Electronic Colloquium on Computational Complexity, Report No. 169 (2010)

The Sherali-Adams System applied to Vertex Cover: Why Borsuk Graphs Fool Strong LPs and some Tight Integrality Gaps for SDPs

Siavosh Benabbas* Konstantinos Georgiou‡ Avner Magen*

*Department of Computer Science, University of Toronto

‡Department of Combinatorics and Optimization, University of Waterloo

November 5, 2010

Abstract

We study the performance of the Sherali-Adams system for VERTEX COVER on graphs with vector chromatic number $2+\epsilon$. We are able to construct solutions for LPs derived by *any number* of Sherali-Adams tightenings by introducing a new tool to establish Local-Global Discrepancy. When restricted to $\Theta(1/\epsilon)$ tightenings we show that the corresponding LP treats the input graph as a nearly perfect matching. Since there exist graphs with $2+o(1)$ vector chromatic number but no linear-sized independent sets, this immediately implies a tight integrality gap for superconstant levels of the Sherali-Adams system.

An important property of our solutions is that they can be slightly perturbed to also satisfy semidefinite conditions. In particular, using this approach we prove the first tight integrality gap for non trivial levels of the Sherali-Adams SDP system. Our argument reduces semidenifiteness to a condition on the Taylor expansion of a reasonably simple function that we are able to establish up to constant-level SDP tightenings. We conjecture that this condition holds even for superconstant levels which would imply that in fact our solution is a valid for superconstant level Sherali-Adams SDPs.

ISSN 1433-8092

1 Introduction

A vertex cover of a graph $G = (V, E)$ is a subset S of the vertices such that for every edge $ij \in E$ at least one vertex among i, j lies in S. In the MINIMUM VERTEX COVER problem the objective is to find the vertex cover of minimum size. Determining the approximability of VERTEX COVER is one of the outstanding open problems in theoretical computer science. While a 2-approximation algorithm is rather straightforward, considerable efforts have failed to yield any algorithm with approximation ratio $2 - \Omega(1)$. Indeed the best known approximation algorithm achieves an approximation ratio of $2 - O(\sqrt{1/\log n})$ [Kar09]. On the other hand, the strongest PCP-based hardness result [DS05] shows that 1.36-approximation of VERTEX COVER is $\mathbf{NP}$-hard. Only if one is willing to assume Khot's Unique Game Conjecture [Kho02] this hardness can be improved to $2 - o(1)$.

Trying to resolve the true approximability of VERTEX COVER, one could study the behavior of prominent algorithmic paradigms, such as Linear Programming (LP) and Semidefinite Programming (SDP) relaxations. There, the measure of efficiency is the Integrality Gap which sets the approximation limitations of any algorithm based on these relaxations. A number of systematic procedures, known as Lift-and-Project systems have been proposed to improve the integrality gap of standard relaxations. These systems build strong hierarchies of either LP relaxations (as in the Lovász-Schrijver and the Sherali-Adams systems) or SDP relaxations (as in Lovász-Schrijver SDP and the Lasserre systems.) In this work we study the limitations of strong LP and SDP relaxations that can be derived by the powerful system of Sherali-Adams enhanced by a positive-semidefiniteness constraint. The performance of the same system has been previously studied for other combinatorial problems, for example, for Constraint Satisfaction Problems [BGMT10], for Quadratic Programing and MaxCutGain [BM10], for Maximum Cut and Sparsest Cut [KS09], and for a variety of different problems in [RS09], but its integrality gap for VERTEX COVER remains open.

The Sherali-Adams system, like all Lift-and-Project systems, can be thought of as being applied in rounds (also called levels). The bigger the number of rounds used, the more accurate the obtained tightened relaxation is. In fact, if as many rounds as the number of variables are used, the final relaxation is exact, all its feasible solutions are distributions over integral solutions, whence no integrality gap exists. On the other hand the size of the derived relaxation grows exponentially with the number of rounds, which implies that the time one needs in order to solve it also grows accordingly. It is natural then to ask if looking at a modest number of rounds (say t being a constant, or $\log \log n$) will result in an algorithm with approximation factor better than 2. In this paper we answer the above question in the negative for the Sherali-Adams and the Sherali-Adams SDP system. That is, we prove that the Sherali-Adams system applied to the canonical LP and the canonical SDP relaxations of VERTEX COVER have integrality gap $2 - o(1)$ for $\omega(1)$ and (some fixed) constant number of rounds respectively.

Lower bounds in the Lift-and-Project systems amount to showing that the integrality gap remains large even after many tightenings of the systems. Integrality gaps for $\omega(1)$ rounds rule out superpolynomial time algorithms albeit for a restricted model of computation. For VERTEX COVER, considerable effort has been invested in strong lower bounds for various hierarchies. For LP hierarchies, [STT07] shows an integrality gap of $2-\epsilon$ for $\Omega(n)$ rounds of the LS system and [CMM09] shows the same integrality gap for the stronger Sherali-Adams system up to $\Omega(n^\delta)$ rounds (with δ going to 0 together with ϵ). Both these results study the system applied to the canonical LP formulation and are not comparable to integrality gaps for even the standard SDP relaxation. Our contribution for LP hierarchies is as follows.

Theorem 1.1. *For every ϵ, there are graphs on n vertices such that the level-$\Omega(\sqrt{\log n/\log\log n})$ LP derived by the Sherali-Adams system for* VERTEX COVER *has integrality gap $2 - \epsilon$.*

Although our integrality gap is weaker than the one of [CMM09] in terms of the number of tightenings, our result follows as an immediate corollary of a novel construction of feasible solutions for a wide family of graphs enjoying nice geometric properties,i.e. Borsuk graphs. In particular, part of the novelty is that we are able to construct solutions for *any number of Sherali-Adams tightenings*. When the number of tightenings is large enough, then unavoidably the solution is inside the integral hull. In contrast, when the number of tightenings is relatively small, then the corresponding LP can be seen to treat the input graph as a nearly perfect matching, i.e. as a graph where just half the vertices form a vertex cover. In particular, in Section 5 we show that for Borsuk graphs there is always a Sherali-Adams solution where each vertex contributes almost 1/2, with the above theorem following as a corollary. The reader should also contrast Theorem 1.1 to the result of Vishwanathan [Vis09] which shows that every hard instance of VERTEX COVER (if the true inapproximability is $2 - o(1)$) must have large subgraphs that "look like Borsuk graphs". In fact one could think of our result as some form of a converse to that of Vishwanathan as far as the Sherali-Adams hierarchy is concerned.

For SDP hierarchies, and for the LS_+ system which is stronger than both the LS system and the canonical SDP formulation (but incomparable to Sherali-Adams), [GMPT07] shows an integrality gap of $2 - \epsilon$ for $\Omega(n)$ levels. The next logical step then is to study the Sherali-Adams SDP system which is also stronger than the LS system, and a significant step closer to the so-called Lasserre system, for which no known tight integrality gaps exist for combinatorial problems with hard constraints. We prove the next theorem.

Theorem 1.2. *For every ϵ, there are graphs on n vertices such that the level-5 SDP derived by the Sherali-Adams SDP system for* VERTEX COVER *has integrality gap $2 - \epsilon$.*

Theorem 1.2 follows as a corollary of Theorem 1.1 after carefully transforming its solution, so as to make sure that the function describing its solution has Taylor expansion with positive coefficients, a condition that we prove is sufficient to satisfy the additional semidefinite constraint. The new solution can be almost trivially seen to satisfy the level-2 Sherali-Adams SDP(see Section 7), which gives an new alternative proof of the well known tight integrality gap for the standard VERTEX COVER SDP tightened by the triangle inequality. For a higher level solution (for which no tight integrality gaps were known prior to this work) we first describe our solution in terms of a seemingly simple function (depending on the number of tightenings). Next we are able to show that the function has Taylor expansion with positive coefficients up to level-5. We should point here that we know empirically that the function maintains positivity in its coefficients even for superconstant tightenings, however the proof is still eluding us. Finally, if one is content with an integrality gap less than 2, integrality gaps of 1.36 for up to $\Omega(n^\delta)$ levels [Tul09] and $7/6$ for up to $\Omega(n)$ levels [Sch08] for the Lasserre system are known.

It might be tempting to ask whether a negative result concerning Lift-and-Project systems is interesting. It turns out that the best algorithms known for many combinatorial optimization problems are based on relaxations weaker than those derived by constant rounds of the Sherali-Adams SDP system which we study here. Examples include the seminal algorithm of Goemans-Williamson [GW95] for MAX CUT, the Karloff-Zwick algorithm [KZ97] for MAX-3-SAT, the Arora-Rao-Vazinari algorithm [ARV09] for SPARSEST CUT and the best algorithm for VERTEX COVER of Karakostas [Kar09]. Lift-and-Project hierarchies have been also used recently in designing approximation algorithms with a runtime-approximation ratio trade off. The list of relevant examples, which includes [dlVKM07] for MAX CUT, [MM09] for VERTEX COVER and INDEPENDENT SET in minor-free graphs, [CS08] for INDEPENDENT SET, [BCG09] for the MaxMin allocation problem, and [ARKN10] for KNAPSACK, is continually growing. Finally, for some particular constraint satisfaction problems, and modulo the UGC, no approximation algorithm can perform better than the one obtained by Sherali-Adams SDP of a constant number of rounds. In fact, a recent result of Raghavendra [Rag08] gives strong evidence that algorithms based on the Sherali-Adams SDP system constitute a strong model of computation.

Outline of the method: To describe the high level idea of our approach let us first give a brief description of the Sherali-Adams SDP system applied to the VERTEX COVER problem. The Sherali-Adams SDP system of level t is a Semidefinite program with the following variables. If G is a graph and U is any subset of its vertices of size at most t the program will have real variables to specify a distribution $\mathcal{D}(U)$ over the subsets of U. Furthermore, the program will have two kinds of constraints. The first kind ensure that any subset of U that is assigned a positive probability covers all the edges inside U, i.e. the distribution $\mathcal{D}(U)$ is over vertex covers of U. The second kind of constraints ensures that the marginals of the distributions for $U_1 \subseteq U$ are consistent on U_1, i.e. any event that only depends on the vertices of U_1 has the same probability according to $\mathcal{D}(U_1)$ and $\mathcal{D}(U)$. The program will also have a semi-definiteness constraint on some of these variables which we will skip in the introduction. The objective value of the program is the sum over all vertices v of the probability that v is in the local vertex covers.

The instances we use for our integrality gap construction are the Frankl-Rödl graphs (which are subgraphs of Borsuk graphs), parameterized by an integer m and a real parameter $\gamma > 0$. The vertex set of these graphs is $\{0,1\}^m$ with an edge between two vertices if their Hamming distance is $m(1-\gamma)$. A result by Frankl-Rödl [FR87] shows that if γ is a positive constant, all vertex covers of this graph have size $2^m(1-o(1))$. In fact [GMPT07] show the same holds even when γ is as small as $\sqrt{\log m/m}$. A tight integrality gap therefore calls for a solution in the system of objective value $1/2+\epsilon$ for an arbitrary small constant $\epsilon > 0$.

Here is a simple description of our solution. Notice that a simple vertex cover of this graph is a Hamming ball (with arbitrary center) with radius more than $(1+\gamma)m/2$. This can be seen by observing that the diameter of the complement of such set is less than γm and hence it does not contain any edge. A geometric way to obtain a distribution of (the same) vertex covers would be to imagine the hypercube in $\mathcal{R}^m$ with its center at origin and take a sufficiently large *spherical cap* centered at a random point on the sphere. Of course, in doing so we have not achieved much since we are defining a (global) distribution of vertex covers, and thus each has to be of size $2^m(1-o(1))$. It is useful to understand this "obstacle" from a geometric point of view: the height of the spherical cap must be at least $1+\sqrt{\gamma}$ (instead of 1 which would give half a hyper-sphere) and by concentration of measure on the sphere the area of such a cap is roughly $1/2+\sqrt{m\gamma}$ of the whole sphere. Then, by averaging, the probability that an arbitrary point on the sphere (and in particular a vertex of the hypercube) is in the cap is $1/2+\sqrt{m\gamma}$ which is rather large.

The main idea is that if one needs to define probabilities only for small sets (that corresponds to a level-t Sherali-Adams relaxation for some "small" t), one can first embed these points in a *small dimensional sphere* and then repeat the experiment to define a random vertex cover. The spherical caps that are required in order to cover the edges in these sets have the same height, but now, due to the lower dimension, their area is greatly reduced! Specifically, if the original set has at most t points the experiment can be performed in a t-dimensional sphere and if $\gamma t = o(1)$, the probability of a vertex participating in the vertex cover will be no more than $1/2+o(1)$. It is critical, of course, that the obtained distributions are consistent. But this is "built-in" in this experiment. Indeed, due to spherical symmetry, the probability that a set of points on a t dimensional sphere belong to a random cap of a fixed radius depends only on t, on the radius of the cap and on the pairwise Euclidean distances of the points in the set. We should also point out that this construction would work for any graph with vector chromatic number $2+o(1)$. In another words, if G is graph that can be embedded into the Unit sphere so that edges may only appear between vertices whose embedded points sum to a vector of norm $o(1)$, then there is a sufficiently "low-level" (but non-trivial) Sherali Adams solution of value $(1/2+o(1))n$.

Unfortunately, we can not show that the above solution satisfies the positive semi-definiteness constraints of the Sherali-Adams SDP system, instead we change our solution in several ways to attain positive semi-definiteness. These changes are somewhat technical and we avoid discussing them in detail here. At a high level the changes are (i) we add a small probability of picking the *whole* graph as the vertex cover. (ii) We apply a transformation of the canonical embedding of the cube in the sphere that ensures that the farthest pairs of vertices are precisely the edges, and also that the inner products have a bias to being positive (compared to the canonical embedding in which the average inner product is 0).

To get some insight into the rationale of these modification, first note that the matrix whose positive definiteness we need to prove happens to be highly symmetric. For such symmetric matrices a necessary condition for positive semi-definiteness is that the average entry is at least as large as the square of the diagonal entries. Manipulation (i) above is precisely the tool we need to ensure this condition, and has no adverse effects otherwise. The second transformation is useful although not clearly necessary. We can, however, argue that without a transformation of this nature, a good SDP solution is possible also for a graphs in which edges connect vertices that are at least as far as $m(1-\gamma)$ (rather than exactly that distance). The existence of solutions for such dense graphs seems intuitively questionable. Last, boosting the typical inner product can be shown to considerably boost the Taylor coefficients of a certain function which we need to show only has positive Taylor coefficients.

Comparison to other works: More than half-dozen different integrality gap constructions for Vertex cover in different Lift-and-Project systems are known. Among these, the most relevant to our work is [CMM09]. In [CMM09], Charikar, Makarychev and Makarychev obtain a Sherali Adams solution that is based on embedding the vertices of the graph in the sphere. The similarity with our work is that Makarychev et al. take a special case of caps, i.e. half-spheres, in order to determine probabilities. Consistency of these distribution is, just as our case, guaranteed by the fact that these probabilities are intrinsic to the local distances of the point-set in question. However, the reason that these distributions behave differently than a global distribution (which is essential for an integrality gap construction) is completely different than ours. It is easy to see that when the caps in the construction are half spheres, the dimension does not play a role at all. However, in [CMM09] there is no *global embedding* of the points in the sphere rather only a local one. In contrast, our distributions can be defined for all dimensions, however as we mentioned we must keep the dimension reasonably small in order to guarantee small objective value. Another big difference pertains to the different instances. While our construction may very well be the one (or close to the one) that will give Lasserre integrality-gap bound, the instances of [CMM09] have no substantial integrality gap even for the standard SDP.

It is also important to put our work in context with the sequence of results dealing with SDP integrality gaps of Vertex-Cover [GK98, Cha02, GMPT07, GMT09a]. In these works the solution can be thought of as an approximation to a very simple set: a dimension cut, that is a face of the cube. This set is not a vertex cover, but in some geometric sense is close to one. The SDP solutions are essentially averaging of such dimension-cuts with some carefully crafted perturbations. Using the same language, the solution we present in the current work is based on Hamming balls of radius $m/2$ (i.e. translations of the majority function) rather than dimension-cuts (i.e. dictatorship functions.) The perturbation we apply to make such a solution valid is simply the small increase in the radius of the Hamming balls.

2 Preliminaries

Definition 2.1. (*Borsuk graphs*) The Borsuk graph B_δ^m is an infinite graph with vertex set S^{m-1}. Two vertices $\mathbf{x}, \mathbf{y}$ are adjacent if they are nearly antipodal, namely $\|\mathbf{x}+\mathbf{y}\| \le 2\sqrt{\delta}$.

In the current paper we will study discrete finite subgraphs of B_δ^m.[1] Next, note that any perfect matching

[1]Note that an alternative characterization of any subgraph of B_δ^m is that its vector chromatic number is $k = 2 + \frac{2\delta}{1-2\delta}$. In other

is a discrete subgraph of B_0^2. Conversely, any discrete subgraph of B_0^m is clearly a perfect matching, and hence, for small values of δ, it is tempting to think of B_δ^m as nearly perfect matchings. In other words, one might expect that for any discrete subgraph (not necessarily induced) of B_δ^m, for some small value δ, there exist vertex covers of size almost half the number of vertices. Interestingly, as we shall see in a while, this is not always the case.

Definition 2.2. (*Frankl-Rödl graphs*) Fix γ, $0 \leq \gamma \leq 1$ and an integer $m \geq 1$. The *Frankl-Rödl graph* G_γ^m is the graph with vertices $\{-1,1\}^m$ and where two vertices $i,j \in \{-1,1\}^m$ are adjacent if $d_H(i,j) = (1-\gamma)m$.

Note that the graph G_γ^m can be embedded in S^{m-1} by just normalizing the hypercube; the vertices of G_γ^m are unit vectors $\mathbf{z}_i$ in $\frac{1}{\sqrt{m}}\{-1,1/\}^m$. For some edge ij, we then have

$$\|\mathbf{z}_i + \mathbf{z}_j\|^2 = 2 + 2\,\mathbf{z}_i \cdot \mathbf{z}_j = 2 + 2(-1+2\gamma) = 4\gamma,$$

showing that G_γ^m is a subgraph of B_γ^m.

Frankl-Rödl graphs exhibit an interesting "extremal" combinatorial property. From the discussion above, G_0^m are perfect matchings and so the minimum vertex cover has size half the number of the vertices. A beautiful theorem by Frankl-Rödl says that by slightly perturbing γ, any vertex cover of the resulting graphs has size $2^m - o(2^m)$. The following slight modification of the original theorem of Frankl and Rödl (Theorem 1.4 in [FR87]) was first proved in [GMPT07].

Theorem 2.3. *Let m be an integer and let $\gamma = \Theta(\sqrt{\log m/m})$ be a sufficiently small number so that γm is an even integer. Then there is no independent set in G_γ^m of size larger than $2^m/m$.*

Consequently, the minimum fraction of vertices that are required to form a vertex cover in G_γ^m, with $\gamma = \Theta(\sqrt{\log m/m})$ is of order $1-o(1)$. Frankl-Rödl graphs have been used as tight integrality gap instances in a series of results [GK98, Cha02, GMPT07, GMT08, GMT09a]. In particular, in [Cha02] it is shown that G_γ^m is an induced subgraph of $B_\gamma^{2^m}$. In the current paper, we show that all discrete subgraphs (not necessarily induced) of B_γ^m are treated almost as perfect matchings by a strong family of LPs and SDPs.

Our construction is based on tensored vectors. Recall that the tensor product $\mathbf{u} \otimes \mathbf{v}$ of vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$ is the vector in $\mathbb{R}^{nm}$ indexed by ordered pairs from $n \times m$ and assuming the value $u_i v_j$ at coordinate (i,j). Define $\mathbf{u}^{\otimes d}$ to be the vector in $\mathbb{R}^{n^d}$ obtained by tensoring $\mathbf{u}$ with itself d times. Let $P(x) = c_1 x^{t_1} + \ldots + c_q x^{t_q}$ be a polynomial with nonnegative coefficients. Then T_P is the function that maps a vector $\mathbf{u}$ to the vector $T_P(\mathbf{u}) = (\sqrt{c_1}\mathbf{u}^{\otimes t_1}, \ldots, \sqrt{c_q}\mathbf{u}^{\otimes t_q})$. Polynomial tensoring can be used to manipulate inner products in the sense that $T_P(\mathbf{u}) \cdot T_P(\mathbf{v}) = P(\mathbf{u} \cdot \mathbf{v})$ and it was used for many integrality gap results such as [GK98, Cha02, GMPT07, GMT08].

3 Strong relaxations for VERTEX COVER

A standard exact formulation of VERTEX COVER can be obtained as follows. Consider some instance $G = (V,E)$ of the VERTEX COVER problem. A valid solution, i.e. a vertex cover, is a partition of the vertices into two sets. We associate every vertex $i \in V$ with a variable $x_i \in \{0,1\}$ with the intended meaning that if a variable is set to 1, then the corresponding vertex will be in the vertex cover. The hard condition that every edge $ij \in E$ needs to be covered by at least one of its endpoints can be clearly formulated as $x_i + x_j \geq 1$. Then, the objective linear function $\sum_{i \in V} x_i$ counts the size of the vertex cover.

words, for any two adjacent vertices i,j and their corresponding vector representation $\mathbf{z}_i, \mathbf{z}_j$ in S^{m-1} we have $\mathbf{z}_i \cdot \mathbf{z}_j \leq -\frac{1}{k-1}$.

We have argued that that following optimization problem, known as an integer programming problem, is an exact formulation for VERTEX COVER.

$$\begin{array}{lll} \min & \sum_{i \in V} x_i & \\ \text{s.t.} & x_i + x_j \geq 1 & \forall ij \in E \\ & x_i \in \{0,1\} & \forall i \in V. \end{array} \tag{1}$$

Integer Programming is an intractable optimization problem. However, the usefulness of the previous formulation becomes transparent by relaxing the integral condition $x_i \in \{0,1\}$ into $x_i \in [0,1]$. The resulting LP

$$\begin{array}{llll} \min & \sum_{i \in V} x_i & & \\ \text{s.t.} & x_i + x_j \geq 1 & \forall ij \in E & (\textit{Edge constraints}) \\ & x_i \in [0,1] & \forall i \in V, & \end{array} \tag{2}$$

is known as the standard LP relaxation for VERTEX COVER, and can be solved in polynomial time. Any such relaxation is inherently associated with what is known as the integrality gap (integrality gap), which measures how much is the relaxed optimal solution off from the true optimal. Formally, for some instance graph, the integrality gap of a relaxation is the ratio between the true optimal solution over the optimal solution of the relaxation. The integrality gap of a relaxation is defined as the supremum of the integrality gaps over all instances. The integrality gap serves as an important measure of effectiveness, as it bounds the approximation that any algorithm can achieve that is based on the relaxation.

It is an easy exercise to show that the integrality gap of (2) is at most 2, showing that no $(2-\Omega(1))$-approximation algorithm can be based on this LP. The tightness of the integrality gap can be easily shown by considering a complete graph on n vertices: since the all $1/2$ vector is always a solution of (2), the integrality gap is $\frac{n-1}{0.5\, n} = 2 - 2/n$.

Sherali and Adams [SA90] proposed a systematic procedure for tightening 0/1 polytopes. For convenience we will present the definition just for the vertex cover polytope (2). In order to give some motivation for the Sherali-Adams system, we need some notation. We abbreviate the set $\{0, 1 \ldots, n\}$ by $[n]$. For $A \subseteq [n]$ we denote by $\mathcal{P}^A, \mathcal{P}_t^A$ the powerset of A, and all subsets of A of size at most t respectively. For $\mathbf{y} \in \mathbb{R}^{\mathcal{P}^{[n]}}$, $t \in [n]$ and $U \in \mathcal{P}^{[n]}$ we define the matrices

$$(M_n(\mathbf{y}))_{I,J} \;:=\; y_{I \cup J}, \;\; \forall I, J \in \mathcal{P}^{[n]} \tag{3}$$

$$(M_U(\mathbf{y}))_{I,J} \;:=\; y_{I \cup J}, \;\; \forall I, J \in \mathcal{P}^{U}. \tag{4}$$

We also set $y_0 = y_\emptyset, y_{\{i\}} = y_i$, and in general we treat the set $\{0\}$ as the empty set $\emptyset$. From the definitions above it is clear that $M_n(\mathbf{y}) = M_{[n]}(\mathbf{y})$. Now, for any $A \subseteq [n]$ we define the "shifting" $\mathbf{y}_{+A}$ of $\mathbf{y} \in \mathbb{R}^{\mathcal{P}^{[n]}}$, as a vector in $\mathbb{R}^{\mathcal{P}^{[n]}}$, with $(y_{+A})_I = y_{A \cup I}$.

Now consider an integral solution $\mathbf{x}$ of (2) for some instance n-vertex graph $G = (V, E)$, and define $\mathbf{y} \in \mathbb{R}^{\mathcal{P}^{[n]}}$ as $y_I = \prod_{i \in I} x_i$. Then it is easy to check that

$$M_{[n]}(\mathbf{y}) \succeq 0, \quad M_{[n]}(\mathbf{y}_{+\{i\}} + \mathbf{y}_{+\{j\}} - \mathbf{y}) \succeq 0, \; \forall ij \in E. \tag{5}$$

Generalizing this, one can show that for a global distribution of vertex covers on $[n]$, if we set y_I equal to the probability that all variables in I are set to 1, then $\mathbf{y}$ satisfies constraints (5). In fact, it is known [LS91, SA90] that adding constraints (5) to the LP (2) yields the convex closure of integral solutions of the exact formulation , and therefore the integrality gap is 1. Clearly, the resulting LP has exponentially many constraints. For this, Sherali and Adams proposed the following relaxation of the vertex cover polytope of some n-vertex graph $G = (V, E)$, using the linear variables variables $\mathbf{y} \in \mathbb{R}^{\mathcal{P}_t^{[n]}}$, for some fixed t.

$$\begin{array}{lll} \min & \sum_{i \in V} y_i & \\ \text{s.t.} & M_U(\mathbf{y}_{+\{i\}} + \mathbf{y}_{+\{j\}} - \mathbf{y}) \succeq 0 \quad \forall ij \in E,\ \forall U \in \mathcal{P}_{t-1}^{[n]} & (\textit{Edge constraints}) \\ & M_U(\mathbf{y}) \succeq 0 \quad \forall U \in \mathcal{P}_t^{[n]} & \\ & y_i \in [0,1] \quad \forall i \in V, & \end{array} \tag{6}$$

Relaxation (6) is known as the level-t Sherali-Adams relaxation for VERTEX COVER, and more interestingly, it is a linear relaxation in disguise (see [Lau03])! We do not need to take advantage of this nice fact, as we know of an alternative argument that guarantees that a $\mathbf{y} \in \mathbb{R}^{\mathcal{P}_t^{[n]}}$ is in fact a solution for (6) (the same argument was first used in [dlVKM07], and later in [CMM09, GM08, GMT09b]). This is based on the following easy observation.

Fact 3.1. *Suppose that we can associate every $U \in \mathcal{P}_t^{[n]}$ with a distribution $\mathcal{D}(U)$ of "local" vertex covers (namely distribution of 0/1 assignments on U with all local edge constraints satisfied). Suppose also that the resulting family of distributions is locally consistent, namely, for all $U, U' \subseteq \mathcal{P}_t^{[n]}$ the distributions $\mathcal{D}(U), \mathcal{D}(U')$ agree on $U \cap U'$. Define the vector $\mathbf{y} \in \mathbb{R}^{\mathcal{P}_t^{[n]}}$ as $y_U = \mathbb{P}_{\mathcal{D}(U)}[U$ is in the vertex cover$]$. Then $\mathbf{y}$ satisfies relaxation (6).*

The reason is almost self evident. If we have such a family of distributions as described in Fact 3.1, then for all $U \in \mathcal{P}_t^{[n]}$, and for all $A \subseteq U$, the distributions $\mathcal{D}(A), \mathcal{D}(U)$ agree on A. Hence, y_A can be seen as the probability of an event defined on U, and therefore all matrices $M_U(\mathbf{y}), M_U(\mathbf{y}_{+\{i\}} + \mathbf{y}_{+\{j\}} - \mathbf{y})$ are convex combinations of positive semidefinite matrices.

One of the most challenging problems in the area is to prove a tight integrality gap in the so-called Lasserre system that requires that the matrices $M_t(\mathbf{y})$ are positive semidefinite. An intermediate system of relaxations between the Sherali-Adams system and the Lasserre system would be to add to the relaxation 6 the level-1 condition of the Lasserre system, namely to require that $M_1(\mathbf{y}) \succeq 0$. We refer to the resulting relaxation (7) as the level-t Sherali-Adams SDP.

$$\begin{array}{lll} \min & \sum_{i \in V} y_i & \\ \text{s.t.} & M_U(\mathbf{y}_{+\{i\}} + \mathbf{y}_{+\{j\}} - \mathbf{y}) \succeq 0 \quad \forall ij \in E,\ \forall U \in \mathcal{P}_{t-1}^{[n]} & (\textit{Edge constraints}) \\ & M_U(\mathbf{y}) \succeq 0 \quad \forall U \in \mathcal{P}_t^{[n]} & \\ & M_1(\mathbf{y}) \succeq 0 & \\ & y_i \in [0,1] \quad \forall i \in V, & \end{array} \tag{7}$$

4 Local Distributions of Vertex Covers for Borsuk Graphs

In this section we study relaxation (6) for discrete subgraphs of B_γ^m on n vertices. In particular, for any $t \leq n$, we show how to find $\mathbf{y} \in \mathbb{R}^{\mathcal{P}_t}$ that satisfies the level-t Sherali-Adams relaxation (6) of the vertex cover polytope. Note that in light of Fact 3.1, our problem reduces in associating every set $U \subseteq [n]$ with a distribution of vertex covers that are locally consistent.

The family of distributions we are looking for arises from the following experiments. For this, we fix some discrete subgraph $G = (V, E)$ of B_δ^m on n vertices, for which we also consider the representation on S^{m-1} to be known. Note that t can assume any value in $[n]$ but it must be fixed beforehand.

Experiment Local-Global

The input is any $I \subseteq V$, of size at most t, and some $\sqrt{\delta} > 0$.
The result of the experiment is a distribution of 0/1 assignments on I.

(a) Embed the I-induced subgraph of G into S^{t-1} preserving all pairwise Euclidean distances.
(b) In S^{t-1} consider the complement C of a random spherical cap of height $1 - \sqrt{\delta}$.
(c) Vertices of I are assigned 1 if the lie in the cap C, otherwise they are assigned 0.

For completeness, we can say that part (a) of Experiment Local-Global can be realized by any orthonormal mapping, and it is possible since $|I| \leq t$. For some I, denote the vectors in S^{t-1} as $(\mathbf{z}_i)_{i \in I}$. [2] Hence, according to the distribution $\mathcal{D}(I)$, vertices $i \in I$ are assigned the value 1 only if $\mathbf{w} \cdot \mathbf{z}_i \leq \sqrt{\delta}\|\mathbf{w}\|$.

We want to use Fact 3.1 now to argue that the vector $\mathbf{y} \in \mathbb{R}^{\mathcal{P}_t^{[n]}}$, such that

$$y_I = \mathop{\mathbb{P}}_{\mathbf{w} \in S^{t-1}} [\mathbf{w} \cdot \mathbf{z}_i \leq \sqrt{\delta},\ \forall i \in I] \tag{8}$$

satisfies the relaxation (6).

Lemma 4.1. *For every finite subgraph of B_δ^m on n vertices, the family of distributions $\mathcal{D}(I)$, $I \in \mathcal{P}_t^{[n]}$, is a distribution of locally consistent vertex covers, in the context of Fact 3.1.*

Proof. Indeed, we can conclude local consistency if we combine the following two observations: i) t is fixed a priori and it is the same for all $I \in \mathcal{P}_t^{[n]}$, ii) the probability in (8) depends only on the pairwise Euclidean distances of vectors in I.

It therefore remains to argue that $\mathcal{D}(I)$ is a distribution of vertex covers. To that end, we need to show that in the Experiment Local-Global, two adjacent vertices cannot be at the same time outside the random cap C. This is true, because i, j are outside the cap only if $\mathbf{w} \cdot \mathbf{z}_i > \sqrt{\delta}$ and $\mathbf{w} \cdot \mathbf{z}_j > \sqrt{\delta}$, in which case,

$$\|\mathbf{z}_i + \mathbf{z}_j\| = \|\mathbf{w}\|\ \|(\mathbf{z}_i + \mathbf{z}_j)\| \geq \mathbf{w} \cdot (\mathbf{z}_i + \mathbf{z}_j) > 2\sqrt{\delta},$$

where the penult inequality is given by Cauchy-Schwarz. Since G is a subgraph of B_γ^m, we conclude that ij cannot be en edge. $\square$

5 Nearly Matchings for the Sherali-Adams System

Roughly speaking, the Sherali-Adams relaxation of the vertex cover polytope treats a graph like an almost perfect matching, if the contribution of every vertex in the objective function of (6) is $1/2+\epsilon$, for some small $\epsilon > 0$. We should therefore ask ourselves, what is the probability that in the Experiment Local-Global, a vertex lies in the random cap.

Observe that our goal is to produce distributions of vertex covers such that the marginals on the singletons are almost $1/2$. Note that if in the Experiment Local-Global, the value of t is very large, then from high dimensional phenomena we will have that the area of the big cap C is very close to that of S^{t-1}, making the probability that a vertex is chosen close to 1. At the same time, if t is relatively small the marginals on the singletons could be relatively close to 1/2. We make this formal in the next technical lemma. Its proof uses a simple concentration argument that we present in the Appendix.

[2] Then, (although it is irrelevant to our analysis), step (b) can be achieved as follows. Consider t independent random variables $w_i,\ i = 1, \ldots, t$, of the normal distribution $N(0, 1)$. It is known that the vector $\mathbf{w}/\|\mathbf{w}\|$, where $\mathbf{w} = (w_1, \ldots, w_t)$ is distributed uniformly on S^{t-1}.

Lemma 5.1. *For any fixed* $\mathbf{z} \in S^{t-1}$*, we have* $\mathbb{P}_{\mathbf{w}\in S^{t-1}}[\mathbf{w}\cdot\mathbf{z} \le \eta] \le \frac{1}{2} + \eta\sqrt{\frac{\pi}{8}(t+1)}$*, where* $\mathbf{w}$ *is distributed uniformly on* S^{t-1}*.*

We are ready to show that the Sherali-Adams system treats discrete finite subgraphs of B^m_δ as nearly perfect matchings.

Theorem 5.2. *Let G be a finite subgraph of B^m_δ on n vertices. Then the level-$\left(\frac{2\epsilon^2}{\pi}\frac{1}{\delta} - 1\right)$ Sherali-Adams relaxation (6) of the vertex cover polytope has objective value $(1/2+\epsilon)n$.*

Proof. We use Experiment Local-Global with parameters $t = \left(2\epsilon^2/\pi\delta - 1\right)$ and $\sqrt{\delta}$ (so that the excluded random caps will have height $1-\sqrt{\delta}$), to define a family of local 0/1 distributions. From Lemma 4.1 these are locally consistent distributions of vertex covers. Then Fact 3.1 gives us a level-t Sherali-Adams solution.By Lemma 5.1, setting $\eta = \sqrt{\delta}$, the contribution of each vertex in the objective function is exactly $1/2+\epsilon$. ☐

As an immediate corollary, we now conclude a tight integrality gap for VERTEX COVER and the Sherali-Adams system, namely Theorem 1.1.

Proof. (of Theorem 1.1) We start with the n-vertex Frankl-Rödl graphs G^m_δ, $n = 2^m$, which are finite subsets of the Borsuk graph B^m_δ. We set $\delta = \Theta(\sqrt{\log m/m})$ so as the conditions of Theorem 2.3 to be satisfied, namely all vertex covers of G^m_δ are of size $n - o(n)$. Theorem 5.2 then implies that the level-$\left(\frac{2\epsilon^2}{\pi}\frac{1}{\delta} - 1\right)$ Sherali-Adams relaxation of the vertex cover polytope has integrality gap at least

$$\frac{n-o(n)}{(1/2+\epsilon)n} = 2 - 2\epsilon - o(1).$$

☐

6 Preparatory Observations for the Sherali-Adams SDP Solution

We need the following sufficient condition for a principal submatrix of $M_1(\mathbf{y})$ to be positive semidefinite. For this, it is convenient to denote by $M_1'(\mathbf{y})$ the principal submatrix $M_1(\mathbf{y})$ indexed by nonempty sets. Note also that for the Sherali-Adams we introduced in the previous section, all $y_{\{i\}}$ attain the same value, say y_R. In that notation, we have

$$M_1(\mathbf{y}) = \begin{pmatrix} 1 & \mathbf{1}y_R \\ \mathbf{1}^T y_R & M_1'(\mathbf{y}) \end{pmatrix}, \tag{9}$$

where by $\mathbf{1}$ we denote the all 1 vector of appropriate size. Then we have the following easy fact.

Fact 6.1. *Suppose that the all 1 vector is an eigenvector for $M_1'(\mathbf{y})$. Then the following are equivalent.*
(a) The matrix $M_1(\mathbf{y})$ is positive semidefinite.
(b) The matrix $M_1'(\mathbf{y})$ is positive semidefinite and for some $j \in V$, $avg_{i\in V} y_{\{i,j\}} \ge y_R^2$.

Proof. It is is to check that $M_1(\mathbf{y}) \succeq 0$ if and only if $M_1'(\mathbf{y}) - y_R^2 J \succeq 0$, where J is the all 1 matrix. Now, since the all 1 vector is an eigenvector for both $M_1'(\mathbf{y}),\ J$, all other eigenvectors of $M_1'(\mathbf{y})$ are perpendicular to the rows of J. By noticing also that this implies that $\sum_{i\in V} y_{\{i,j\}}$ does not depend on $j \in V$, we conclude the claim. In particular, this shows that the value $n\left(\mathrm{avg}_{i\in V} y_{\{i,j\}} - y_R^2\right)$, is an eigenvalue of $M_1'(\mathbf{y}) - y_R^2 J$. ☐

The next Lemma establishes a sufficient condition for solutions fooling SDP relaxations for Borsuk graphs. The proof uses the standard tool of tensoring.

Lemma 6.2. *Let $\mathbf{y}$ be a level-t Sherali-Adams solution for* VERTEX COVER *for a Borsuk graph with vector representation $\mathbf{u}_i$ and suppose that the value $y_{\{i,j\}}$ can be expressed as a function $f(x)$ of the inner product $\mathbf{u}_i \cdot \mathbf{u}_j = x$. Then if $f(x)$ has Taylor expansion with nonnegative coefficients, then the matrix $M_1'(\mathbf{y})$ is positive semidefinite.*

Proof. Consider the Taylor expansion of $f(x) = \sum_{i=0}^{\infty} a_i x^i$, where $a_i \geq 0$. We map $\mathbf{u}_i \in S^{m-1}$ to an infinite dimensional space as follows

$$\mathbf{u}_i \mapsto T_f(\mathbf{u}_i).$$

Then the vectors $T_f(\mathbf{u}_i)$ constitute the Cholesky decomposition of $M_1'(\mathbf{y})$, and therefore $M_1'(\mathbf{y}) \succeq 0$. □

Now we examine the Sherali-Adams solution of some special case that will be instructive for our general argument. Consider some n vertex subgraph $G = (V, E)$ of $B^m_{\rho^2}$ with vector representation $\mathbf{z}_i \in S^{m-1}$. Suppose also that edges $ij \in E$ appear exactly when $\mathbf{z}_i \cdot \mathbf{z}_j = -1 + 2\rho^2$, and that for all other pairs $i, j \in V$ we have $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1 + 2\rho^2$. Run Experiment Local-Global with parameters $t = 2$ and $\delta = \rho^2$ to define the level-2 Sherali-Adams solution $\mathbf{y}$

$$y_I = \mathop{\mathbb{P}}_{\mathbf{w} \in S^1} [\mathbf{w} \cdot \mathbf{z}_i \leq \rho, \ \forall i \in I] \tag{10}$$

for all I of size at most 2, where $\mathbf{w}$ is distributed uniformly on the circle.

Claim 6.3. *The values $y_{\{i,j\}}$ of equations (10) depends on the inner product $\mathbf{z}_i \cdot \mathbf{z}_j = x$ and the rounding parameter ρ. As such, its value is given by*

$$f(x, \rho) = \begin{cases} 1 & , \rho \geq 1 \\ 1 - \frac{2\theta_\rho}{\pi} & , 2\rho^2 \geq x + 1 \ \ \& \ \ \rho \leq 1 \\ 1 - \frac{\theta_\rho}{\pi} - \frac{\theta_x}{2\pi} & , 2\rho^2 \leq x + 1 \ \ \& \ \ \rho \leq 1, \end{cases} \tag{11}$$

from which we also derive that $f(1, \rho) = 1 - \frac{\theta_\rho}{\pi}$.

Proof. The vector $\mathbf{w}$ of (10) can be obtained by taking $W_1, W_2 \sim N(0, 1)$ and then normalizing the vector $\mathbf{w} = (W_1, W_2)$. Next we find an analytic expression for $y_{\{i,j\}}$ that depends on the inner product $\mathbf{z}_i \cdot \mathbf{z}_j = x$ and the rounding parameter ρ. Since this value depends on x, ρ, we denote it by $f(x, \rho)$. Formally,

$$f(x, \rho) := \mathop{\mathbb{P}}_{\mathbf{w} \in S^1} [\mathbf{w} \cdot \mathbf{z}_i \leq \rho \ \& \ \mathbf{w} \cdot \mathbf{z}_j \leq \rho, \text{ with } \mathbf{z}_i \cdot \mathbf{z}_j = x] \ \text{(which is also denoted by } y_{\{i,j\}}) \tag{12}$$

Note that in that language, we have $y_{\{i\}} = f(1, \rho)$. Since $W_1, W_2 \sim N(0, 1)$, the random variable $W_1^2 + W_2^2$ follows the chi-square distribution $\chi^2(2)$ with 2 degrees of freedom. Conditioning on $W_1^2 + W_2^2 = w$, we can normalize the expression in (12) by w so that the probability can be expressed as the ratio of arcs of the unit circle. Call θ_x the angle between $\mathbf{z}_i, \mathbf{z}_j$, namely $\theta_x = \arccos(\mathbf{z}_i \cdot \mathbf{z}_j) = \arccos(x)$. Set also $\theta_\rho = \arccos(\rho)$. In other words, we interpret the event in (12) as saying that two vectors with angle θ_x remain on the unit circle after removing a random cap of measure $\arccos(\rho)$. Note that when $\theta_x \leq 2\arccos(\rho)$, namely when $x \geq 2\rho^2 - 1$, we have $y_{\{i,j\}} = 1 - \frac{\theta_\rho}{\pi} - \frac{\theta_x}{2\pi}$. In contrast, when $x \leq 2\rho^2 - 1$, we have $y_{\{i,j\}} = 1 - 2\frac{\theta_\rho}{\pi}$. □

Fact 6.4. *If $\rho \in [0, 1]$, the function $1 - \frac{\theta_\rho}{\pi} - \frac{\theta_x}{2\pi}$ has Taylor expansion with nonnegative coefficients.*

Proof. (of Fact 6.4) We use the Taylor expansion of $\arccos(x)$ to write the above functions as

$$\frac{3}{4} - \frac{\arccos(\rho)}{\pi} + \frac{1}{2\pi}\sum_{k=0}^{\infty} \frac{(2k)!}{2^{2k}(k!)^2(2k+1)} x^{2k+1}.$$

Since $\rho > 0$, we have $\arccos(\rho) < \pi/2$ and therefore the constant term above is non negative. $\square$

Note here that if we start with a configuration of vectors $\mathbf{z}_i$ for which $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1 + 2\rho^2$ for all pairs $i, j \in V$, then the value $y_{\{i,j\}}$ will be described as a function on the inner product $\mathbf{z}_i \cdot \mathbf{z}_j = x$, and this function on x will have Taylor expansion with nonnegative coefficients. Unfortunately, for our Sherali-Adams solution of the previous sections this is not the case. We establish this extra condition in Section 7, making sure that the principal submatrix $M_1'(\mathbf{y})$ of $M_1(\mathbf{y})$ is positive semidefinite. Proving that the matrix $M_1(\mathbf{y})$ is positive semidefinite will require one extra simple argument, which is self evident from fact 6.1.

7 An easy level-2 Sherali-Adams SDP Solution

In this section we apply the techniques we developed in Section 6 to show a tight integrality gap for VERTEX COVER in the level-2 Sherali-Adams SDP system. This will serve as an instructive example of the general case, whose proof will be obtained as a smooth generalization of the arguments below. In what follows, we show that:

Theorem 7.1. *For every $\epsilon > 0$, there exist $\delta > 0$ and sufficiently big m, such that the objective value of the level-2 Sherali-Adams SDP system for the* VERTEX COVER *polytope and the family of graphs G_δ^m is* $2^m(1/2 + \epsilon)$.

As the theorem states, we start with the Frankl-Rödl graph $G_\delta^m = (V, E)$, which is a subset of B_δ^m, with vector representation $\mathbf{u}_i$. Our goal is to define $\mathbf{y}$ in the context of Theorem 5.2, so as the matrix $M_1(\mathbf{y})$ to be positive semidefinite. Our Sherali-Adams solution as it appears in Theorem 5.2 does not satisfy the constraint $M_1(\mathbf{y}) \succeq 0$, for reasons that will be clear shortly. For this, we need to apply the transformation $\mathbf{u}_i \mapsto \mathbf{z}_i := (\sqrt{\zeta}, \sqrt{1-\zeta}\, T_P(\mathbf{u}_i))$, for some appropriate tensoring polynomial $P(x)$, and some $\zeta > 0$ (that is allowed to be a function of (m, δ)). We will use the following fact, first proved by Charikar [Cha02].

Fact 7.2. *There exist a polynomial $P(x)$, with nonnegative coefficients and $P(1) = 1$, such that for all $x \in [-1, 1]$, we have $P(x) \geq P(-1 + 2\delta) = -1 + 2\delta_0$, for some $\delta_0 = \Theta(\delta)$. Moreover, for every constant $c > 0$ and for every $x \in (-c/\sqrt{m}, c/\sqrt{m})$, we have $|P(x)| = O(\sqrt{1/m})$.*

We use the polynomial P of Fact 7.2 to map the vectors $\mathbf{u}_i$ to the new vectors $\mathbf{z}_i := (\sqrt{\zeta}, \sqrt{1-\zeta}\, T_P(\mathbf{u}_i))$. Note that with this transformation, for an edge $ij \in E$ we have $\mathbf{z}_i \cdot \mathbf{z}_j = \zeta + (1-\zeta)P(-1+2\delta) = \zeta + (1-\zeta)(-1+2\delta_0) = -1 + 2\left(\zeta(1-\delta_0) + \delta_0\right)$. If we denote $\sqrt{\zeta(1-\delta_0) + \delta_0}$ by ρ, then the above transformation maps G_δ^m to $G_{\rho^2}^{m'}$, where m' is the degree of the polynomial P. We are therefore eligible to run Experiment Local-Global with parameters $t = 2$ and ρ^2 on the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\, T_P(\mathbf{u}_i))$. Then Lemma 4.1, together with Fact 3.1, imply that $\mathbf{y}$ as defined in (10) is a level-2 Sherali-Adams solution (the parameters δ, ζ will be fixed later). Next we show that for a slightly perturbed $\mathbf{y}$ we have that $M_1(\mathbf{y})$ is positive semidefinite.

First we observe that the context of Section 6 is relevant to the current configuration of vectors $\mathbf{z}_i$ and to our graph instances, since $\mathbf{z}_i \cdot \mathbf{z}_j \geq -1 + 2\rho^2$. If $\mathbf{u}_i \cdot \mathbf{u}_j = x$, then the value of $y_{\{i,j\}}$ is exactly $g(\zeta + (1-\zeta)P(x))$, where $g(x) = 1 - \frac{\theta_\rho}{\pi} - \frac{\arccos(x)}{2\pi}$. By Fact 6.4 we know that the function $g(x)$ has

Taylor expansion with nonnegative coefficients. Since $\zeta + (1-\zeta)P(x)$ is a polynomial with nonnegative coefficients, it follows that $g(\zeta + (1-\zeta)P(x))$ has Taylor Expansion with nonnegative coefficients. Hence, we can apply Lemma 6.2 to obtain that

Lemma 7.3. *The matrix $M_1'(\mathbf{y})$ is positive semidefinite.*

In what follows we describe a way to extend the positive semidefinteness of $M_1'(\mathbf{y})$ to that of $M_1(\mathbf{y})$. In fact what we will show is general and holds for any level t. Since the entries of $M_1'(\mathbf{y})$ are a function the inner product of the corresponding vectors of the hypercube, it follows that the all 1 vector is an eigenvector for $M_1'(\mathbf{y})$. By Fact 6.1 it follows that we need to show that $\text{avg}_{i\in V} y_{\{i,j\}} - y_{\{i\}}^2 \geq 0$. It turns out that this is not the case, but we can establish a weaker condition (described here in terms of a general sphere dimension D).

Lemma 7.4. *There exist a constant $c > 0$ (not depending on m, ρ), such that $\text{avg}_{i\in V} y_{\{i,j\}} - y_{\{i\}}^2 \geq -cD\rho$.*

We omit the proof of this lemma from this extended abstract. A rough estimate that suffices is that whenever two points have positive inner product, the probability that both are in a random cap is at least $1/4$. It can be shown that due to the affine transformation, all but exponentially small fraction of the pairs will have positive inner products, hence we get that the average of $y_{\{i,j\}}$ is at least $1/4 - o(1)$. On the other hand, by Section 5 we know that $y_{\{i\}} \leq 1/2 + O(D\rho)$.

Boosting: It remains to show how to "boost" the solution to move from the relaxed condition to the exact, and necessary one. The idea is simple. Consider a ridiculously wasteful integral solution to Vertex Cover, namely the solution that takes all vertices. Clearly, if we take a convex combination of this solution with the existing one we still get a Sherali-Adams solution. If the weight of the integral solution is some small number $\xi > 0$ then the objective value increases by no more than $\xi/2$ which can be absorbed for arguments to go through as long as $\xi \leq \epsilon$. Owing to the strict convexity of the quadratic function, however, this simple perturbation does allow to improve the bound on averages as we describe now and prove in the Appendix.

Lemma 7.5. *Let y' be the matrix $y' = (1-\xi)y + \xi J$ where J represents the all 1 solution. Also let $s = y_{\{i\}}$ and $s' = y'_{\{i\}}\}$. Then $\text{avg}_{i,j} y'_{\{i,j\}} - s'^2 = \Omega(\xi)$.*

We are now ready to formally prove Theorem 7.1.

Proof. (of Theorem 7.1) We start with the n-vertex Frankl-Rödl graph G_δ^m, with $\delta = \theta(\sqrt{\log n/\log\log n})$ so as to satisfy the conditions of Theorem 2.3. We use the polynomial of Fact 7.2 to obtain the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\; T_P(\mathbf{u}_i))$, with $\zeta = \delta_0$ (where $\delta_0 = \Theta(\delta)$ by Fact 7.2). We set $\rho = \sqrt{\zeta(1-\delta_0)+\delta_0} = \sqrt{\Theta(\zeta)}$, and we run the Experiment Local-Global on the vectors $\mathbf{z}_i$ with parameters $t = 2$ and ρ^2, to obtain the vector $\mathbf{y}$. By Lemma 4.1 and Fact 3.1, we have that $\mathbf{y}$ as defined in (10) is a level-2 Sherali-Adams solution. Note that since $\delta = o(1)$ we conclude from Lemma 5.1 that $y_{\{i\}} = 1/2 + \Theta(\delta) < 2/3$.

Next we define $\mathbf{y}'$ as $(1-\xi)y + \xi J$. We already argued that $M_1'(\mathbf{y}')$ is positive semidefinite. By the above discussion (and Lemma 7.5) we conclude that $\text{avg}_{i,j} {y'}_{\{i,j\}} - {y'}_{\{i\}}^2 \geq 0$. We can therefore use Fact 6.1 to conclude that $M(\mathbf{y}')$ is positive semidefinite. The last thing to note is that the contribution of every vertex in the objective value is $1/2 + O(\delta)$ □

8 The level-$(t+2)$ Sherali-Adams SDP Tight Integrality Gap

For the level-$(t+2)$ SDP, we start with the n-vertex Frankl-Rödl graphs G_δ^m, $n = 2^m$ with vector representation $\mathbf{u}_i$. The value of δ is chosen so as to satisfy Theorem 2.3, namely $\delta = \Theta(\sqrt{\log m/m})$. As in Section 7 we apply to $\mathbf{u}_i$ two transformations; one using the tensoring polynomial of Fact 7.2 and one affine transformation. Then we use the resulting vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\, T_P(\mathbf{u}_i))$ to define a level-$(t+2)$ Sherali-Adams solution that we denote by $\mathbf{y}$.

Our goal is to meet the conditions of Fact 6.1. Namely, the first thing to ensure is that $M_1'(\mathbf{y})$ is positive semidefinite. In this direction, from Lemma 6.2 it suffices to show that the Taylor expansion of the function that describes the value of $y_{\{i,j\}}$, when $\mathbf{u}_i \cdot \mathbf{u}_j = u$, has Taylor expansion with nonnegative coefficients. Given that this function at 0 will always represent some probability, the problem is equivalent to showing that the first derivative of this function has such a good Taylor expansion. Our transformation on the vectors $\mathbf{u}_i$ can be thought as mapping their inner product u first to $x = P(u)$, and second x to $y(x) = \zeta + (1-\zeta)x$. Under this notation, we can show the following lemma that involves a number of technical calculations. We deal with its proof in Section 10 in the Appendix.

Lemma 8.1. *The first derivative of the functional description of $y_{\{i,j\}}$ is*

$$D_\zeta(x) := -(\arccos(y(x)))' \left(\frac{1-2\rho^2+y(x)}{1+y(x)}\right)^{t/2},$$

where the subscript of D_ζ we emphasize that $y(x) = \zeta + (1-\zeta)x$.

Therefore, to conclude that $M_1'(\mathbf{y})$ is positive semidefinite it suffices to show the next technical lemma,

Lemma 8.2. *For $t \le 3$ and for $\rho^2 < 1.01\zeta$, the function $D_\zeta(x)$ as it reads in Lemma 8.1 has Taylor expansion with nonnegative coefficients.*[3]

whose proof requires arguments along the lines of our proof for Claim 7.3. Due to its technicality, Lemma 8.2 is proved in the Appendix in Section 11.

Now we obtain a level-$(t+2)$ Sherali-Adams solution from the vectors $\mathbf{z}_i = (\sqrt{\zeta}, \sqrt{1-\zeta}\, T_P(\mathbf{u}_i))$, namely we prove Theorem 1.2 (that requires $t = 3$). We need to set $\zeta = 1000\delta_0$, where $\delta_0 = (1 + \min(P(x))/2$. Since the rounding parameter we need is $\rho = \sqrt{\zeta(1-\delta_0)+\delta_0}$, it is easy to see that $\rho^2 < 1.01\zeta$. It follows by Lemma 8.2 that the matrix $M_1'(\mathbf{y})$ is positive semidefinite.

Now call c the constant for which $\text{avg}_{i\in V} y_{\{i,j\}} - y_{\{i\}}^2 \ge -ct\rho^2$. We also know that if $t\rho^2$ is no more than a small constant $\epsilon/10$, then $y_{\{i\}} \le 1/2 + \epsilon$. Then define

$$\mathbf{y}' = (1-4c\epsilon)\mathbf{y} + (4c\epsilon)\mathbf{1}.$$

As we did for the level-2 Sherali-Adams SDP solution, the vector $\mathbf{y}'$ is a level-$(t+2)$ Sherali-Adams solution. Moreover, the matrix $M_1'(\mathbf{y}')$ is positive semidefinite, and $\text{avg}_{i\in V} y'_{\{i,j\}} - y'^2_{\{i\}} \ge 0$. All conditions of Fact 6.1 are satisfied implying that $M_1(\mathbf{y}')$ is positive semidefinite. Finally, note that the contribution of the singletons is no more than $1/2 + \Theta(ct\rho^2)$. Hence, if we start[4] with $t\rho^2 = o(1)$, the contribution of the singletons remains $1/2 + o(1)$. On the other hand, choosing $\delta = \Theta(\sqrt{\log m/m})$ results in graphs G_δ^m with no vertex covers smaller than $n - o(n)$. We should point here that the limitations in the level of our integrality gap are set by Lemma 8.2, which is already messy to prove. We dare to conjecture that Lemma 8.2 holds true for superconstant values of t, for which we also have strong numerical evidence. This further allows us to conjecture that the exact same solution we have can prove that

Conjecture 8.3. *For every $\epsilon, \epsilon_0 > 0$ and for $t = (\sqrt{\log n/\log\log n})^{1-\epsilon_0}$, there exist a family of graphs on n vertices, such that the integrality gap of the level-t Sherali-Adams SDP relaxation for* VERTEX COVER *has integrality gap $2 - o(1)$.*

[3]We have numerical evidence that the lemma holds true as long as $t = o(\zeta)$ and $\rho^2 < 1.4\zeta$, however we were not able to show this stronger version. In particular, this lemma sets the restrictions on the rounds of the Sherali-Adams SDPrelaxation.

[4]It even suffices to have $t = (\sqrt{\log n/\log\log n})^{1-\epsilon_0}$.

References

[ARKN10] Claire Mathieu Anna R. Karlin and C. Thach Nguyen. Integrality gaps of linear and semi-definite programming relaxations for knapsack. Manuscript, 2010. 3

[ARV09] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):1–37, 2009. 3

[BCG09] Mohammad Hossein Bateni, Moses Charikar, and Venkatesan Guruswami. MaxMin allocation via degree lower-bounded arborescences. In *STOC'09: Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 543–552, New York, NY, USA, 2009. ACM Press. 3

[BGMT10] Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. Sdp gaps from pairwise independence. submitted manuscript available at http://www.cs.toronto.edu/∼siavosh/pdf/pairwise-full-1.pdf, 2010. 2

[BM10] Siavosh Benabbas and Avner Magen. Extending SDP integrality gaps to Sherali-Adams with applications to quadratic programming and MaxCutGain. In *IPCO'10: Proceedings of the 14th International Conference on Integer Programming and Combinatorial Optimization*, volume 6080 of *LNCS*, pages 299–312. Springer, 2010. 2

[Cha02] Moses Charikar. On semidefinite programming relaxations for graph coloring and Vertex Cover. In *SODA'02: Proceedings of the 13th annual ACM-SIAM symposium on Discrete algorithms*, pages 616–620, Philadelphia, PA, USA, 2002. ACM Press. 5, 6, 12

[CMM09] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations. In *STOC'09: Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 283–292, New York, NY, USA, 2009. ACM Press. 2, 5, 8

[CS08] Eden Chlamtac and Gyanit Singh. Improved approximation guarantees through higher levels of SDP hierarchies. In *APPROX'08/RANDOM'08: Proceedings of the 11th International Workshop on Approximation and the 12th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 49–62, Berlin, Heidelberg, 2008. Springer-Verlag. 3

[dlVKM07] Wenceslas Fernandez de la Vega and Claire Kenyon-Mathieu. Linear programming relaxations of Max Cut. In Nikhil Bansal, Kirk Pruhs, and Clifford Stein, editors, *SODA'07 Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 53–61. ACM Press, 2007. 3, 8

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

[FR87] Peter Frankl and Vojtech Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987. 4, 6

[GK98] Michel X. Goemans and Jon Kleinberg. The Lovász theta function and a semidefinite programming relaxation of Vertex Cover. *SIAM J. Discrete Math.*, 11(2):196–204 (electronic), 1998. 5, 6

[GM08] Konstantinos Georgiou and Avner Magen. Expansion fools the Sherali-Adams system: Compromising local and global arguments. Technical Report CSRG-587, University of Toronto, November 2008. 8

[GMPT07] Konstantinos Georgiou, Avner Magen, Toniann Pitassi, and Iannis Tourlakis. Integrality gaps of $2-o(1)$ for Vertex Cover SDPs in the Lovász-Schrijver hierarchy. In *FOCS'07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 702–712. IEEE Computer Society, 2007. 3, 4, 5, 6

[GMT08] Konstantinos Georgiou, Avner Magen, and Iannis Tourlakis. Vertex Cover resists SDPs tightened by local hypermetric inequalities. In Andrea Lodi, Alessandro Panconesi, and Giovanni Rinaldi, editors, *IPCO'08: Proceedings of the 13th International Conference on Integer Programming and Combinatorial Optimization*, volume 5035 of *Lecture Notes in Computer Science*, pages 140–153. Springer, 2008. 6

[GMT09a] Konstantinos Georgiou, Avner Magen, and Iannis Tourlakis. On the tightening of the standard SDP for Vertex Cover with ℓ_1 inequalities. In Ravi Kannan and K Narayan Kumar, editors, *FSTTCS'09: Proceedings of the 29th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 4 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 203–214, Dagstuhl, Germany, 2009. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. 5, 6

[GMT09b] Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. Optimal Sherali-Adams gaps from pairwise independence. In *APPROX'09/RANDOM'09: Proceedings of the 12th International Workshop on Approximation and the 13th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 5687 of *Lecture Notes in Computer Science*, pages 125–139. Springer, 2009. 8

[GW95] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for Maximum Cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995. 3

[Kar09] George Karakostas. A better approximation ratio for the Vertex Cover problem. *ACM Trans. Algorithms*, 5(4):1–8, 2009. 2, 3

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *STOC'02: Proceedings of the 34th annual ACM symposium on Theory of computing*, pages 767–775, New York, 2002. ACM Press. 2

[KS09] Subhash Khot and Rishi Saket. SDP integrality gaps with local ℓ_1-embeddability. In *FOCS'09: Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 565–574. IEEE Computer Society, 2009. 2

[KZ97] Howard Karloff and Uri Zwick. A 7/8-approximation algorithm for MAX 3SAT? In *FOCS'07: Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science*, pages 406–415. IEEE Computer Society, 1997. 3

[Lau03] Monique Laurent. A comparison of the Sherali-Adams, Lovász-Schrijver, and Lasserre relaxations for 0-1 programming. *Math. Oper. Res.*, 28(3):470–496, 2003. 8

[LS91] László Lovász and Alexander Schrijver. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, May 1991. 7

[MM09] Avner Magen and Mohammad Moharrami. Robust algorithms for maximum independent set on minor-free graphs based on the Sherali-Adams hierarchy. In Irit Dinur, Klaus Jansen, Joseph Naor, and José D. P. Rolim, editors, *APPROX'09/RANDOM'09: Proceedings of the 12th International Workshop on Approximation and the 13th International Workshop on Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 5687 of *Lecture Notes in Computer Science*, pages 258–271. Springer, 2009. 3

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *STOC'08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 245–254. ACM Press, 2008. 3

[RS09] Prasad Raghavendra and David Steurer. Integrality gaps for strong SDP relaxations of unique games. In *FOCS'09: Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 575–585. IEEE Computer Society, 2009. 2

[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 J. Discrete Math.*, 3(3):411–430, 1990. 7

[Sch08] Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *FOCS'08: Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602, Washington, DC, USA, 2008. IEEE Computer Society. 3

[STT07] Grant Schoenebeck, Luca Trevisan, and Madhur Tulsiani. A linear round lower bound for Lovász-Schrijver SDP relaxations of Vertex Cover. In *CCC'07: Proceedings of the 22nd Annual IEEE Conference on Computational Complexity*, pages 205–216. IEEE Computer Society, 2007. 2

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *STOC'09: Proceedings of the 41th annual ACM symposium on Theory of computing*, pages 303–312, New York, NY, USA, 2009. ACM Press. 3

[Vis09] Sundar Vishwanathan. On hard instances of approximate Vertex Cover. *ACM Trans. Algorithms*, 5(1):Art. 7, 6, 2009. 3

Appendix

9 Proofs of Lemmata

Proof. (of Lemma 5.1) A vector $\mathbf{w} \in S^{t-1}$ defines the complement of a spherical cap of height $1-\eta$. We therefore need to determine the ratio between the measure of a cap over the measure of the t dimensional sphere.

Denote by $\mu(S^{t-1}_\eta)$ the surface area of a spherical cap of S^{t-1}, with height $1-\eta$ (as the one defined by $\mathbf{w}$ in the statement of the lemma). Let also

$$\Phi_t\left(x,y\right) := \int_x^y \sin^{t-2} r dr,$$

for which the following bounds hold

$$\sqrt{\frac{2\pi}{t+1}} \leq \Phi_t\left(0,\pi\right) \leq \sqrt{\frac{2\pi}{t}}. \tag{13}$$

We will need the following fact

$$\mu(S^{t-1}_\eta) = 2\pi\ \Phi_{t-2}\left(0, \arccos\eta\right)\prod_{i=1}^{t-3}\Phi_i\left(0,\pi\right).$$

Note that in this language, the surface area of the hypersphere is $2\ \mu(S^{t-1}_0)$. We therefore have

$$\begin{aligned}
\mathop{\mathbb{P}}_{\mathbf{w}\in S^{t-1}}[\mathbf{w}\cdot\mathbf{z}\leq\eta] &= \frac{2\mu(S^{t-1}_0)-\mu(S^{t-1}_\eta)}{2\mu(S^{t-1}_0)} \\
&= \frac{1}{2}+\frac{\mu(S^{t-1}_0)-\mu(S^{t-1}_\eta)}{2\mu(S^{t-1}_0)} \\
&= \frac{1}{2}+\frac{\Phi_{t-2}\left(0,\pi/2\right)-\Phi_{t-2}\left(0,\arccos\eta\right)}{2\Phi_{t-2}\left(0,\pi/2\right)} \\
&= \frac{1}{2}+\frac{\Phi_{t-2}\left(\arccos\eta,\pi/2\right)}{\Phi_{t-2}\left(0,\pi\right)} \\
&\overset{(13)}{\leq} \frac{1}{2}+\sqrt{\frac{t+1}{2\pi}}\Phi_{t-2}\left(\arccos\eta,\pi/2\right) \\
&\leq \frac{1}{2}+\sqrt{\frac{t+1}{2\pi}}(\pi/2-\arccos\eta).
\end{aligned} \tag{14}$$

Now we need the Taylor expansion of $\arccos\eta$, according to which

$$\arccos\eta = \frac{\pi}{2}-\sum_{k=0}^{\infty}\frac{(2k)!}{2^{2k}(k!)^2(2k+1)}\eta^{2k+1} \geq \frac{\pi}{2}-\eta\sum_{k=0}^{\infty}\frac{(2k)!}{2^{2k}(k!)^2(2k+1)} = \frac{\pi}{2}-\eta\frac{\pi}{2} \tag{15}$$

We can therefore upper bound (14) by

$$\frac{1}{2}+\eta\frac{\pi}{2}\sqrt{\frac{t+1}{2\pi}} = \frac{1}{2}+\eta\sqrt{\frac{\pi}{8}(t+1)}.$$

□

Proof. (of Lemma 7.5)

$$\begin{aligned}
\mathrm{avg}_{i,j}y'_{\{i,j\}}-s'^2 &= (1-\xi)\mathrm{avg}_{i,j}y_{\{i,j\}}+\xi-((1-\xi)s+\xi)^2 \\
&= (1-\xi)\mathrm{avg}_{i,j}A_{\{i,j\}}+\xi-(1-\xi)^2s^2-2(1-\xi)\xi s-\xi^2 \\
&\geq (1-\xi)(\mathrm{avg}_{i,j}A_{\{i,j\}}-s^2)+\xi+\xi s^2-\xi^2s^2-2\xi s-\xi^2 \\
&\geq -O(D\rho+\xi^2)+(1-s)^2\xi \\
&= \Omega(\xi)
\end{aligned}$$

The last inequality we use the fact that we only take D so as to make $D\rho = O(\epsilon)$, that ξ is roughly the same as ϵ, and that s is bounded away from 1. □

10 A Functional Description of the Sherali-Adams Solution (proof of Lemma 8.1)

Our goal is to prove the analog of Theorem 7.1 for a higher level of the Sherali-Adams system, satisfying the positive semidefinite constraint as well. We start with the configuration of Section 6, namely two vectors $\mathbf{z}_i, \mathbf{z}_j$ having inner product $x \geq -1 + 2\rho^2$. We now revisit (12) and (10) taking into consideration that our experiment takes place in S^{t+1}, $t \geq 0$ (this will give rise to a level-$(t+2)$ Sherali-Adams solution). As in the construction of the level-2 Sherali-Adams SDP solution, we give a functional description of the value of the doubletons $y_{\{i,j\}}$. Then we only need to show that the Taylor expansion of this function has nonnegative coefficients.

Remember that the experiment can be realized by considering $t+2$ independent random variables W_i chosen from the normal distribution. Note that the random variables $\sum_{i=3}^{t} W_i^2 := U_1$ and $W_1^2 + W_2^2 := U_2$ follow the distributions $\chi^2(t)$ and $\chi^2(2)$ respectively. We will need the following fact.

Fact 10.1. *Let U_1, U_2 be two stochastic independent random variables of the distributions $\chi^2(t)$ and $\chi^2(2)$ respectively. Then the random variable $\frac{U_1/t}{U_2/2}$ has probability density function*

$$g_{t,2}(u) = \frac{t}{u}\sqrt{\frac{(tu)^t}{(tu+2)^{t+2}}}.$$

Moreover, the associated distribution is known as the $F_{t,2}$-distribution.

Conditioning on $U_1 = u_1,\ U_2 = u_2$, as in Section 6, we can normalize the events by u_2, in which case we can interpret the probability of the event associated with $y_{\{i,j\}}$ again in terms of an experiment that takes place in S^1. The difference is that this time, the rounding parameter becomes $\rho\sqrt{(1+u_1/u_2)}$. We just showed the following fact.

Fact 10.2. *Let $y_{\{i_1,i_2\}} = \mathbb{P}_{\mathbf{w}\in S^{t+1}}[\mathbf{w}\cdot\mathbf{z}_i \leq \rho,\ i = 1,2]$, namely the values $y_{\{i_1,i_2\}}$ are the result of a random cut that takes place in S^{t+1}. The same cut is a distribution of cuts in S^1 with rounding parameter $\eta = \rho\sqrt{1-tu/2}$. The probability density function of u is the $F_{t,2}$-distribution of Fact 10.1.*

For two vectors with $\mathbf{z}_i \cdot \mathbf{z}_j = x$, we can adjust the notation of (10) so as to define $F_t(x,\rho) = y_{\{i,j\}}$. Using now (11) we have

$$F_t(x,\rho) = \int_{u=0}^{\infty} g_{t,2}(u) f(x, \rho\sqrt{1+tu/2}) du. \tag{16}$$

Now define

$$C := 2\frac{1-\rho^2}{t\rho^2},\ \ h(x) := \frac{1-2\rho^2+x}{t\rho^2}$$

emphasizing that C does not depend on x. Then taking advantage of (11), we have

$$f(x,\rho\sqrt{1+tu/2}) = \begin{cases} 1 & , u \geq C \\ 1 - \frac{2\theta(\rho\sqrt{1+tu/2})}{\pi} & , h(x) \leq u \leq C, \\ 1 - \frac{\theta(\rho\sqrt{1+tu/2})}{\pi} - \frac{\theta_x}{2\pi} & , 0 \leq u \leq h(x) \end{cases} \tag{17}$$

Note that this is well defined, since $\rho > 0$ and for all $x \in [-1+2\rho^2, 1]$ we have $0 \leq h(x) \leq C$ (recall that x represents the inner product of unit vectors, that have always angle at most $2\arccos(\rho)$).

This allows us to find a nicer expression for (16). For notational simplicity, let $\eta = \rho\sqrt{1+tu/2})$. Then,

$$\begin{aligned} F_t(x,\rho) &= \int_{u=C}^{\infty} g_{t,2}(u)du + \int_{u=0}^{h(x)} g_{t,2}(u)\left(1 - \frac{\theta_\eta}{\pi} - \frac{\theta_x}{2\pi}\right)du + \int_{u=h(x)}^{C} g_{t,2}(u)\left(1 - \frac{2\theta_\eta}{\pi}\right)du \\ &= 1 - \frac{1}{\pi}\int_{u=0}^{C} g_{t,2}(u)\theta_\eta du + \frac{1}{2\pi}\left(-\theta_x \int_{u=0}^{h(x)} g_{t,2}(u)du - 2\int_{u=h(x)}^{C} g_{t,2}(u)\theta_\eta du\right) \end{aligned}$$

where the component of $F_t(x,\rho)$ that depends on x now becomes clear.

Next we find the first partial derivative of $F_t(x,\rho)$, with respect to x. For this define

$$H_1(x) := \theta_x \int_{u=0}^{h(x)} g_{t,2}(u)du, \qquad H_2(x) := \int_{u=h(x)}^{C} g_{t,2}(u)\theta_\eta du,$$

Now observe that

$$\frac{\partial}{\partial x}H_1(x) = (\arccos(x))' \int_{u=0}^{h(x)} g_{t,2}(u)du + \arccos(x)\, g_{t,2}(h(x))\frac{\partial h(x)}{\partial x}$$

and that

$$\begin{aligned} \frac{\partial}{\partial x}H_2(x) &= -g_{t,2}(h(x))\ \arccos\left(\rho\sqrt{1+\frac{t}{2}h(x)}\right)\ \frac{\partial h(x)}{\partial x} \\ &= -g_{t,2}(h(x))\ \arccos\left(\sqrt{\frac{1+x}{2}}\right)\ \frac{\partial h(x)}{\partial x} \\ &= -\frac{1}{2}\, g_{t,2}(h(x))\ \arccos(x)\ \frac{\partial h(x)}{\partial x} \end{aligned}$$

where the last equality follows from the identity $\arccos\sqrt{(1+x)/2} = \arccos(x)/2$. Combining the above, we conclude that

$$\begin{aligned} \frac{\partial}{\partial x}F_t(x,\rho) &= -(\arccos(x))' \int_{u=0}^{h(x)} g_{t,2}(u)du \\ &= -(\arccos(x))' \left[\left(\frac{tu}{tu+2}\right)^{t/2}\right]_0^{h(x)} \\ &= -(\arccos(x))' \left(\frac{t\,h(x)}{t\,h(x)+2}\right)^{t/2} \\ &= -(\arccos(x))' \left(\frac{1-2\rho^2+x}{1+x}\right)^{t/2}. \end{aligned} \tag{18}$$

The proof of Lemma 8.1 is now complete.

11 Positive coefficients of the Taylor Expansion (proof of Lemma 8.2)

We show that the Taylor expansion of

$$D_\zeta(x) = -(\arccos(y(x)))' \left(\frac{1-2\rho^2+y(x)}{1+y(x)}\right)^{t/2} \tag{19}$$

where $y(x) = \zeta + (1-\zeta)x$, has strictly positive coefficients. Taking into consideration that $-(\arccos(y(x)))' = \zeta/\sqrt{1-y^2(x)}$, the above can be rewritten as

$$\begin{aligned}
& \frac{\zeta}{\sqrt{1-y^2(x)}}\left(\frac{1-2\rho^2+y(x)}{1+y(x)}\right)^{t/2} \\
= \quad & \frac{\zeta}{\sqrt{1-y(x)}}\frac{\left(1-2\rho^2+y(x)\right)^{t/2}}{(1+y(x))^{t/2+1/2}} \\
= \quad & \frac{\zeta}{\sqrt{1-\zeta-(1-\zeta)x}}\frac{\left(1-2\rho^2+\zeta+(1-\zeta)x\right)^{t/2}}{(1+\zeta+(1-\zeta)x)^{t/2+1/2}} \\
= \quad & \frac{\zeta(1-2\rho^2+\zeta)^{t/2}}{\sqrt{1-\zeta}(1+\zeta)^{t/2+1/2}}\frac{1}{\sqrt{1-x}}\frac{\left(1+\frac{1-\zeta}{1-2\rho^2+\zeta}x\right)^{t/2}}{\left(1+\frac{1-\zeta}{1+\zeta}x\right)^{t/2+1/2}} \\
= \quad & \frac{\zeta(1-2\rho^2+\zeta)^{t/2}}{\sqrt{1-\zeta}(1+\zeta)^{t/2+1/2}}\frac{1}{\sqrt{1-x^2}}\frac{\left(1+\frac{1-\zeta}{1-2\rho^2+\zeta}x\right)^{t/2}\sqrt{1+x}}{\left(1+\frac{1-\zeta}{1+\zeta}x\right)^{t/2+1/2}}
\end{aligned} \tag{20}$$

For the shake of convenience, we apply a simplification. We remind the reader that $\rho^2 = \zeta(1-\delta_0)+\delta_0$, where $\delta_0 = \Omega(\sqrt{\log n/\log\log n})$ is the parameter of the Frankl-Rödl graph on n vertices. Parameter ζ can be chosen so as $\zeta = o(\delta_0)$ and hence, $\rho^2 = (1+o(1))\zeta$. In what follows, we think of ρ^2 equal ζ, in which case (20) can be written as

$$\frac{1}{\sqrt{1-x^2}}\frac{(1+x)^{t/2+1/2}}{\left(1+\frac{1-\zeta}{1+\zeta}x\right)^{t/2+1/2}}$$

after we omit the constants. For convenience, denote $\frac{1-\zeta}{1+\zeta}$ by $1-\epsilon$, and $t/2+1/2$ by $\tau/2$. What we need to show is that the function

$$f_{\tau,\epsilon}(x) := \frac{1}{\sqrt{1-x^2}}\left(\frac{1+x}{1+(1-\epsilon)x}\right)^{\tau/2} \tag{21}$$

has Taylor expansion with positive coefficients. In what follows, we will need the sequence

$$\alpha_j^{(\tau)} := \binom{j-(\tau-1)/\tau}{j},$$

for which the following properties can be proven.

Lemma 11.1.

$$\sum_{j=0}^{D-1}\alpha_j^{(\tau)} = \tau D\alpha_D^{(\tau)}$$

$$\alpha_j^{(\tau)}/\alpha_{j+1}^{(\tau)} = \frac{j+1}{j+1/\tau}$$

$$\alpha_j^{(\tau)} \approx \Gamma\left(\frac{1}{\tau}\right)\frac{1}{j^{(\tau-1)/\tau}} \textit{ for big enough values of } j.$$

11.1 The case $t = 1$ ($\tau = 2$)

We show that the function $f_{2,\epsilon}(x)$, namely

$$\frac{1}{\sqrt{1-x^2}}\left(1+\epsilon\frac{x}{1+(1-\epsilon)x}\right),$$

has Taylor expansion with positive coefficients. First we observe that

$$\frac{1}{\sqrt{1-x^2}} = \sum_{j=0}^{\infty}\binom{j-1/2}{j}x^{2j} = \sum_{j=0}^{\infty}\alpha_j^{(2)}x^{2j}, \tag{22}$$

using the sequence $\alpha_j^{(\tau)}$ we defined above. Also,

$$\frac{1}{1+(1-\epsilon)x} = \sum_{j=0}^{\infty}\binom{-1}{j}(1-\epsilon)^j x^j = \sum_{j=0}^{\infty}(-1)^j(1-\epsilon)^j x^j. \tag{23}$$

Note that $f_{1,\epsilon}(x)$ is the product of two relatively simple functions. The factor (30) contributes only monomials of even degrees, all with positive coefficients. The second factor, namely the function $1+\epsilon\frac{x}{1+(1-\epsilon)x}$ has it's 0th term and all odd degree monomials with positive coefficients, and the rest of even degree monomials with negative coefficients. It follows that in the Taylor expansion of $f_{1,\epsilon}(x)$, all monomials of odd degree have strictly positive coefficients. It therefore suffices to show that the even degree monomials have positive coefficients as well.

Taking into consideration the convolution of (30) and (23), the coefficient of the monomial x^{2D} in the Taylor expansion of $f_{2,\epsilon}(x)$ equals

$$\alpha_D^{(2)} - \epsilon\sum_{j=0}^{D-1}\alpha_j^{(2)}(1-\epsilon)^{2D-1-2j}. \tag{24}$$

Therefore, it suffices to show that (24) is non negative.

For simplicity and for the remaining of this section where $\tau = 1$, we drop the superscript of $\alpha_j^{(2)}$ which we denote by α_j.

11.1.1 The subcase $D \leq \frac{1}{2\epsilon}$

We observe that

$$\sum_{j=0}^{D-1}\alpha_j(1-\epsilon)^{2D-1-2j} \overset{Lemma\ (11.1)}{\leq} (1-\epsilon)2D\alpha_D$$

Then, expression (24) is at least

$$\alpha_D\left(1-2\epsilon D(1-\epsilon)\right),$$

which is positive as long as $D \leq 1/2\epsilon$.

11.1.2 The subcase $\frac{1}{2\epsilon} < D \leq \frac{7}{10\epsilon}$

Let $D = c_2/\epsilon$, with $c_2 \leq 7/10$ and $c_1 := c_2/2$ (note that $c_1 \leq 0.35$). Then we have

$$\begin{aligned}\sum_{j=0}^{D-1} \alpha_j (1-\epsilon)^{2D-1-2j} &= \sum_{j=0}^{c_1/\epsilon-1} \alpha_j (1-\epsilon)^{2D-1-2j} + \sum_{j=c_1/\epsilon}^{c_2/\epsilon-1} \alpha_j (1-\epsilon)^{2D-1-2j} \\ &\leq \alpha_{c_1/\epsilon} \frac{2c_1}{\epsilon} (1-\epsilon)^{(2c_2-2c_1)/\epsilon} + \alpha_{c_1/\epsilon} \frac{c_2-c_1}{\epsilon}\end{aligned}$$

Since $\frac{1}{2\epsilon} < D < \frac{7}{10\epsilon}$ we then conclude that the coefficient of the monomial x^{2D} is

$$\begin{aligned}\alpha_D - \epsilon \sum_{j=0}^{D-1} \alpha_j (1-\epsilon)^{2D-1-2j} &\geq \alpha_{c_1/\epsilon} \left(\frac{\alpha_{c_2/\epsilon}}{\alpha_{c_1/\epsilon}} - 2c_1 (1-\epsilon)^{2(c_2-c_1)/\epsilon} - (c_2 - c_1) \right) \\ &\approx \alpha_{c_1/\epsilon} \left(\sqrt{\frac{c_1}{c_2}} - 2c_1 e^{-2(c_2-c_1)} - (c_2 - c_1) \right) \\ &= \alpha_{c_1/\epsilon} \left(\frac{\sqrt{2}}{2} - 2c_1 e^{-2c_1} - c_1 \right).\end{aligned}$$

It is easy it check that the last expression is positive as long as $c_1 \leq 0.35$.

11.1.3 The subcase $D \geq 7/10\epsilon$

Let $c_1 < 1/2 < c_2 < 1$ be some constants that will be fixed later. As before we have

$$\sum_{j=0}^{c_1/\epsilon-1} \alpha_j (1-\epsilon)^{2D-1-2j} \leq \frac{2c_1}{\epsilon} \alpha_{c_1/\epsilon} (1-\epsilon)^{2D-1-2c_1/\epsilon} \tag{25}$$

$$\sum_{j=c_2/\epsilon}^{c_2/\epsilon-1} \alpha_j (1-\epsilon)^{2D-1-2j} \leq \frac{c_2-c_1}{\epsilon} \alpha_{c_1/\epsilon} (1-\epsilon)^{2D-1-2c_2/\epsilon} \tag{26}$$

Next, we observe that the term $\alpha_j (1-\epsilon)^{2D-1-2j}$ is increasing for sufficiently big j. In particular, one can show that

$$j > \frac{1}{4\epsilon - 2\eta} \quad \Rightarrow \quad \frac{\binom{j+1/2}{j+1} (1-\epsilon)^{2D-1-2j}}{\binom{j-1/2}{j} (1-\epsilon)^{2D+1-2j}} > 1 + \eta,$$

for any $\eta = c_0 \epsilon$, with $c_0 < 2$. Then, for $c_2 \geq \frac{\epsilon}{4\epsilon - 2\eta}$ we have

$$\begin{aligned}\sum_{j=c_2/\epsilon}^{D-1} \binom{j-1/2}{j} (1-\epsilon)^{2D-1-2j} &\leq \sum_{j=c_2/\epsilon}^{D-1} \frac{1}{(1+\eta)^{D-1-j}} \binom{D-1/2}{D} (1-\epsilon) \\ &\leq \binom{D-1/2}{D} \sum_{j=c_2/\epsilon}^{D-1} \frac{1}{(1+\eta)^{D-1-j}} \\ &\leq \binom{D-1/2}{D} \frac{1 + \eta - (1+\eta)^{-D+1+c_2/\epsilon}}{\eta}\end{aligned} \tag{27}$$

Then using (25), (26), (27), and taking into consideration that $D = k/\epsilon$, for some $k \geq 7/10$, we can write for the coefficient of the monomial x^{2D} that

$$\begin{aligned}
&\alpha_D - \epsilon \sum_{j=0}^{D-1} \alpha_j (1-\epsilon)^{2D-1-2j} \\
= \;& \alpha_D - \epsilon \sum_{j=0}^{c_1/\epsilon - 1} \alpha_j (1-\epsilon)^{2D-1-2j} - \epsilon \sum_{j=c_1/\epsilon}^{c_2/\epsilon - 1} \alpha_j (1-\epsilon)^{2D-1-2j} - \epsilon \sum_{j=c_2/\epsilon}^{D-1} \alpha_j (1-\epsilon)^{2D-1-2j} \\
\geq \;& \alpha_D - 2c_1 \alpha_{c_1/\epsilon} (1-\epsilon)^{2D-1-2c_1/\epsilon} - (c_2 - c_1)\alpha_{c_1/\epsilon}(1-\epsilon)^{2D-1-2c_2/\epsilon} - \alpha_D \frac{1+\eta-(1+\eta)^{-D+1+c_2/\epsilon}}{c_0} \\
\geq \;& \alpha_D \left(1 - \frac{1.01 - (1+\eta)^{-D+1+c_2/\epsilon}}{c_0}\right) - \alpha_{c_1/\epsilon}\left(\frac{2c_1}{e^{2k-2c_1}} + \frac{c_2 - c_1}{e^{2k-2c_2}}\right) \\
\approx \;& \alpha_{c_1/\epsilon}\left(\sqrt{\frac{c_1}{k}}\left(1 - \frac{1.01 - (1+c_0\epsilon)^{-k/\epsilon+1+c_2/\epsilon}}{c_0}\right) - \left(\frac{2c_1}{e^{2k-2c_1}} + \frac{c_2-c_1}{e^{2k-2c_2}}\right)\right). \\
\approx \;& \alpha_{c_1/\epsilon}\left(\sqrt{\frac{c_1}{k}}\left(1 - \frac{1.01 - e^{-c_0 k + c_0 c_2}}{c_0}\right) - \left(\frac{2c_1}{e^{2k-2c_1}} + \frac{c_2-c_1}{e^{2k-2c_2}}\right)\right).
\end{aligned} \tag{28}$$

Now we set $c_0 := 6/5$, $c_1 := 2/5$ and $c_2 := \frac{1}{4-2c_0} = 0.625$ (observe that $c_2 < 0.7 \leq k$ as required by the split of the sum). Then, expression (28) becomes a function only of k. One can easily see then that this expression remains positive for all $k \geq 7/10$.

11.2 The positivity of small coefficients for general τ

We observe that $f_{\tau,\epsilon}(x)$ can be rewritten as

$$f_{\tau,\epsilon}(x) = \left((1-x^2)^{-1/\tau} \frac{1+x}{1+(1-\epsilon)x}\right)^{\tau/2}.$$

The advantage is that for $f_{\tau,\epsilon}(x)$ to have positive Taylor expansion, it suffices that the seemingly easier function

$$(1-x^2)^{-1/\tau} \frac{1+x}{1+(1-\epsilon)x} \tag{29}$$

has positive Taylor expansion. We proceed by determining the Taylor expansion of (29). First we note that

$$(1-x^2)^{-1/\tau} = \sum_{j=0}^{\infty} \binom{j-(\tau-1)/\tau}{j} x^{2j} = \sum_{j=0}^{\infty} \alpha_j^{(\tau)} x^{2j}, \tag{30}$$

using again the sequence $\alpha_j^{(\tau)}$. The same argument as in the previous section shows that it suffices to check the positivity of the coefficients of the monomials of even degree x^{2D}, namely

$$\alpha_D^{(\tau)} - \epsilon \sum_{j=0}^{D-1} \alpha_j^{(\tau)} (1-\epsilon)^{2D-1-2j}. \tag{31}$$

The reader can also verify that the coefficients of (31) are exactly the coefficients of the function

$$\frac{1}{(1-x)^{1/\tau}} \frac{1-(1-\epsilon)x}{1-(1-\epsilon)^2 x}.$$

11.2.1 The subcase $D \leq \frac{1}{\tau\epsilon}$

We observe that

$$\sum_{j=0}^{D-1} \alpha_j^{(\tau)} (1-\epsilon)^{2D-1-2j} \overset{Lemma\ (11.1)}{\leq} (1-\epsilon)\tau D \alpha_D$$

Then, expression (31) is at least

$$\alpha_D^{(\tau)} \left(1 - \tau\epsilon D(1-\epsilon)\right),$$

which is positive as long as $D \leq 1/\tau\epsilon$.

11.3 The case $t = 3$ ($\tau = 4$)

We show that the function $f_{4,\epsilon}(x)$, namely

$$\frac{1}{\sqrt{1-x^2}} \left(1 + \epsilon \frac{x}{1+(1-\epsilon)x}\right)^2,$$

has Taylor expansion with positive coefficients. Recall that from section 11.2.1, we know that all coefficients of degree at most $1/4\epsilon$ are positive. Therefore, if we wish, we may restrict our attention to coefficients of degree at least $1/4\epsilon$.

First we can easily prove that all odd coefficients of $f_{4,\epsilon}(x)$ are non negative.

Lemma 11.2. *All odd coefficients of $f_{4,\epsilon}(x)$ are non negative.*

Proof. The reader can first verify that

$$\frac{1}{2}\left(f_{4,\epsilon}(x) - f_{4,\epsilon}(-x)\right) = \frac{1}{2}\left(f_{2,\epsilon}(x) + f_{2,\epsilon}(-x)\right) \frac{2x\epsilon}{1-(1-\epsilon)^2 x}. \tag{32}$$

Observe that the left-hand side of (32), namely $\frac{1}{2}\left(f_{4,\epsilon}(x) - f_{4,\epsilon}(-x)\right)$, gives exactly the odd coefficients of $f_{4,\epsilon}(x)$. Now in the right hand side of (32), $\frac{1}{2}\left(f_{2,\epsilon}(x) + f_{2,\epsilon}(-x)\right)$ are the even coefficients of $f_{2,\epsilon}(x)$ which we have proven in section 11.1 that have positive Taylor expansion. Finally, the function $\frac{2x\epsilon}{1-(1-\epsilon)^2 x}$ has positive Taylor expansion as well, completing the proof. □

Now we focus on the even coefficients of $f_{4,\epsilon}(x)$. Using the same trick as before, we can isolate the even coefficients using the expression

$$\frac{1}{2}\left(f_{4,\epsilon}(x) + f_{4,\epsilon}(-x)\right) = \frac{1}{\sqrt{1-x^2}} \frac{(1-(1-\epsilon)x^2)^2 + \epsilon^2 x^2}{(1-(1-\epsilon)^2 x^2)^2}$$

Therefore we need to show that the function

$$h(x) := \frac{1}{\sqrt{1-x}} \left(\left(\frac{1-(1-\epsilon)x}{1-(1-\epsilon)^2 x}\right)^2 + \left(\frac{\epsilon}{1-(1-\epsilon)^2 x}\right)^2 x \right) \tag{33}$$

has positive Taylor expansion. The second term in (33) can only contribute positively in the Taylor coefficients. Clearly it suffices then to show that the function

$$\frac{1}{(1-x)^{1/4}} \frac{1-(1-\epsilon)x}{1-(1-\epsilon)^2 x} \tag{34}$$

has positive Taylor coefficients, since positivity is preserved after squaring (34). In what follows we prove the next lemma.

Lemma 11.3. *The function of* (34) *has Taylor expansion with positive coefficients.*

We already know from Section 11.2.1 that all coefficients of degree at most $0.25/\epsilon$ are non negative. Next we cover the rest of the degree-spectrum of coefficients, from $0.25/\epsilon$ to infinity.

Claim 11.4. *All coefficients of the function* (34)*, of degree* $0.5/\epsilon \le D \le 1/\epsilon$ *are non negative.*

Proof. We need to show that for D in the above interval,

$$\beta_D := \alpha_D^{(4)} - \epsilon \sum_{j=0}^{D-1} \alpha_j^{(4)} (1-\epsilon)^{2D-1-2j} \ge 0.$$

We will first show that in the above sum the terms first decrease and then start increasing again at around $j \sim 3/8\epsilon$. We will then use this to break the sum into three parts. The first part which we bound in the same fashion as in 11.1.1, the second part we bound using the fact that its terms are always decreasing and the third we bound using a geometric sum.

To show that the first few elements of the sum are decreasing notice that it follows from Lemma 11.1 that,

$$\frac{\alpha_j^{(4)}(1-\epsilon)^{2D-1-2j}}{\alpha_{j+1}^{(4)}(1-\epsilon)^{2D-1-2(j+1)}} = \frac{(j+1)(1-\epsilon)^2}{j+1/4} = (1+3/(4j+1))(1-\epsilon)^2,$$

which is clearly more than 1 for $j=1$ and small ϵ and decreasing with j. Letting the ratio be equal to 1 and solving for j gives us the point at which the elements of the sum start increasing for $j_0 = 3/(8\epsilon) - O(1)$, where the $O(1)$ notation is with respect to ϵ going to 0. In what follows we assume ϵ to be small enough that we can ignore these low order terms. Let $D = c/\epsilon$, $0 \le \lambda, \tau \le 1$ be two parameters to be set later such that τ/ϵ is less than the j_0 calculated above and $\lambda c/\epsilon \le \tau/\epsilon$. We have,

$$\begin{aligned}
\beta_D &= \alpha_D^{(4)} - \epsilon \sum_{j=0}^{D-1} \alpha_j^{(4)} (1-\epsilon)^{2D-1-2j} \\
&= \alpha_D^{(4)} - \epsilon \left(\sum_{j=0}^{\lambda c/\epsilon - 1} \alpha_j^{(4)} (1-\epsilon)^{2D-1-2j} + \sum_{j=\lambda c/\epsilon}^{\tau/\epsilon - 1} \alpha_j^{(4)} (1-\epsilon)^{2D-1-2j} + \sum_{j=\tau/\epsilon}^{c/\epsilon - 1} \alpha_j^{(4)} (1-\epsilon)^{2D-1-2j} \right)
\end{aligned}$$

using the maximum value of the $(1-\epsilon)$ term in the first sum, the fact that the elements of the second sum are decreasing, and using the maximum value of the $\alpha_j^{(4)}$ term in the last sum,

$$\begin{aligned}
\ge \alpha_D^{(4)} - \epsilon \Bigg((1-\epsilon)^{2D-1-2(\lambda c/\epsilon - 1)} \sum_{j=0}^{\lambda c/\epsilon - 1} \alpha_j^{(4)} + \sum_{j=\lambda c/\epsilon}^{\tau/\epsilon - 1} \alpha_{\lambda c/\epsilon}^{(4)} (1-\epsilon)^{2D-1-2(\lambda c/\epsilon)} \\
+ \sum_{j=\tau/\epsilon}^{c/\epsilon - 1} \alpha_{\tau/\epsilon}^{(4)} (1-\epsilon)^{2D-1-2j} \Bigg)
\end{aligned}$$

using first part of Lemma 11.1,

$$\geq \alpha_D^{(4)} - \epsilon(1-\epsilon)^{-1}\Bigg((1-\epsilon)^{2c(1-\lambda)/\epsilon} 4\alpha_{\lambda c/\epsilon}^{(4)} \lambda c/\epsilon + \alpha_{\lambda c/\epsilon}^{(4)} (1-\epsilon)^{2c(1-\lambda)/\epsilon}(\tau - \lambda c)/\epsilon + \alpha_{\tau/\epsilon}^{(4)} \sum_{j=\tau/\epsilon}^{c/\epsilon - 1} (1-\epsilon)^{2D-2j}\Bigg)$$

using $1 - \epsilon \leq e^{-\epsilon}$,

$$\geq \alpha_D^{(4)} - \epsilon(1-\epsilon)^{-1}\left(\exp(-2c(1-\lambda))\alpha_{\lambda c/\epsilon}^{(4)}(4\lambda c + \tau - \lambda c)/\epsilon + \alpha_{\tau/\epsilon}^{(4)} \frac{1-(1-\epsilon)^{2c/\epsilon - 2\tau/\epsilon}}{1-(1-\epsilon)^2}(1-\epsilon)^2\right)$$
$$\geq \alpha_D^{(4)} - \epsilon(1-\epsilon)^{-1}\left(\exp(-2c(1-\lambda))\alpha_{\lambda c/\epsilon}^{(4)}(3\lambda c + \tau)/\epsilon + \alpha_{\tau/\epsilon}^{(4)} \frac{1-(e-O(\epsilon))^{2c-2\tau}}{2\epsilon}(1+O(\epsilon))\right)$$

dividing by $\alpha_{\lambda c/\epsilon}^{(4)}$ and using the last part of Lemma 11.1 and observing that c, λc, and τ are constants bounded away from 0 while $1/\epsilon$ goes to infinity,

$$\geq \lambda^{3/4} - (1+o(1))\epsilon\left(\exp(-2c(1-\lambda))(3\lambda c+\tau)/\epsilon + (\lambda c/\tau)^{3/4} - o(1)\frac{1-\exp(2c-2\tau)}{2\epsilon}\right)$$
$$= \lambda^{3/4} - \exp(-2c(1-\lambda))(3\lambda c + \tau) - (\lambda c/\tau)^{3/4}(1-\exp(2c-2\tau))/2 - o(1),$$

where the $o(1)$ term depends on ϵ and not any other parameters (in paticular D).

The rest of the proof is the game of adjusting our parameters so that the high order part of the above expression is strictly positive. Set $\lambda = \tau = 1/3$ and observe that $\tau/\epsilon < 3/(8\epsilon) - O(1)$ and $\tau/\epsilon \geq \lambda c/\epsilon$ for the range of c we are interested in as required. Now,

$$\beta_D \geq 1/3^{3/4} - \exp(-4c/3)(c+1/3) - c^{3/4}(1-\exp(2c-2/3))/2 - o(1),$$

which can be checked to be bigger than 0.05 for $c \in [0.5, 1]$. □

Claim 11.5. *All coefficients of the function (34), of degree $0.25/\epsilon \leq D \leq 0.63/\epsilon$ and $D \geq 0.89/\epsilon$ are non negative.*

Proof. We will heavily rely on the fact that $D = C/\epsilon \geq 0.25/\epsilon$ and that ϵ can be chosen arbitrarily small. Recall that what we need to show is that

$$\alpha_D^{(4)} - \epsilon \sum_{j=0}^{D-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} \geq 0, \tag{35}$$

where the values $\alpha_D^{(4)}$ obey the rules of Lemma 11.1. We use two kinds of upper bounds. The fist, gives the obvious upper bound on the low terms of the above sum, and it reads as

$$\sum_{j=0}^{c/\epsilon-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} \leq \frac{4c}{\epsilon}\alpha_{c/\epsilon}^{(4)}(1-\epsilon)^{2(C-c)/\epsilon} \approx \frac{4c}{\epsilon}\alpha_{c/\epsilon}^{(4)}\exp(-2(C-c)),$$

for every constant $c > 0$ such that $c < C$. The second bound that will use says that for two constants $c, c' > 0$, we have

$$\begin{aligned}\sum_{j=c/\epsilon-1}^{c'/\epsilon-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} &\leq \alpha_{c/\epsilon}^{(4)} \sum_{j=c/\epsilon}^{c'/\epsilon-1} (1-\epsilon)^{2D-1-2j} \\ &= \alpha_{c/\epsilon}^{(4)} \frac{(1-\epsilon)^2}{2-\epsilon} (1-\epsilon)^{2C/\epsilon}((1-\epsilon)^{-2c'/\epsilon} - (1-\epsilon)^{-2c/\epsilon}) \\ &\approx \alpha_{c/\epsilon}^{(4)} \frac{(1-\epsilon)^2}{2-\epsilon} \big(\exp(-2C+2c') - \exp(-2C+2c)\big) \\ &\approx \frac{\alpha_{c/\epsilon}^{(4)}}{2} \big(\exp(-2C+2c') - \exp(-2C+2c)\big)\end{aligned}$$

Now consider a sequence $c_1, c_2, \ldots, c_k$ of k constants such that $c_1 = c$ and $c_i = i\,c$. Then

$$\begin{aligned}&\sum_{j=0}^{D-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} \\ =\ & \sum_{j=0}^{c_1/\epsilon-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} + \sum_{i=2}^{k} \left(\sum_{j=c_{i-1}/\epsilon}^{j=c_i/\epsilon-1} \alpha_j^{(4)}(1-\epsilon)^{2D-1-2j} \right) \\ \leq\ & \frac{4c_1}{\epsilon} \alpha_{c_1/\epsilon}^{(4)} \exp(-2(c_k-c_1)) + \sum_{i=2}^{k} \left(\frac{\alpha_{c_{i-1}/\epsilon}^{(4)}}{2} (\exp(-2c_k+2c_i) - \exp(-2c_k+2c_{i-1})) \right) \\ \approx\ & \frac{\alpha_{c_1/\epsilon}^{(4)}}{\epsilon} \left(4c \left(\frac{c_k}{c_1}\right)^{3/4} \exp(-2(c_k-c_1)) + \frac{1}{2} \sum_{i=2}^{k} \left(\left(\frac{c_k}{c_{i-1}}\right)^{3/4} (\exp(-2c_k+2c_i) - \exp(-2c_k+2c_{i-1})) \right) \right) \\ =\ & \frac{\alpha_{c/\epsilon}^{(4)}}{\epsilon} \left(4ck^{3/4} \exp(-2c(k-1)) + \frac{1}{2} \sum_{i=2}^{k} \left(\left(\frac{k}{i-1}\right)^{3/4} (\exp(-2c(k-i)) - \exp(-2c(k-i+1))) \right) \right)\end{aligned}$$

It follows that for (35) to hold true (after factoring $\alpha_D^{(4)} = \alpha_{kc/\epsilon}^{(4)}$ out), we need

$$1 - 4ck^{3/4} \exp(-2c(k-1)) - \frac{1}{2} \sum_{i=2}^{k} \left(\left(\frac{k}{i-1}\right)^{3/4} (\exp(-2c(k-i)) - \exp(-2c(k-i+1))) \right) \geq 0. \tag{36}$$

It should be intuitive that the the bigger the value k is, the wider the spectrum of values of c will satisfy (36). This can be verified by any mathematical software program, e.g. MAPLE. Indeed, we can take $k = 1000$, and plot the left-hand side function on c of (36) to verify that it remains positive, as long as $c < 0.00063$ and $c > 0.00089$. Since $k = 1000$, this translates into that $D \leq 0.63/\epsilon$ and $D \geq 0.89/\epsilon$, as promised.

□

ECCC ISSN 1433-8092
http://eccc.hpi-web.de