root only affects vertices upto a distance $\ell = \frac{8}{\varepsilon} \log \frac{1}{\varepsilon}$ of i.

Definition 6 *For $b \in \{0, 1\}$ we define a b-Splash around a vertex i as the distribution which modifies vertices upto a distance of 2ℓ as follows*

1. *$i = b$*
2. *For every vertex upto distance ℓ (and at distance greater than $\ell + 1$), we independently decide to include it with probability 1 if its parent is not in the vertex cover and with probability $2\varepsilon/(\frac{1}{2} + \varepsilon)$ if its parent is already in the vertex cover.*

3. *For u and v at distances ℓ, $\ell+1$ respectively, we include v with probability 1 if u is not in the vertex cover and with probability*

$$\frac{\mathbf{Pr}[u=1|i=b]-\left(\frac{1}{2}-\varepsilon\right)}{\mathbf{Pr}[u=1|i=b]}$$

otherwise.

Where $u=1$ denotes the event $u \in D$ for a random variable D (with distribution defined by the splash) ranging over the vertex covers of the graph.

For the above to be well defined, we need $\mathbf{Pr}[u=1|i=b] > 1/2-\varepsilon$ for a vertex u at distance ℓ from i. Claim 7 shows that in fact $\mathbf{Pr}[u=1|i=b] \in [1/2+\varepsilon-\varepsilon^4, 1/2+\varepsilon+\varepsilon^4]$ for u at distance greater than $\ell/2$ and hence the probability at distance ℓ is non-negative.

Claim 7 *Consider a b-Splash around any vertex i such that all vertices upto distance ℓ are labeled $\frac{1}{2}+\varepsilon$. Let j be a vertex such that $d(i,j) \le \ell$. Then,*

1. $\mathbf{Pr}[j=1|i=1, d(i,j)=k] = (1/2+\varepsilon)\left[1+(-1)^k\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{k+1}\right]$ *for* $0 \le k \le \ell$

   $\mathbf{Pr}[j=1|i=0, d(i,j)=k] = \mathbf{Pr}[j=1|i'=1, d(i',j)=k-1]$ *for* $1 \le k \le \ell$
2. $|\mathbf{Pr}[j=1|i=b, d(i,j)=\ell/2]-(1/2+\varepsilon)| \le \varepsilon^4$
3. $\mathbf{Pr}[j=1|i=1, d(i,j)=k] + \mathbf{Pr}[j=1|i=1, d(i,j)=k+1] \ge 1+4\varepsilon^2$ *for* $0 \le k \le \ell$

Note, in particular, that the probabilities are independent of i and j and depend only on their distance $d(i,j)$. Also, the difference of the probabilities from $1/2+\varepsilon$ decreases exponentially with distance. The following claim shows that the vertices outside a radius of ℓ from i are independent of whether or not i is in the cover.

Claim 8 *If we pick a 0-Splash with probability $1/2-\varepsilon$ and a 1-Splash with probability $1/2+\varepsilon$, then all vertices have probability $1/2+\varepsilon$. Furthermore, vertices at distance $\ell+1$ or more from i have weight $1/2+\varepsilon$ in the 0-Splash as well as 1-Splash around i.*

The vectors that appear in our argument may involve conditioning on a vertex i that has value different from $1/2+\varepsilon$ based on a splash distribution around a vertex r close to it. The following claims allow us to compute $\mathbf{Pr}[i=1, j=1|r=b]$, the probability of two vertices i,j being simultaneously present in a b-Splash at r, and also $\mathbf{Pr}[i=0, j=1|r=b]$, which is the probability that j is present and i is not. We defer the proofs to the appendix.

Claim 9 *Let $i=v_0, v_1, \ldots, v_{m-1}, v_m=j$ be the path to j, $m \le \ell$, and let u be the vertex on this path which is closest to r. Then*

1. $$\begin{aligned}\mathbf{Pr}[i=1, j=1|r=b] &= \mathbf{Pr}[u=1|r=b]\cdot\mathbf{Pr}[i=1|u=1]\cdot\mathbf{Pr}[j=1|u=1] \\ &+ \mathbf{Pr}[u=0|r=b]\cdot\mathbf{Pr}[i=1|u=0]\cdot\mathbf{Pr}[j=1|u=0]\end{aligned}$$

2. *If* $\mathbf{Pr}[u = 1|r = b] = 1/2 + \varepsilon$*, then* $\mathbf{Pr}[i = 1, j = 1|r = b] = (1/2 + \varepsilon)\,\mathbf{Pr}[j = 1|i = 1]$

The first part of the above claim states that once we condition on u, then i and j are independent. The second part states that if u is sufficiently far r, we can ignore r completely and just compute the probability of j as determined by a splash around i.

Claim 10 *Let i be a vertex and (j,k) be an edge in a b-Splash around r and let $b' \in \{0, 1\}$.*

$$\mathbf{Pr}[i = b', j = 1|r = b] + \mathbf{Pr}[i = b', k = 1|r = b] \quad \geq \quad \mathbf{Pr}[i = b'|r = b] \cdot (1 + 4\varepsilon^3)$$

The next claim allows us to treat vertices that are sufficiently far from each other as almost independent in the dirtibution conditioned on r.

Claim 11 *Let i and j be two vertices in a b-Splash around r, such that $d(i, j) \geq \ell$. Then*

$$\left|\mathbf{Pr}[i = b', j = 1|r = b] - \mathbf{Pr}[i = b'|r = b] \cdot \mathbf{Pr}[j = 1|r = b]\right| \quad \leq \quad 2\varepsilon^4$$

6 Distribution of Vertex Covers in Sparse Graphs

To reduce solutions with more complicated structure to simpler solutions, we will need to show that if we look at a sufficiently small subgraph of our original graph obtained in Lemma 1, then the more complicated solution can be expressed as a convex combination of 0/1 solutions.

The following result is proved in [ABLT06].

Lemma 12 ([ABLT06]) *Let $\eta \leq \frac{2\varepsilon}{3+10\varepsilon}$ and let $G = (V, E)$ be a graph such that*

1. *for each $S \subseteq V$, $G(S) = (V_{G(S)}, E_{G(S)})$, then $|E_{G(S)}| \leq (1 + \eta)|V_{G(S)}|$.*
2. *girth$(G) \geq \frac{1+2\varepsilon}{\varepsilon}$.*

Then there exists a distribution over vertex covers on G such that each vertex belongs to the vertex cover with probability $1/2 + \varepsilon$.

We will need a slight generalization. Instead of requiring the solution to have the value $1/2 + \varepsilon$ everywhere, we only require that the sum of the values on each edge should be at least $1 + 2\varepsilon$, if both of its endpoints are not already fixed.

Definition 13 *We call a fractional solution y for a graph G* ε-saturated *if for each edge (i, j) in graph G either:*

- *Both i and j are fixed and $y_i + y_j \geq 1$ or,*
- $y_i + y_j \geq 1 + 2\varepsilon$.

We now show that the under the conditions of the previous lemma, every ε-saturated solution can be written as a convex combination of vertex covers of the graph.

Lemma 14 *Let $\eta \leq \frac{2\varepsilon}{3+10\varepsilon}$ and let $G = (V, E)$ be a graph such that*

1. *for each $S \subseteq V$, $G(S) = (V_{G(S)}, E_{G(S)})$, then $|E_{G(S)}| \leq (1+\eta)|V_{G(S)}|$.*
2. *girth$(G) \geq \frac{1+2\varepsilon}{\varepsilon}$.*

and let $\mathbf{y}$ be an ε-saturated solution. Then there exists a distribution over vertex covers on G such that each vertex i belongs to the vertex cover with probability y_i.

Proof: For the graph G, we will create a set of feasible fractional solutions $\mathbf{y}(k) \in \{0, 1/2+\varepsilon, 1\}^{|V|}$ such that $\mathbf{y}$ is a convex combination of these vectors.

We partition V into V_0, $V_{1/2+\varepsilon}$, and V_1, as follows:

$$i \in \begin{cases} V_0 & y_i < 1/2+\varepsilon \\ V_{1/2+\varepsilon} & y_i = 1/2+\varepsilon \\ V_1 & y_i > 1/2+\varepsilon \end{cases}$$

We define $t(i)$ as follows:

$$t(i) = \begin{cases} 1 - \frac{y_i}{1/2+\varepsilon} & i \in V_0 \\ 1 & i \in V_{1/2+\varepsilon} \\ \frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon} & i \in V_1 \end{cases}$$

We can order the $t(i)$'s: $0 \leq t(i_1) \leq t(i_2) \leq \cdots \leq t(i_{|V|}) \leq 1$. For each $k : 1 \leq k \leq |V|$ we create the vector $\mathbf{y}(k)$ where

$$\mathbf{y}(k)_i = \begin{cases} 0 & i \in V_0 \text{ and } t(i) \leq t(i_k) \\ 1 & i \in V_1 \text{ and } t(i) \leq t(i_k) \\ 1/2+\varepsilon & \text{otherwise} \end{cases}$$

We claim the distribution where $\mathbf{y}(k)$ occurs with probability $t_{i_k} - t_{i_{k-1}}$ gives us $\mathbf{y}$.

If $i \in V_0$, then it will be 0 with probability t_i and $1/2+\varepsilon$ with probability $1-t_i = \frac{y_i}{1/2+\varepsilon}$. Therefore the probability that i is in the vertex cover is y_i. If $i \in V_1$, then it will be 1 with probability $t_i = \frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon}$ and $1/2+\varepsilon$ with probability $1-t_i = 1-\frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon}$. Therefore the probability that i is in the vertex cover is $\frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon} + (1/2+\varepsilon)(1-\frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon}) = y_i$. If $i \in V_{1/2+\varepsilon}$, then it is clear that the probability that i is in the vertex cover is $1/2+\varepsilon$.

Note that all the weights in each $\mathbf{y}(k)$ are 0, 1 or $1/2+\varepsilon$. It remains to show that in each of these $\mathbf{y}(k)$ any edge which contains one vertex fixed to 0 has the other vertex fixed to 1. First, note that all neighbors of vertices in V_0 are in V_1. It suffices to show that if i and j are adjacent, $i \in V_1$, $j \in V_0$, that $t(i) \geq t(j)$. However

$$\begin{aligned} t(i)-t(j) &= \frac{y_i-(1/2+\varepsilon)}{1/2-\varepsilon}-\frac{(1/2+\varepsilon)-y_j}{1/2+\varepsilon} \\ &= \frac{(y_i+y_j)/2+\varepsilon(y_i-y_j)-(1/2+\varepsilon)}{1/4-\varepsilon^2} \\ &\geq \frac{(1+2\varepsilon)/2+\varepsilon(y_i-y_j)-(1/2+\varepsilon)}{1/4-\varepsilon^2} \\ &= \frac{\varepsilon(y_i-y_j)}{1/4-\varepsilon^2} \geq 0 \end{aligned}$$

which concludes the proof of the lemma. □

7 The Main Lemma

We now define the type of solutions that will occur in our recursive argument.

Let $G=(V,E)$ be an $(\alpha,\delta,\gamma,\eta)$ graph with n vertices and degree at most d, as in the assumption of Lemma 2. We define the constant $C=\sum_{i=1}^{\ell+1} d^i$ as the maximum number of vertices within a distance ℓ from some vertex and $D=5\ell C$ as the maximum number of vertices within distance ℓ of all the vertices in a path of length 5ℓ. Choose $\eta=\frac{1}{3D}$. Note that η depends on only ε and d. Also, we assume that n is large enough that the girth of the graph is larger than various fixed constants throughout. We fix G for the rest of this section.

Let $R=\frac{\gamma n}{C+2D}$

Let $G_{|S}=(S,E_{|S})$ be the subgraph of G induced by $S\subseteq V$. For some set $S\subseteq V$, define $N_S(i)=\{j:$ there exists path of length ℓ from i to j using only edges in $E\backslash E_{|S}\}$.

Definition 15 *We say that a vector $y=(y_0,\ldots,y_n)$ is r-canonical if there exists a set $S\subseteq V$ such that:*

- $\forall j\in S$ $y_j\in\{0,1\}$ *and* $\mathbf{y}_{|S}$ *is a vertex cover of* $G_{|S}$
- *For every two vertices in S the shortest path between them that uses only vertices not in S has length $>5\ell$. (Therefore if $i,j\in S$, $i\neq j$, then $N_S(i)\cap N_S(j)=\emptyset$).*
- $$y_i=\begin{cases}\mathbf{Pr}[i=1|j=y_j] & \exists j\in S \text{ s.t. } i\in N_S(j)\\ 1/2+\varepsilon & o.w\end{cases}$$
- $|S|\leq rC+2rD$
- *Let $|S|=rC+kD$ ($k\leq 2r$) and $G_{|S}=(S,E_{|S})$ is the subgraph of G induced by S, then*
$$|E_{|S}|-|S|\geq k-r$$

We call a set S as in Definition 15 a witness.

Claim 16 *If $\mathbf{y}$ is an r-canonical vector then, $\mathbf{y} \in VC(G)$. Moreover, $\mathbf{y}$ is ε^2-saturated.*

PROOF: This follows from the fact all edges are either internal to S, internal to some $N_S(i)$, internal to $V \setminus \cup_{i\in S} N(i)$ or between some $N(i)$ and $V \setminus \cup_{i\in S} N(i)$. In the first case, it follows because $\mathbf{y}_{|S}$ is a valid vertex cover having only $0/1$ values. In the second because of the fact that a $N(i)$ is weighted according to a splash and Claim 7. In the third case, because the weights are all $1/2+\varepsilon$. The final case just concerns the vertices at distance ℓ and $\ell+1$ from the center of a splash and again follows from Claim 7. $\square$

Lemma 2 follows from the above claim, the following result and the fact that $(1, 1/2+\varepsilon, \ldots, 1/2+\varepsilon)$ is 0-canonical.

Lemma 17 *Let $\mathbf{y}$ be an r-canonical solution, and $r \le R$. Then $\mathbf{y}$ is in $N^{R-r}(VC(G))$.*

PROOF: We prove it by induction on $R-r$. By Claim 16, an R-canonical solution is feasible for $VC(G)$, and this gives the basis for the induction.

Let $\mathbf{y}$ be an r-canonical solution and let S be a witness to $\mathbf{y}$. We show that there is a protection matrix Y for $\mathbf{y}$ such that $(\mathbf{Y}_i)/y_i$ and $(\mathbf{Y}_0 - \mathbf{Y}_i)/(y_0 - y_i)$ are distributions over $(r+1)$-canonical vectors for $y_i \neq 0, y_0$. If $y_i = 0$, then we take $Y_i = 0$ which is in $N^k(VC(G))$ for all k and $\mathbf{Y}_0 - \mathbf{Y}_i = \mathbf{Y}_0$ which is r-canonical.

The protection matrix is defined as follows. (When we talk about distance between vertices, we mean distance via paths that do not go through any vertex in S.)

- $Y_{i,0} = Y_{0,i} = Y_{i,i} = y_i$.
- If i and j are at distance greater than ℓ from each other, then $Y_{i,j} = y_i \cdot y_j$
- If i is at distance greater than 2ℓ from the closest vertex in S, and j is at distance at most ℓ from i, then $Y_{i,j}$ is the probability that i and j both belongs to a vertex cover selected according to a splash distribution around $Y_{ij} = y_i \, \mathbf{Pr}[j=1|i=1]$
- If i is at distance at most 2ℓ from a vertex $r \in S$, and j is at distance at most ℓ from i, then Y_{ij} is the probability that i and j both belong to a vertex cover selected according to a b-Splash distribution around r i.e. $Y_{ij} = \mathbf{Pr}[i=1, j=1|r=b]$

Claim 18 *The matrix Y is symmetric.*

PROOF: If $d(i,j) > \ell$, clearly $Y_{ij} = Y_{ji}$. There remain three additional cases.

- First, if both i and j are at distance greater than 2ℓ from any vertex in S, then $y_i = y_j = 1/2+\varepsilon$ and also $\mathbf{Pr}[j=1|i=1] = \mathbf{Pr}[j=1|i=1]$ as it depends only on the distance by Claim 7, and hence $Y_{ij} = Y_{ji}$.
- Second, both i and j are at distance at most 2ℓ from any vertex in S. Both i and j cannot be close to two different vertices in S because then $d(i,j) \le \ell$ would imply a path of length at most 5ℓ between the two vertices which is not possible. Hence, in this case, $Y_{ij} = Y_{ji} = \mathbf{Pr}[i=1, j=1|r=b]$, where r is the vertex in S close to both i and j.

- Finally, if $d(i,r) \leq 2\ell$ for some $r \in S$ and $d(j,r) > 2\ell$ $\forall r \in S$, then the path from i to j cannot come closer than distance $\ell+1$ to r. If l is the vertex on this path closest to r, then we have $P_r^b(l) = 1/2+\varepsilon$ and by Claim 9, $\mathbf{Pr}[i=1, j=1|r=b] = (1/2+\varepsilon)\,\mathbf{Pr}[j=1|i=1] = y_i\,\mathbf{Pr}[j=1|i=1]$. Therefore, $Y_{ij} = \mathbf{Pr}[i=1, j=1|r=b] = y_i\,\mathbf{Pr}[j=1|i=1] = Y_{ji}$.

□

Let us fix a vertex i, and consider the vectors $\mathbf{z} := \mathbf{Y}_i/y_i$ and $\mathbf{w} := (\mathbf{Y}_i - \mathbf{Y}_0)/y_i$. We will show that they are (convex combinations of) $(r+1)$-canonical vectors. (If $y_i = 0$ we do not need to analyse $\mathbf{z}$, and if $y_i = 1$ we do not need to analyse $\mathbf{w}$.)

Note that $\mathbf{z}$ and $\mathbf{w}$ are same as $\mathbf{y}$ except for vertices that are within distance ℓ of i.

Lemma 19 *If $\mathbf{y}$ is an r-canonical solution and Y is the matrix as defined above, then $\forall 1 \leq i \leq n$, the solutions $\mathbf{z} := Y_i/y_i$ and $\mathbf{w} := (\mathbf{Y}_i - \mathbf{Y}_0)/y_i$ are ε^3-saturated*

PROOF: We first give the proof for $\mathbf{z}$. Note that for $d(i,j) > \ell$ $z_j = y_j$ and hence edges as distance greater than ℓ from i are ε^2 saturated because they were in $\mathbf{y}$ by Claim 16. If $d(i,r) > 2\ell$ $\forall r \in S$ then the distribution up to distance 2ℓ from i is same as a $1-Splash$, which is in fact ε^2-saturated by Claim 7 and Claim 8.

Let i be within distance 2ℓ of $r \in S$ and let (j,k) be an edge such that $d(i,j) \leq \ell$ or $d(i,k) \leq \ell$. If both j and k are within distance ℓ of i, then by Claim 10

$$\begin{aligned} Y_{ij} + Y_{ik} = &\ \mathbf{Pr}[i=1, j=1|r=b] + \mathbf{Pr}[i=1, k=1|r=b] \\ &\geq (1+4\varepsilon^3)\,\mathbf{Pr}[i=1|r=b] = (1+4\varepsilon^3)y_i \end{aligned}$$

and we are done. Finally, if $d(i,j) = \ell$ and $d(i,k) = \ell+1$, then we know by Claim 11 that $|\,\mathbf{Pr}[i=1, k=1|r=b] - \mathbf{Pr}[i=1|r=b]\,\mathbf{Pr}[k=1|r=b]| \leq 2\varepsilon^4$. This gives

$$\begin{aligned} Y_{ij} + Y_{ik} = &\ \mathbf{Pr}[i=1, j=1|r=b] + \mathbf{Pr}[i=1|r=b]\,\mathbf{Pr}[k=1|r=b] \\ \geq &\ \mathbf{Pr}[i=1, j=1|r=b] + \mathbf{Pr}[i=1, k=1|r=b] - 2\varepsilon^4 \\ &\geq (1+4\varepsilon^3)\,\mathbf{Pr}[i=1|r=b] - 2\varepsilon^4 \geq (1+3\varepsilon^3)y_i \end{aligned}$$

using the fact that $\mathbf{Pr}[i=1|r=b]$ is at least 2ε. We prove this for $\mathbf{w}$ similarly. □

We shall now express $\mathbf{z}$ and $\mathbf{w}$ as a convex combination of $(r+1)$-canonical vectors.

Claim 20 *If $i \in S$, or if $\forall r \in S$, $d(i,r) > 5\ell$, then $\mathbf{z}$ is $r+1$ canonical.*

PROOF: If $i \in S$, then $z_k = y_k$ (or $w_k = y_k$) for all $k \in V$ by construction of protection matrix. Because $\mathbf{y}$ is r-canonical $\mathbf{z}$ (or $\mathbf{w}$) is also and this thus also $(r+1)$-canonical.

If $\forall r \in S$, $d(i,r) > 5\ell$, then it is easily seen that $S \cup \{i\}$ is a witness to $\mathbf{z}$ and $\mathbf{w}$ being $(r+1)$-canonical. □

If neither of these cases is true, we treat only $\mathbf{z}$, because the same argument works for $\mathbf{w}$. We first define the subset of vertices which is fixed in these vectors.

Recall that for $i \in S$, $N_S(i) = \{j$: there exists path of length at most ℓ from i to j using only edges in $E \backslash E_{|S}\}$. In addition let $\partial N_S(i) = \{j : d(i,j) = \ell + 1$ in the graph $(V, E \backslash E_{|S})\}$. Also, let $N'_S(i) = N_S(i) \cup \partial N_S(i)$.

Then we make the following definition:

Definition 21 *For a fixed vertex i, we construct $F \subseteq V \setminus S$ as follows:*
Start with $F = N'_S(i)$. If there is a path P of length less that 5ℓ between any two vertices in $F \cup S$ that uses only edges in $V \setminus (F \cup S)$, then $F = F \cup P$. Also, if P intersects $N_S(j)$ for $j \in S$, then $F = F \cup P \cup (N'_S(j) \backslash \{j\})$.

Note that it follows from the above definition that for every $j \in S$, either $N_S(j) \cap F = \emptyset$ or $N'_S(j) \subseteq F$. Also if $\partial F = \{j \in F : j \textit{ has neighbors in } V \backslash (S \cup F)\}$, then $\forall j \in \partial F$, $z_j = 1/2 + \varepsilon$ (because for every intersecting $N_S(j')$, we also included $\partial N_S(j')$). We now bound the size of F.

Claim 22 $|F| \leq C + (2r + 2 - k)D$, *where* $|S| = rC + kD$.

Proof: Every path added in the construction of F has length at most 5ℓ. Also, each vertex in a path can be within distance ℓ of at most one $j \in S$. Thus, the number of vertices added due to a path is at most $5\ell C = D$. Thus, if p paths are added during the construction, then $|F| \leq C + pD$ since C is the size of the $N'_S(i)$, which we start with.

Since the paths are added incrementally, it suffices to show that adding $2r + 2 - k$ paths implies a contradiction. This would imply that $p \leq 2r + 2 - k$ and hence the claim. Let F' be F after addition of $2r + 2 - k$ paths. Then

$$\frac{|E_{|S \cup F'}|}{|S \cup F'|} = 1 + \frac{|E_{|S \cup F'}| - |S \cup F'|}{|S \cup F'|}$$

Note that $|E_{|S \cup F'}| - |S \cup F'| \geq (k - r) + (2r + 2 - k) - 1$, since $|E_{|S}| - |S| \geq (k - r)$ to begin with and addition of $N'_S(i)$, which is a tree adds one more vertex than edge (hence contributing -1), while the addition of each path adds one more edge than vertex. For any $j \in S$ including the region $N_S(j)$ intersected by the path includes a tree of which at least one vertex is already in F and can only contribute positively. This gives

$$\frac{|E_{|S \cup F'}|}{|S \cup F'|} \geq 1 + \frac{k - r + 2r - k + 1}{|S| + |F'|} \geq 1 + \frac{r + 1}{rC + kD + C + (2r + 2 - k)D} = 1 + \frac{1}{C + 2D} > 1 + \eta$$

since $\eta = \frac{1}{3D} < \frac{1}{C+2D}$. But this is a contradiction since $|S \cup F'| \leq \gamma n$ and hence $|E_{|S \cup F'}| \leq (1 + \eta)|S \cup F'|$. □

Now, because $r \leq R = \frac{\gamma n}{C+2D}$, $|S \cup F| \leq \gamma n$ and we employ Lemma 14 to $T = S \cup F$ using the fact that $\mathbf{z}$ is ε^3-saturated.

We obtain vertex covers on $S \cup F, T^1, \ldots, T^m$ such that $\lambda_1 T^1 + \ldots + \lambda_m T^m = \mathbf{z} \mid_T$ where $\sum_{l=1}^m \lambda_l = 1$. Note that the values for the vertices in S are 0/1 and are hence unchanged in all these solutions. To extend these solutions to fractional solutions over the whole graph, we look at each vertex j on the boundary of the set F and change the values of vertices upto a distance ℓ from it in $V \setminus (S \cup F)$ according to a splash around j. We first prove that all the vertices upto distance ℓ from the boundary of F have value $1/2 + \varepsilon$ in $\mathbf{z}$.

Claim 23 *For all $j \in F$, either*

- *all neighbors of j are in $S \cup F$, or*
- *For all $k \in N_{S\cup F}(j)$, $z_k = 1/2 + \varepsilon$*

PROOF: Assume not, then for some $j \in F$ which has some neighbor not in $S \cup F$, there exists $k \in N_{S\cup F}(j)$ such that $z_k \neq 1/2 + \varepsilon$. First, we show that it must be that $z_j = 1/2 + \varepsilon$. The only elements of $\mathbf{z}$ which do not have weight $1/2+\varepsilon$ are elements of $N_S(l)$ for $l \in F$ and $N_S(i)$. However, $N'_S(i) \subseteq F \cup S$ so no element of $N_S(i)$ has a neighbor outside of F. Similarly, if $j \in N_S(l)$, then because $j \in F$, it must be that $N'_S(l) \subseteq F \cup S$ and thus j has no neighbors outside $S \cup F$.

So, say that $k \neq j$, then $k \notin S \cup F$. But there exists a path P of length $\leq \ell$ which avoids $S \cup F$ from j to k. Because $\mathbf{y}$ is r-canonical, and $\mathbf{z}$ is the same as $\mathbf{y}$ except possibly at the vertices in $N_S(i)$, it must be that $k \in N_S(i)$ or $k \in N_S(j')$ for some $j' \in S$. But, it cannot be that $k \in N_S(i)$ because $N_S(i) \subseteq F$. Also if $k \in N_S(j')$ for some $j' \in S$, then there is a path from j to j' length at most 2ℓ and so either k must be in $S \cup F$ or $j = j'$. The former cannot be true by assumption. The later cannot be true because $j \in F$ which is disjoint from S. $\square$

Create $\mathbf{y}^{(l)}$ as follows.

$$y_k^{(l)} = \begin{cases} \mathbf{Pr}\left[k = 1 | j = y_j^{(l)}\right] & k \in N_{S\cup F}(j) \text{ for some } j \in F \\ y_k^{(l)} = z_i & \text{o. w.} \end{cases}$$

First note that this is well defined, because if any vertex were in $N_{S\cup F}(j)$ and $N_{S\cup F}(j')$ for $j, j' \in F$, $j \neq j'$, then there would be path between two vertices in F of length 2ℓ which does not go through $S \cup F$.

We wish to show that $\lambda_1 \mathbf{y}^{(1)}, \ldots, \lambda_m \mathbf{y}^{(m)} = \mathbf{z}$. Consider first some $k \in N_{S\cup F}(j)$ for some $j \in F$. First note that $\lambda_1 \mathbf{y}_j^{(1)} + \ldots + \lambda_m y_j^{(m)} = z_j$. By Claim 23 if $k \neq j$, then it must be that $z_j = z_k = 1/2 + \varepsilon$. Therefore by Claim 8

$$\lambda_1 y_k^{(1)} + \ldots + \lambda_m y_k^{(m)} = z_j \,\mathbf{Pr}[k = 1 | j = 1] + (1 - z_j)\,\mathbf{Pr}[k = 1 | j = 0] = 1/2 + \varepsilon = z_k$$

If $k \notin \cup_{j\in F} N_{S\cup F}(j)$, then $y_k^{(l)} = z_k$ for all k, and so $\lambda_1 y_k^{(1)}, \ldots, \lambda_m y_k^{(m)} = z_k$. We now must show that for each k, $\mathbf{y}^{(k)}$ is an $(r+1)$-canonical solution. We show that $T = S \cup F$ is a witness for $\mathbf{y}^{(k)}$. Since the solution $T^{(k)}$ given by Lemma 14 is a vertex cover $\mathbf{y}_{|T}^{(k)} = T^{(k)}$ is a vertex cover for T. Also, by construction of F, there is no path of length less than 5ℓ between any vertices of $S \cup F$ using only vertices outside $S \cup F$. By Claim 22 $|T| = |S| + |F| \leq rC + kD + C + (2r + 2 - k)D = (r+1)C + 2(r+1)D$. If the number of paths added in constructing F is p, then $|T| \leq (r+1)C + (k+p)D$. Also, as argued in Claim 22, $|E_{|S\cup F}| - |S \cup F| \geq (k - r) + p - 1 = (k + p) - (r + 1)$.

Finally, we need to show that $y_j^{(k)} = \mathbf{Pr}[j = 1 | j' = y_{j'}]$ if $j \in N_{S\cup F}(j')$ and $1/2 + \varepsilon$ otherwise. Let $y_j^{(k)} \neq 1/2 + \varepsilon$. Then either $j \in N_{S\cup F}(j')$ for some $j' \in F$ (since these vertices were set according to a splash pattern while creating $\mathbf{y}^{(k)}$) and we are done, or $z_k \neq 1/2 + \varepsilon$. However, $\mathbf{z} = \mathbf{Y}_i / y_i$ differs from $\mathbf{y}$ only in $N_S(i)$. Therefore, $z_k \neq 1/2 + \varepsilon$ in turn implies $j \in N_S(i)$ and hence $j \in F$, or $y_j \neq 1/2 + \varepsilon$. To finish off, we note that $y_j \neq 1/2 + \varepsilon$ would mean $j \in N_S(j')$ for some $j' \in S$ (by assumption on S). Since $N_S(j')$ is either contained in or disjoint with F, we must have $j \in S \cup F$ or $j \in N_{S\cup F}(j')$ respectively.

Since each $\mathbf{y}^{(k)}$ is an $(r+1)$-canonical solution, by our inductive hypothesis $\forall 1 \le k \le m$ $\mathbf{y}^{(k)} \in N^{R-r-1}(VC(G))$ and hence $\mathbf{z} \in N^{R-r-1}(VC(G))$. Using a similar argument for show $\mathbf{w}$, we get that $\mathbf{y} \in N^{R-r}(VC(G))$. This completes the proof of Lemma 17. $\square$

8 Lower bounds for MAX-CUT

Let $G = (V, E)$ be a graph with n vertices and m edges. We prove a $1/2 + \zeta$ integrality gap for $\Omega(n)$ rounds of LS on MAX-CUT.

The solutions we define for MAX-CUT are simple extensions of vertex cover solutions. For a vector $\mathbf{y} \in \mathbb{R}^{n+1}$, we define an extension $Ext(\mathbf{y})$ as the vector $\mathbf{u} \in \mathbb{R}^{n+m+1}$ such that, $u_i = y_i \;\; \forall 0 \le i \le n$ and $u_e = 2y_0 - y_i - y_j$ for $e = (i,j) \in E$. Also, we define $Res(\mathbf{u})$ as the inverse operation i.e. the projection of the first $n+1$ coordinates of $\mathbf{u}$. It is easy to verify that if $\mathbf{y} \in VC(G)$ then $Ext(\mathbf{y}) \in MC(G)$. Notice that with $R = \frac{\gamma n}{C+D}$ as defined in the previous section, it is sufficient to prove the following

Lemma 24 *If $\mathbf{y} \in \mathbb{R}^{n+1}$ is a 2r-canonical solution for VC(G), then Ext($\mathbf{y}$) $\in N^{R/2-r}(MC(G))$.*

The integrality gap follows because $\mathbf{y} = (1, 1/2+\varepsilon, \dots, 1/2+\varepsilon)$ is 0-canonical and for $\mathbf{u} = Ext(\mathbf{y})$, $\sum_{e\in E} u_e = (1-2\varepsilon)m$.

PROOF: We proceed by induction on $R/2 - r$. The base case follows because if $\mathbf{y}$ is an R-canonical solution, then $\mathbf{y} \in VC(G)$ which implies $Ext(\mathbf{y}) \in MC(G) = N^0(MC(G))$. For the inductive step, let $\mathbf{y}$ be an $2r$-canonical solution and let $\mathbf{u} = Ext(\mathbf{y})$. We create a protection matrix U, such that $\forall 1 \le i \le n$ and $\forall e \in E$, $Res(\mathbf{U}_i), Res(\mathbf{U}_e), Res(\mathbf{U}_0 - \mathbf{U}_i)$ and $Res(\mathbf{U}_0 - \mathbf{U}_e)$ can be expressed as convex combinations of $(2r+2)$-canonical solutions. This suffices because for a vector $\mathbf{u}$ if $Res(\mathbf{u}) = \lambda_1 \mathbf{u}^{(1)} + \ldots + \lambda_m \mathbf{u}^{(m)}$ then $\mathbf{u} = Ext(\lambda_1 \mathbf{u}^{(1)}) + \ldots + Ext(\lambda_m \mathbf{u}^{(m)})$, since the coordinates of $Ext(\mathbf{v})$ are affine functions of the coordinates of $\mathbf{v}$.

Let Y be the protection matrix of a $2r$-canonical solution as defined in the previous section. We define the matrix U as

$$\begin{aligned} \mathbf{U}_i &= Ext(\mathbf{Y}_i) & \forall 0 \le i \le n \\ \mathbf{U}_e &= Ext(2\mathbf{Y}_0 - (\mathbf{Y}_i + \mathbf{Y}_j)) & \forall e = (i,j) \in E \end{aligned}$$

We can write out the entries of U as follows, showing that it is symmetric.

$$\begin{aligned} U_{i,j} &= Y_{ij} & 0 \le i,j \le n \\ U_{i,e} = U_{e,i} &= 2Y_{i0} - Y_{ij} - Y_{ik} & 0 \le i \le n,\; e = (j,k) \in E \\ U_{e_1,e_2} &= 4Y_{00} - 2(Y_{i0} + Y_{j0} + Y_{k0} + Y_{l0}) + (Y_{ik} + Y_{jk} + Y_{il} + Y_{jl}) & e_1 = (i,j), e_2 = (k,l) \in E \end{aligned}$$

Note that for $i \in V$ and $e = (j,k) \in E$, $Res(\mathbf{U}_i) = \mathbf{Y}_i, Res(\mathbf{U}_0 - \mathbf{U}_i) = \mathbf{Y}_0 - \mathbf{Y}_i$ and $Res(\mathbf{U}_e) = \mathbf{Y}_0 - \mathbf{Y}_j + \mathbf{Y}_0 - \mathbf{Y}_k$, which are convex combinations of $(2r+1)$-canonical solutions as proved in the previous section. It only remains to tackle $Res(\mathbf{U}_0 - \mathbf{U}_e) = \mathbf{Y}_j + \mathbf{Y}_k - \mathbf{Y}_0$. We first prove that it is ε^3-saturated.

Claim 25 *If Y is the protection matrix of a 2r-canonical solution and $(i,j),(u,v)$ are two edges, then*

$$\frac{(\mathbf{Y}_i+\mathbf{Y}_j-\mathbf{Y}_0)_u}{y_i+y_j-y_0}+\frac{(\mathbf{Y}_i+\mathbf{Y}_j-\mathbf{Y}_0)_v}{y_i+y_j-y_0}\geq 1+4\varepsilon^3$$

PROOF: Without loss of generality, we can assume that j and u are the closer endpoints of the edges (i,j) and (u,v). We first handle the case when $d(j,u)>\ell$. Then $Y_{iu}=y_iy_u, Y_{iv}=y_iy_v, Y_{ju}=y_jy_u$ and $Y_{jv}=y_jy_v$. Hence, the LHS is y_u+y_v, which is greater than $1+2\varepsilon^2$ since a $2r$-canonical solution is ε^2 saturated.

When $d(j,u)\leq\ell$, all the four vertices are within distance $\ell+2$ of each other. Now, in any subgraph H of diameter 3ℓ, we may think of the restriction of $\mathbf{y}$ to H as the probabilities of the vertices being present in a distribution over vertex covers of H. Notice that if $\mathbf{y}$ is a $2r$-canonical solution, H may contain vertices close to (within distance ℓ of) at most one fixed vertex. In case there is such a vertex r, $\forall i\in H$ $y_i=\mathbf{Pr}[i=1|r=1]$. If there is no such vertex, all vertices in H have $y_i=1/2+\varepsilon$ and we can these as probabilities for a distribution which chooses a 1-splash with probability $1/2+\varepsilon$ and 0-splash with probability $1/2-\varepsilon$ around any arbitrary vertex in H (Claim 8). Also, we can interpret Y_{pq} as $\mathbf{Pr}[p=1,q=1]$ for the same distribution as above.

Consider the distribution over the subgraph within a radius $\ell+2$ from i. We first note that since $(\mathbf{Y}_0-\mathbf{Y}_i)/(1-y_i)$ is a valid vertex cover solution and $(\mathbf{Y}_0-\mathbf{Y}_i)_i=0$, $(\mathbf{Y}_0-\mathbf{Y}_i)_j/(1-y_i)=1$ which gives $y_i+y_j-1=Y_{ij}$. Using this and the fact that $\mathbf{Pr}[(i=1)\vee(j=1)|u=1]=1$, we have

$$\begin{aligned}\frac{(\mathbf{Y}_i+\mathbf{Y}_j-\mathbf{Y}_0)_u}{y_i+y_j-y_0}&=\frac{y_u(\mathbf{Pr}[i=1|u=1]+\mathbf{Pr}[j=1|u=1]-1)}{\mathbf{Pr}[i=1,j=1]}\\&=\frac{y_u\,\mathbf{Pr}[(i=1)\wedge(j=1)|u=1]}{\mathbf{Pr}[i=1,j=1]}\\&=\mathbf{Pr}[u=1|i=1,j=1]\end{aligned}$$

Therefore, we get

$$\begin{aligned}\frac{(\mathbf{Y}_i+\mathbf{Y}_j-\mathbf{Y}_0)_u}{y_i+y_j-y_0}+\frac{(\mathbf{Y}_i+\mathbf{Y}_j-\mathbf{Y}_0)_v}{y_i+y_j-y_0}-1&=\mathbf{Pr}[u=1|i=1,j=1]+\mathbf{Pr}[v=1|i=1,j=1]-1\\&=\mathbf{Pr}[(u=1)\wedge(v=1)|i=1,j=1]\\&=\mathbf{Pr}[(u=1)\wedge(v=1)|j=1]\end{aligned}$$

The last equality following from the fact that it is sufficient to condition on the closer of the two vertices i and j. Also,

$$\begin{aligned}\mathbf{Pr}[(u=1)\wedge(v=1)|j=1]=&\ \ \mathbf{Pr}[u=1|j=1]+\mathbf{Pr}[v=1|j=1]-1\\=&\ \ \frac{Y_{uj}}{y_j}+\frac{Y_{vj}}{y_j}-1\\\geq&\ \ 4\varepsilon^3\qquad\text{(by Lemma 19)}\end{aligned}$$

□

We now want to express $\mathbf{w} = (\mathbf{Y}_i + \mathbf{Y}_j - \mathbf{Y}_0)/(y_i + y_j - 1)$ as a convex combination of $(2r+2)$-canonical solutions. Let S be the witness to $\mathbf{y}$ being $2r$-canonical. We now find a set $T \supseteq S$ such that $\mathbf{w}$ is a convex combination of solutions $\mathbf{w}^{(1)}, \ldots, \mathbf{w}^{(m)}$ which take 0/1 values over T and which are $(2r+2)$-canonical, with T being the witness. There are two cases:

Case 1: *$i \notin S$ and $\exists r \in S$ s.t. $d(i,r) \leq 5\ell$* (with $d(i,r)$ being length of the shortest path not passing through S)
By the proof in the previous section, we know that the vector $z = \mathbf{Y}_i/y_i$ is a convex combination of $(2r+1)$-canonical solutions with a set S_1 being the witness for all of them. Also, $j \in S_1$ as it includes every vertex within distance ℓ of i. We take $T = S_1$.

Case 2: *$i \in S$ or $d(i,r) > 5\ell$ $\forall r \in S$*
In this case $\mathbf{z} = \mathbf{Y}_i/y_i$ is $(2r+1)$-canonical with $S \cup \{i\}$ the witness. We now look at the protection matrix Z for $\mathbf{z}$ and consider the vector $\mathbf{z}' = \mathbf{Z_j}/z_j$. This is a convex combination of $(2r+2)$-canonical solutions having a common witness S_2 which contains $S \cup \{i\}$. Take $T = S_2$.

In both cases $|T| \leq (2r+2)C + (4r+4)D$. We now employ Lemma 14 to T to obtain vertex covers $T^1, \ldots, T^m$ on T such that $\lambda_1 T^1 + \ldots + \lambda_m T^m = \mathbf{w}_{|T}$ with $\sum_{l=1}^m \lambda_l = 1$. We can extend them to create $(2r+2)$-canonical solutions $\mathbf{w}^{(1)}, \ldots, \mathbf{w}^{(m)}$ as in the previous section. By the arguments in the previous section, all these have T as the witness. This completes the proof. $\square$

Acknowledgements

We are grateful to James Lee for his suggestion to consider integrality gaps for Max Cut.

A Proof of Lemma 1

Lemma A.1 *For every $0 < \alpha < 1/125$, $\eta > 0$, there exists a $d = d(\alpha) \in \mathbb{N}$, δ, $\gamma > 0$, and $N \in \mathbb{N}$ such that for $n \geq N$ there exists an $(\alpha, \delta, \gamma, \eta)$ graph with max cut less than $\frac{1}{2}|E|(1+\alpha)$ and maximum degree at most d on n vertices. Here $d(\alpha)$ is an explicit function that depends only on α.*

We use the following lemma from [ABLT06]

Lemma A.2 *For every $1 < \alpha < 1/250$, $\eta > 0$, there exists a $\delta, \gamma > 0$ such that a random graph from the $G_{n,p}$ distribution where $p = \alpha^{-2}/n$ has the following properties with probability $1 - o(n)$:*

- *after $O(\sqrt{n})$ edges are removed, the girth is $\delta \log n$.*
- *the minimum vertex cover contains at least $(1-\alpha)n$ vertices*
- *every induced subgraph on a subset S of at most γn vertices has at most $(1+\eta)|S|$ edges.*

Proof:[of Lemma A.1] Given $\alpha, \eta > 0$, set $\alpha' = \alpha/2$. Use Lemma A.2 with inputs α', η to randomly pick a graph on n vertices. Set $p = (\alpha')^{-2}/n$ as in Lemma A.2. Now, with high probability, we can remove set of edges R to obtain a $(\alpha/2, \delta, \gamma, \eta)$-graph on n vertices. Do not yet remove edges.

Also, it is well known that w.h.p. the max-cut in a random $G_{n,p}$ has size less than $\frac{1}{2}|E|(1+1/\sqrt{d})$, where d is the average degree. The average degree of a vertex in this model is $\lambda = pn = 4\alpha^{-2}$. Hence the size of the max-cut is at most $\frac{1}{2}|E|(1+\alpha/2)$. The probability that some fixed vertex v_0 has degree greater than 2λ is less than $\exp(-\lambda/3)$ by a Chernoff bound. So by Markov's inequality the probability that more than $\exp(-\lambda/6)n$ vertices have degree greater than 2λ is at most $\exp(-\lambda/6) \leq exp(-10000)$.

If this is the case, then first remove the edge set R. By removing edges we could only decrease the maximum degree. Then simply remove all vertices with degree more than 2λ from the graph and any other subset to obtain a graph G' with $n(1-exp(-d/6))$ vertices. Now, it is easy to check that G' is a $(\alpha,\delta,\gamma,\eta)$-graph with maximum degree at most $d(\alpha) = 2\lambda = 8/\alpha^2$. Removing the edges and vertices changes the max cut to $\frac{1}{2}|E|(1+\alpha/2+o(1)) < \frac{1}{2}|E|(1+\alpha)$. $\Box$

B Proofs of claims about splashes

We use the following notation for the proofs in this appendix. We denote $\mathbf{Pr}[i=1|r=b]$ and $\mathbf{Pr}[i=1,j=1|r=b]$ by $P_r^b(i)$ and $P_r^b(i,j)$ respectively. $\mathbf{Pr}[i=0|r=b]$ and $\mathbf{Pr}[i=0,j=1|r=b]$ are expressed as $1-P_r^b(i)$ and $P_r^b(j)-P_r^b(i,j)$ respectively. Also, in cases where $\mathbf{Pr}[j=1|i=b]$ depends only on $d(i,j)$, we denote it by $Q^b(d(i,j))$.

Claim B.1 *Consider a b-Splash around a vertex i such that all vertices upto distance ℓ are labeled $\frac{1}{2}+\varepsilon$. Then,*

1. $Q^1(k) = (1/2+\varepsilon)\left[1+(-1)^k\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{k+1}\right]$ *for* $0 \leq k \leq \ell$

   $Q^0(0) = 0$ *and* $Q^0(k) = Q^1(k-1)$ *for* $1 \leq k \leq \ell$
2. $\left|Q^0(\ell/2) - (1/2+\varepsilon)\right| \leq \varepsilon^4$
3. $\forall 0 \leq k \leq \ell$, $Q^1(k) + Q^1(k+1) \geq 1 + 4\varepsilon^2$

Proof: We prove the formula for $Q^1(k)$ by induction. For $k=0$,

$$(1/2+\varepsilon)\left[1+(-1)^k\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{k+1}\right] = (1/2+\varepsilon)\left[\frac{1}{1/2+\varepsilon}\right] = 1 = Q^1(0)$$

Assuming the correctness of the formula for $k=n$, we start with the recurrence

$$Q^1(n+1) = (1-Q^1(n)) + \left(\frac{2\varepsilon}{1/2+\varepsilon}\right)Q^1(n) = 1 - \left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)Q^1(n)$$

since the vertex at distance n (in the same path) might not be present with probability $1-Q^1(n)$ in which case the one at distance $n+1$ is present with probability 1, and it is present with probability $Q^1(n)$ in which case the one at distance $n+1$ is included with probability $\left(\frac{2\varepsilon}{1/2+\varepsilon}\right)$. Therefore, we have

$$Q^1(n+1) = 1 - \left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)(1/2+\varepsilon)\left[1+(-1)^n\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{n+1}\right]$$
$$= 1-(1/2-\varepsilon)+(-1)^{n+1}(1/2+\varepsilon)\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{n+2} \quad = \quad (1/2+\varepsilon)\left[1+(-1)^{n+1}\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{n+2}\right]$$

Also note that if i is labeled 0, then all its neighbors must be set to 1. Hence $Q^0(0) = 0$ and $Q^0(1) = 1$. The rest of the induction works exactly as above.
Note that

$$\left|Q^0(\ell/2) - (1/2+\varepsilon)\right| = (1/2+\varepsilon)\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{\ell/2} < (1-2\varepsilon)^{\ell/2} = (1-2\varepsilon)^{(\frac{4}{\varepsilon}\log\frac{1}{\varepsilon})} \leq \varepsilon^4$$

Finally for $0 \leq k < \ell$,

$$Q^1(k)+Q^1(k+1) = (1/2+\varepsilon)\left[2+(-1)^k\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{k+1}\left(1-\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)\right]$$
$$= (1/2+\varepsilon)\left[2+(-1)^k\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^{k+1}\left(\frac{2\varepsilon}{1/2+\varepsilon}\right)\right]$$
$$\geq (1/2+\varepsilon)\left[2-\left(\frac{2\varepsilon}{1/2+\varepsilon}\right)\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^2\right] = 1+2\varepsilon-2\varepsilon\left(\frac{1/2-\varepsilon}{1/2+\varepsilon}\right)^2 \geq 1+4\varepsilon^2$$

The claim for $k=\ell$ follows from part 2 and the fact that $Q^1(d) = 1/2+\varepsilon$ for $d > \ell$. □

Claim B.2 *If we pick a 0-Splash with probability* $1/2-\varepsilon$ *and a 1-Splash with probability* $1/2+\varepsilon$, *then all vertices have probability* $1/2+\varepsilon$. *Furthermore, vertices at distance* $\ell+1$ *or more from* i *have weight* $1/2+\varepsilon$ *in the 0-Splash as well as 1-Splash around* i.

Proof: We prove it by induction on the length of the path from i to j. Let $P_i(j) = (1/2-\varepsilon)P_i^0(j)+(1/2+\varepsilon)P_i^1(j)$. The base case, when the path is of length 0 is clear. If the path between i and j is $i = v_0, v_1, \ldots, v_{m-1}, v_m = j$, then there are two cases. In the first case v_{m-1} and v_m are both within distance ℓ of i. Then

$$P_i(j) = 1-(1-\frac{2\varepsilon}{1/2+\varepsilon})P_i(v_{m-1})$$

because v_m is only excluded with probability $\frac{2\varepsilon}{1/2+\varepsilon}$ when v_{m-1} is present and this event is independent of whether or not each vertex $i = v_0, v_1, \ldots, v_{m-1}$ is included in the cover. By induction, $P_i(v_{m-1}) = 1/2+\varepsilon$, and so $1-(1-\frac{2\varepsilon}{1/2+\varepsilon})P_i(v_{m-1}) = 1/2+\varepsilon$.
In the second case v_{m-1} is at distance ℓ. However,

$$P_i^b(j) = 1-(1-\frac{P_i^b(v_{m-1})-(1/2-\varepsilon)}{P_i^b(v_{m-1})})P_i^b(v_{m-1}) = 1/2+\varepsilon$$

because the probability v_{m-1} is included in a b-Splash is $P_i^b(v_{m-1})$ and the probability of including v_m when v_{m-1} is present is $\frac{P_i^b(v_{m-1})-(1/2-\varepsilon)}{P_i^b(v_{m-1})}$. □

Claim B.3 *Let $i = v_0, v_1, \ldots, v_{m-1}, v_m = j$ be the path to j, $m \le \ell$, and let u be the vertex on this path which is closest to r. Then*

1. $P_r^b(i,j) = P_r^b(u) \cdot P_u^1(i) P_u^1(j) + [1 - P_r^b(u)] \cdot P_u^0(i) P_u^0(j)$
2. *If $P_r^b(u) = 1/2 + \varepsilon$, then $P_r^b(i,j) = (1/2 + \varepsilon) P_i^1(j)$*

PROOF:

1. Let E be the event that both i and j are in a vertex cover and $r = b$. Then $P_r^b(i,j) = \Pr[E \mid r = b]$. We can also condition on whether u is in the vertex cover.

$$\begin{aligned} P_r^b(i,j) &= \Pr[u \in VC \mid r = b] \cdot \Pr[E \mid r = b \text{ and } u \in VC] \\ &+ \Pr[u \notin VC \mid r = b] \cdot \Pr[E \mid r = b \text{ and } u \notin VC] \end{aligned}$$

But $\Pr[E \mid r = b \text{ and } u \in VC] = \Pr[E \mid u \in VC]$. Because given that u is in or out of the vertex cover, we can determine if i and j are in the vertex cover by following the edges from u to each of them. But this information is independent of whether r is in the vertex cover. For the same reason $\Pr[E \mid r = b \text{ and } u \in VC] = \Pr[E \mid u \in VC]$. Therefore

$$P_r^b(i,j) = P_r^b(u) \cdot P_u^1(i) P_u^1(j) + [1 - P_r^b(u)] \dot{P}_u^0(i) P_u^0(j)$$

as claimed.

2. The probability that i and j are in a vertex cover (assume r is not yet fixed) is just $(1/2 + \varepsilon) P_i^1(j)$. Now, we can just condition on l, and rewrite this as

$$\Pr[u \in VC] \cdot P_u^1(i,j) + \Pr[u \notin VC] \cdot P_u^0(i,j)$$

We can also not condition on $r = b$ because once l is fixed, that does not affect anything, and in addition, $\Pr[u \in VC] = 1/2 + \varepsilon = P_r^b(u)$. So this becomes

$$P_r^b(u) \cdot P_u^1(i,j) + [1 - P_r^b(u)] \cdot P_u^0(i,j)$$

Finally, if we note that $P_u^b(i,j) = P_u^b(i) P_u^b(j)$, we see that we get

$$P_r^b(l) \cdot P_u^1(i) P_u^1(j) + [1 - P_r^b(u)] \cdot P_u^0(i) P_u^0(j)$$

which by 1) is simply $P_r^b(i,j)$ as claimed.

□

Claim B.4 *Let i be a vertex and (j,k) be an edge in a b-Splash around r. Then if j and k are not already fixed*

$$P_r^b(i,j) + P_r^b(i,k) \ge P_r^b(i)(1 + 4\varepsilon^3)$$

and

$$[P_r^b(j) - P_r^b(i,j)] + [P_r^b(k) - P_r^b(i,k)] \ge (1 - P_r^b(i))(1 + 4\varepsilon^3)$$

PROOF: We consider separately the cases when (j,k) lies on or outside the path between r and i.

***Case 1*: (j,k) lies outside the path connecting r and i**

Without loss of generality, let j be the vertex closer to the path from r to i. Let u be the vertex in the path closest to j. Then by Claim B.3

$$\begin{aligned} P_r^b(i,j) &= P_r^b(u)\cdot P_u^1(i)P_u^1(j) + [1-P_r^b(u)]\cdot P_u^0(i)P_u^0(j) \\ P_r^b(i,k) &= P_r^b(u)\cdot P_u^1(i)P_u^1(k) + [1-P_r^b(u)]\cdot P_u^0(i)P_u^0(k) \end{aligned}$$

Therefore,

$$P_r^b(i,j)+P_r^b(i,k) = P_r^b(u)P_l^1(i)\cdot\left[P_u^1(j)+P_u^1(k)\right] + [1-P_r^b(u)]P_u^0(i)\cdot\left[P_u^0(j)+P_u^0(k)\right]$$

Also by Claim B.1 we know that $P_u^b(j)+P_u^b(k)\geq 1+4\varepsilon^2$, if j and k are not already fixed, which gives

$$P_r^b(i,j)+P_r^b(i,k) \geq \left[P_r^b(u)P_u^1(i)+[1-P_r^b(u)]P_u^0(i)\right](1+4\varepsilon^2) = P_r^b(u)(1+4\varepsilon^2)$$

***Case 2*: (j,k) lies on the path connecting r and i**

Let j be the vertex closer to r. Also, let $\alpha = P_r^b(j)$ and $\beta = P_j^1(i)$. Then,

$$\begin{aligned} P_r^b(i,j) &= P_r^b(j)P_j^1(i) = \alpha\beta \\ P_r^b(i,k) &= P_r^b(k)P_k^1(i) = \left[1-\alpha+\frac{2\varepsilon}{1/2+\varepsilon}\alpha\right]\left[(1-\beta)\frac{1/2+\varepsilon}{1/2-\varepsilon}\right] \\ &= (1-\alpha)(1-\beta)\left(\frac{1/2+\varepsilon}{1/2-\varepsilon}\right)+\alpha(1-\beta)\left(\frac{2\varepsilon}{1/2-\varepsilon}\right) \end{aligned}$$

where the second equation follows from the recurrence $Q^1(n+1) = (1-Q^1(n)) + \left(\frac{2\varepsilon}{1/2+\varepsilon}\right)Q^1(n)$ used in Claim B.1. Also,

$$\begin{aligned} P_r^b(i) &= P_r^b(j)P_j^1(i) + (1-P_r^b(j))P_j^0(i) = P_r^b(j)P_j^1(i) + (1-P_r^b(j))P_k^1(i) \\ &= \alpha\beta + (1-\alpha)(1-\beta)\left(\frac{1/2+\varepsilon}{1/2-\varepsilon}\right) \end{aligned}$$

This gives

$$\frac{P_r^b(i,j)+P_r^b(i,j)}{P_r^b(i)} = 1+\frac{\alpha(1-\beta)\left(\frac{2\varepsilon}{1/2-\varepsilon}\right)}{\alpha\beta+(1-\alpha)(1-\beta)\left(\frac{1/2+\varepsilon}{1/2-\varepsilon}\right)} \geq 1+4\varepsilon^3$$

since $\alpha, (1-\beta) > 2\varepsilon$ (all probabilities in a splash are at least 2ε, unless one is 0 and the other is 1, but then both are fixed).

The proof of the second statement follows similarly. □

Claim B.5 *Let i and j be two vertices in a b-Splash around r, such that $d(i,j) \geq \ell$. Then*

$$\left| P_r^b(i,j) - P_r^b(i) P_r^b(j) \right| \leq 2\varepsilon^4$$

and

$$\left| [P_r^b(j) - P_r^b(i,j)] - (1 - P_r^b(i)) P_r^b(j) \right| \leq 2\varepsilon^4$$

Proof: Let u be the vertex closest to r on the path from i to j. Without loss of generality, assume that $d(i,u) \geq \ell/2$. Then

$$\begin{aligned} \left| P_r^b(i,j) - P_r^b(i) P_r^b(j) \right| &= \left| P_r^b(u) \cdot P_u^1(i) P_u^1(j) + [1 - P_r^b(u)] \cdot P_u^0(i) P_u^0(j) - P_r^b(i) P_r^b(j) \right| \\ &\leq \left| (1/2 + \varepsilon) \left[P_r^b(u) \cdot P_u^1(j) + [1 - P_r^b(u)] \cdot P_u^0(j) \right] - P_r^b(i) P_r^b(j) \right| + \varepsilon^4 \\ &= \left| (1/2 + \varepsilon) P_r^b(j) - P_r^b(i) P_r^b(j) \right| + \varepsilon^4 \leq 2\varepsilon^4 \end{aligned}$$

where the two inequalities follow from the fact that $|P_r^b(i) - (1/2 + \varepsilon)| \leq \varepsilon^4$ if $d(i,r) \geq \ell/2$ as proved in Claim B.1.

The second statement can be proven in a similar fashion. □

References

[AAT05] Michael Alekhnovich, Sanjeev Arora, and Iannis Tourlakis. Towards strong nonapproximability results in the Lovasz-Schrijver hierarchy. In *Proceedings of the 37th ACM Symposium on Theory of Computing*, pages 294–303, 2005. 4

[ABLT06] Sanjeev Arora, Béla Bollobás, László Lovász, and Iannis Tourlakis. Proving integrality gaps without knowing the linear program. *Theory of Computing*, 2(2):19–51, 2006. 2, 3, 6, 7, 11, 20

[BOGH+03] Josh Buresh-Oppenheim, Nicola Galesi, Shlomo Hoory, Avner Magen, and Toniann Pitassi. Rank bounds and integrality gaps for cutting planes procedures. In *Proceedings of the 44th IEEE Symposium on Foundations of Computer Science*, pages 318–327, 2003. 4

[Cha02] Moses Charikar. On semidefinite programming relaxations for graph coloring and vertex cover. In *Proceedings of the 13th ACM-SIAM Symposium on Discrete Algorithms*, pages 616–620, 2002. 4

[DS05] Irit Dinur and Shmuel Safra. On the hardness of approximating minimum vertex-cover. *Annals of Mathematics*, 162(1):439–486, 2005. 2

[FdlVKM07] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of Maxcut. In *Proceedings of the 18th ACM-SIAM Symposium on Discrete Algorithms*, 2007. 2, 7

[GW95] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995. Preliminary version in *Proc. of STOC'94*. 2

[KG98] Jon M. Kleinberg and Michel X. Goemans. The Lovász Theta function and a semidefinite programming relaxation of vertex cover. *SIAM Journal on Discrete Mathematics*, 11:196–204, 1998. 4

[KR03] Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2 - \varepsilon$. In *Proceedings of the 18th IEEE Conference on Computational Complexity*, 2003. 2

[LS91] L. Lovasz and A. Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM J. on Optimization*, 1(12):166–190, 1991. 1, 3

[SA90] Hanif D. Sherali and Warren P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3:411–430, 1990. 2

[STT06] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovasz-Schrijver SDP relaxations of vertex cover. Technical Report TR06-098, Electronic Colloquium on Computational Complexity, 2006. 4

[Tou06] Iannis Tourlakis. New lower bounds for vertex cover in the Lovasz-Schrijver hierarchy. In *Proceedings of the 21st IEEE Conference on Computational Complexity*, 2006. 2, 3, 7

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/