Vertex Cover resists SDPs tightened by local hypermetric inequalities

Konstantinos Georgiou* Avner Magen* Iannis Tourlakis*

Department of Computer Science
University of Toronto

January 18, 2008

Abstract

We consider the standard semidefinite programming (SDP) relaxation for the VERTEX COVER problem to which all hypermetric inequalities supported on at most k vertices are added and show that the integrality gap for such SDPs remains $2 - o(1)$ even for $k = O(\sqrt{\log n / \log \log n})$. This extends results by Kleinberg-Goemans, Charikar and Hatami et al. who considered VERTEX COVER SDPs tightened using the triangle and pentagonal inequalities, respectively.

Our result is complementary to a recent result by Georgiou et al. proving integrality gaps for VERTEX COVER SDPs in the Lovász-Schrijver hierarchy. However, the SDPs we consider are incomparable to the SDPs analyzed by Georgiou et al. In particular we show that VERTEX COVER SDPs in the Lovász-Schrijver hierarchy fail to satisfy any hypermetric constraints supported on an independent set of the input graph. This constrasts with the LP Lovász-Schrijver hierarchy where all local LP constraints are derived.

*Funded in part by NSERC

1 Introduction

A *vertex cover* for a graph is any subset of vertices that touches all edges in the graph. The approximability of the minimum VERTEX COVER problem on graphs remains one of the outstanding problems in theoretical computer science. While there exists a trivial 2-approximation algorithm, considerable efforts have failed to obtain an approximation ratio better than $2 - o(1)$. Moreover, the strongest PCP-based hardness result known [?] only shows that 1.36-approximation of VERTEX COVER is NP-hard. Assuming Khot's Unique Game Conjecture [?], Khot and Regev [?] show that $2 - o(1)$-approximation is NP-hard.

Several recent papers [?, ?, ?, ?, ?, ?] examine whether semidefinite programming (SDP) relaxations of VERTEX COVER might yield better approximations. Goemans and Williamson [?] introduced semidefinite programming relaxations as an algorithmic technique to obtain a 0.878 approximation for MAX-CUT. Since then semidefinite programming has arguably become our most powerful tool for designing approximation algorithms. Indeed, for many NP-hard optimization problems, the best approximation ratios are achieved using SDP-based algorithms.

For a graph $G = (V, E)$, the standard SDP relaxation for VERTEX COVER is

$$\begin{array}{lll} \min & \sum_{i \in V} (1 + \mathbf{v}_0 \cdot \mathbf{v}_i)/2 & \\ \text{s.t.} & (\mathbf{v}_0 - \mathbf{v}_i) \cdot (\mathbf{v}_0 - \mathbf{v}_j) = 0 & \forall ij \in E \\ & \|\mathbf{v}_i\| = 1 & \forall i \in \{0\} \cup V \end{array} \tag{1}$$

Halperin [?] employed this relaxation (and an appropriate rounding technique) to obtain a $(2 - \Omega(\log\log\Delta/\log\Delta))$-approximation for VERTEX COVER on graphs with maximal degree Δ. On the other hand, Kleinberg and Goemans [?] showed that in general this relaxation has an integrality gap of $2 - o(1)$.

One possible avenue for decreasing this integrality gap comes from the following simple observation: for any integral (or rather, one-dimensional) solution, $\|\mathbf{v}_i - \mathbf{v}_j\|^2$ is an ℓ_1 metric. Therefore the addition of inequalities on the distances $\|\mathbf{v}_i - \mathbf{v}_j\|^2$ that are valid for ℓ_1 metrics may yield a possible tightening of the SDP (note that the constraint $(\mathbf{v}_0 - \mathbf{v}_i) \cdot (\mathbf{v}_0 - \mathbf{v}_j) = 0$ in SDP (1) is in fact the following distance constraint "in disguise": $\|\mathbf{v}_i - \mathbf{v}_0\|^2 + \|\mathbf{v}_j - \mathbf{v}_0\|^2 = \|\mathbf{v}_i - \mathbf{v}_j\|^2$).

For example, since ℓ_1 metrics satisfy the triangle inequality, we could add the following constraint to SDP (1):

$$\|\mathbf{v}_i - \mathbf{v}_j\|^2 + \|\mathbf{v}_j - \mathbf{v}_k\|^2 \geq \|\mathbf{v}_i - \mathbf{v}_k\|^2 \qquad \forall i, j, k \in \{0\} \cup V. \tag{2}$$

This ℓ_2^2 triangle inequality is the crucial addition yielding the breakthrough Arora-Rao-Vazirani SPARSEST CUT algorithm [?]. This suggests that the addition of such inequalities to the standard VERTEX COVER SDP might give a $2 - \Omega(1)$ approximation.

Indeed, Hatami et al. [?] prove that if SDP (1) is strengthened by requiring that the distances $\|\mathbf{v}_i - \mathbf{v}_j\|^2$ satisfy *all* ℓ_1 inequalities (i.e., the vectors $\mathbf{v}_i$ equipped with the ℓ_2^2 norm $\|\cdot\|^2$ are ℓ_1-embeddable), then the resulting relaxation has no integrality gap. Of course, the caveat here is that the resulting relaxation has exponentially many constraints and is hence intractable. To obtain a tractable relaxation (or at least one computable in subexponential time), our relaxation must use only a limited subset of ℓ_1 inequalities.

One canonical subclass of ℓ_1 inequalities is the discrete and easily-described class of *hypermetric* inequalities (see the Preliminaries for definitions). These include the triangle inequalities, as well as the so-called pentagonal, heptagonal, etc., inequalities. For example, the usefulness of these inequalities is witnessed in a work by Avis and Umemoto [**?**]. They show that for dense graphs linear programming relaxations of MAX CUT based on the k-gonal inequalities have integrality gap at most $1 + 1/k$. This, in a sense, gives rise to an LP-based PTAS.

Charikar [**?**] showed that even with the addition of the triangle inequality (2) the integrality gap of SDP (1) remains $2 - o(1)$. Nevertheless, Karakostas [**?**] showed that adding the triangle inequality (as well as the "antipodal" triangle inequalities $(\pm\mathbf{v}_i - \pm\mathbf{v}_j) \cdot (\pm\mathbf{v}_i - \pm\mathbf{v}_k) \geq 0$) yields a $(2 - \Omega(1/\sqrt{\log n}))$-approximation for VERTEX COVER, currently the best ratio achievable by any algorithm. Hatami et al. [**?**] subsequently showed that Karakostas's SDP even with the addition of the pentagonal inequalities has integrality gap $2 - o(\sqrt{\log\log n/\log n})$.

In this work we rule out the possibility that adding local hypermetric constraints improves the integrality gap of VERTEX COVER SDPs:

Theorem 1 *The tightening of the standard SDP for* VERTEX COVER *with all hypermetrics that are supported on* $O(\sqrt{\log n/\log\log n})$ *points has integrality gap* $2 - o(1)$.

As mentioned above, Hatami et al. [**?**] show that adding the constraint that solutions to SDP (1) be ℓ_1-embeddable results in an SDP with no integrality gap. Theorem 1 then immediately gives the following corollary about ℓ_2^2 metrics:

Corollary 1 *There exist* ℓ_2^2 *metrics that are not* ℓ_1*-embeddable yet satisfy all hypermetric inequalities supported on* $O(\sqrt{\log n/\log\log n})$ *points.*

It is interesting to compare Corollary 1 with recent results about local-global phenomena in metric spaces. In [**?**, **?**] the authors describe metric spaces that cannot be well-embedded into ℓ_1 but locally every small subset embeds into ℓ_1 isometrically. In contrast, our corollary shows the existence of a metric that locally resembles ℓ_1 (although not provably ℓ_1) but globally does not embed isometrically into ℓ_1. From that standpoint, this is far weaker than [**?**, **?**]. However, the metric we supply is also an ℓ_2^2 metric. Finding ℓ_2^2 metrics that are far from being ℓ_1 proved to be a very challenging task (see Khot and Vishnoi's work [**?**] motivated by integrality gap instances for SPARSEST CUT). To the best of our knowledge, there are no known results that point to such metrics which further satisfy any local conditions beyond the obvious triangle inequality.

A result related to Theorem 1 was proved by Georgiou et al. in [**?**]. The main result of that paper showed that SDP relaxations obtained by tightening the standard linear programming relaxation for VERTEX COVER using $O(\sqrt{\log n/\log\log n})$ rounds of the LS_+ "lift-and-project" method of Lovász and Schrijver [**?**] have integrality gap $2 - o(1)$. The SDPs considered in [**?**] seem intimately related to those obtained by adding local ℓ_1 or hypermetric constraints. However, the VERTEX COVER SDP relaxation obtained after k rounds of the LS_+ method is incomparable to the relaxation obtained by adding all order k hypermetric inequalities to SDP (1) In section 4 we show in a strong sense the incomparability of these relaxations: Fix any subset S of vertices that is an independent. We then find a hypermetric inequality supported on all points of S that is nevertheless not valid for any VERTEX COVER SDPs in the Lovász Schrijver hierarchy. In particular, before the current work it was conceivable that adding such concrete constraints as, say, all hypermetric inequalities on 7-points (e.g., the "heptagonal" inequalities) may result in a non-trivial SDP relaxation. This was true even in light of the integrality gaps proved in [**?**].

We briefly describe now how we prove Theorem 1. We use the same graph family as in [**?**, **?**, **?**, **?**]. The SDP solution can be though of as an ℓ_1 metric to which a small perturbation was applied.

This perturbation is characterized by two "infinitesimal" parameters, γ and ϵ relating to the graph and the integrality gap, respectively. We show that hypermetric inequalities that are supported on $k \geq 4$ points, one of which is $\mathbf{v}_0$, must have a slack component that depends on k and on ϵ and γ, that will be maintained as long as $k\gamma = O(\epsilon)$. The case of the triangle inequality is covered by [**?**] and [**?**], and the case where $\mathbf{v}_0$ does not participate in the inequality is handled by the fact that the metric formed by the remaining vectors is an ℓ_1 metric. Setting ϵ to an arbitrary small constant, and setting γ to $\Theta(\sqrt{\log\log n/\log n})$ provides the bound in our theorem.

2 Preliminaries

Given two vectors $\mathbf{x}, \mathbf{y} \in \{-1,1\}^n$ their *Hamming distance* $d_H(\mathbf{x},\mathbf{y})$ is $|\{i \in [n] : x_i \neq y_i\}|$. For two vectors $\mathbf{u} \in \mathbb{R}^n$ and $\mathbf{v} \in \mathbb{R}^m$ denote by $(\mathbf{u},\mathbf{v}) \in \mathbb{R}^{n+m}$ the vector whose projection on the first n coordinates is $\mathbf{u}$ and on the last m coordinates is $\mathbf{v}$.

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 that assumes the value $\mathbf{u}_i\mathbf{v}_j$ 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 $P(x) = c_1x^{t_1} + \ldots + c_qx^{t_q}$ be a polynomial with nonnegative coefficients. Then T_P is the function that maps a vector $\mathbf{u}$ to the vector $T_P(\mathbf{u}) = (\sqrt{c_1}u^{\otimes t_1}, \ldots, \sqrt{c_q}u^{\otimes t_q})$.

Fact: For all vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^d$, $T_P(\mathbf{u}) \cdot T_P(\mathbf{v}) = P(\mathbf{u} \cdot \mathbf{v})$.

Metrics and ℓ_1 inequalities We quickly review the facts we need about ℓ_1 inequalities. The book [**?**] of Deza and Laurent is a good source for more information.

A finite metric space is called an ℓ_1 *metric* if it can be embedded in ℓ_1-normed space so that all distances remain unchanged. It is easy to see that the set $\mathcal{C}$ of all ℓ_1 metrics on a fixed number of points is a convex cone. Let X be a set of size n. A subset S of X is associated with a metric $\delta_S(x,y)$ that is called a *cut metric* and is defined as $|\chi_S(x) - \chi_S(y)|$, where $\chi_S(\cdot)$ is the characteristic function of S. These metrics are the extreme rays of $\mathcal{C}$; namely, every ℓ_1 metric is a positive linear combination of cut metrics. This fact leads to a simple characterization of all inequalities that are valid for ℓ_1 metrics as follows. Consider the polar cone of $\mathcal{C}$,

$$\mathcal{C}^* = \{B \in \mathbb{R}^{n\times n} | B \cdot D \leq 0 \text{ for all } D \in \mathcal{C}\},$$

where by $B \cdot D$ we denote the matrix inner product of B and D, that is $B \cdot D = \text{trace}(BD^t) = \sum_{i,j} B_{ij}D_{ij}$. Notice that for B to be in $\mathcal{C}^*$ it is enough to require that $B \cdot \delta_S \leq 0$ for all cuts S. By definition it is clear that any $B \in \mathcal{C}^*$ defines a valid inequality such that $\sum_{i,j} B_{ij}d_{ij} \leq 0$ whenever d is an ℓ_1 metric. Conversely, (strong) duality implies that if d satisfies all inequalities of this type for every $B \in \mathcal{C}^*$ then d is an ℓ_1 metric.

A special canonical class of ℓ_1 inequalities is the class of *hypermetric inequalities*. Let $\mathbf{b} \in \mathbb{Z}^k$, with $\sum_{i=1}^k b_i = 1$. It can be easily verified that $B = \mathbf{b}\mathbf{b}^t$ is in $\mathcal{C}^*$, and the inequality $\sum_{i,j} b_ib_jd_{ij} \leq 0$ is called a *hypermetric*. If we further require $\mathbf{b} \in \{-1,1\}^k$, in which case the hypermetric is called *pure*, we obtain the k-gonal inequalities, e.g., the triangle inequality for $k = 3$, pentagonal inequality for $k = 5$, etc.

3 Construction and proof

Fix arbitrarily small constants $\gamma, \epsilon > 0$ such that $\epsilon > 3\gamma$, and let m be a sufficiently large integer.

The *Frankl-Rödl graph* G_m^γ is the graph with vertices $\{-1,1\}^m$ and where two vertices $i, j \in \{-1,1\}^m$ are adjacent if $d_H(i,j) = (1-\gamma)m$. A classical result of Frankl and Rödl [**?**] implies that the size of a minimal vertex cover in G_m^γ is $2^m(1-o(1))$ whenever $\gamma = \Omega(\sqrt{\log m/m})$. We denote the vertices V of G as vectors $\mathbf{w}_i \in \{-1,1\}^m$ (the association of index i with a vector in the cube is arbitrary) and normalize these to get unit vectors $\mathbf{u}_i = \frac{1}{\sqrt{m}}\mathbf{w}_i$.

Consider the polynomial

$$P(x) = \beta x(x+1)^{\frac{2m}{\gamma}} + \alpha x^{\frac{1}{\gamma}} + (1-\alpha-2\beta)x,$$

where the constants $\alpha, \beta > 0$ will be defined below. Let $\mathbf{z}_0 = (1, 0 \ldots, 0)$, $\mathbf{z}_i = (2\epsilon, \sqrt{1-4\epsilon^2}T_P(\mathbf{u}_i))$, where $T_P(\mathbf{v})$ is the tensoring of $\mathbf{v}$ induced by the polynomial P. We fix the values of α and β defining P (and hence, defining the vectors $\mathbf{z}_i$) according to the following lemma implicit in [?]:

Lemma 1 ([?]) *Suppose $\frac{2m}{\gamma}$ and $\frac{1}{\gamma}$ are even and that m is significantly larger than $1/\gamma$. Suppose further that $\epsilon > 3\gamma$. Then there exist constants $\alpha, \beta > 0$ satisfying*

$$\alpha < 7.5\gamma,$$
$$2\beta + \alpha > \frac{4\epsilon}{1+2\epsilon} - 4\gamma,$$

such that the vectors $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_n$ satisfy both the standard VERTEX COVER *SDP* (1) *and the triangle inequality* (2)*.*

Note that a translated version of the vector set $\{\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_n\}$ lay at root of the LS_+ lower bounds proved in [?]. Specifically, the Gram matrix of the vectors $\mathbf{v}_i = \frac{\mathbf{z}_0+\mathbf{z}_i}{2}$ was shown to be a solution for the VERTEX COVER SDP resulting from $O(\sqrt{\log n/\log\log n})$ rounds of LS_+ lift-and-project.

The remainder of this section is devoted to proving the following theorem.

Theorem 2 *The vectors $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_n$ satisfy all hypermetric inequalities on r points, $r \leq \frac{2}{45}\frac{\epsilon}{\gamma}$.*

We claim that Theorem 1 follows immediately from Theorem 2. Indeed, note first that the value of SDP (1) on the vectors $\mathbf{z}_0, \mathbf{z}_1, \ldots, \mathbf{z}_n$ is $(1+\epsilon)2^{m-1}$. On the other hand, recall that the underlying graph G_m^γ has minimal vertex cover size $(1-o(1))2^m$ whenever $\gamma = \Omega(\sqrt{\log m/m})$. Hence, Theorem 1 follows by taking $\epsilon > 0$ to be any arbitrarily small constant and $\gamma = \Omega(\sqrt{\log m/m})$.

As an aside, we note that our vectors $\{\mathbf{z}_i\}$ also satisfy the "antipodal" triangle inequalities $(\pm\mathbf{z}_i - \pm\mathbf{z}_j)\cdot(\pm\mathbf{z}_i - \pm\mathbf{z}_k) \geq 0$ for all $i, j, k \in \{0\} \cup V$. Recall that these inequalities define the SDP at root of Karakostas's [?] VERTEX COVER algorithm. That our vectors satisfy these inequalities can be seen as follows. Consider the subset $\{\mathbf{z}_i\}_{i\geq 1}$. For each coordinate, the vectors in this subset take on at most 2 different values, and hence this subset is ℓ_1-embeddable. Moreover, this remains true even if we replace some (or all) of the $\mathbf{z}_i$ by $-\mathbf{z}_i$. Hence, it suffices to consider only the "antipodal" triangle inequalities involving $\mathbf{z}_0$. The validity of these inequalities then follows easily from the fact that the $\mathbf{z}_i$ satisfy the standard triangle inequalities (by Lemma 1) and the fact that the value of $\mathbf{z}_i \cdot \mathbf{z}_0$ does not depend on i.

Before giving the proof of Theorem 2 we give some intuition. Note that the vector set $\{\mathbf{z}_i\}$ is the result of a perturbation applied to the following simple-minded ℓ_1 metric: Let $D = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be the metric obtained by taking $\mathbf{v}_i$ to be the (normalized version of) the vectors of the m-dimensional cube, and let $\mathbf{v}_0$ be a unit vector perpendicular to all $\mathbf{v}_i$. Notice that these vectors are precisely the vectors we would have obtained if we had used the polynomial $P(x) = x$ to define the tensored vectors $\mathbf{z}_i$ (corresponding to taking $\epsilon = \gamma = 0$). The metric D is easily seen to be an ℓ_1

metric: take the Hamming cube and place the zeroth point at the origin to get an ℓ_1 embedding that is an isometry. Since D is ℓ_1, every hypermetric inequality is valid for it. On the other hand, D does not satisfy even the basic conditions of SDP (1) (e.g., the edge constraints) with respect to our graph of interest, i.e., G_m^γ with $\gamma > 0$, and the basic attempts to remedy that will already violate the triangle inequality. By focusing on the pure hypermetrics, we can give some intuition of why our construction works, and why the critical value of k is $O(\epsilon/\gamma)$ (for the remaining hypermetrics, this intuition is less accurate). Given any choice of $\alpha, \beta > 0$ we get distances that are a perturbation of D by D_Δ. As was mentioned in the proof outline in the introduction and in light of Lemma 1, we may concentrate on inequalities supported on more than three points. Since any given hypermetric inequality is satisfied by D, it is sufficient to prove that it is satisfied for the perturbated component of the metric, i.e., D_δ. Analyzing the inequality on D_Δ then shows that $\sum_{i,j} b_i b_j d_{ij} \leq -2\epsilon + C\gamma k$, where C is a universal constant, and the d_{ij} are the distances defined by D_Δ. Consequently, as long as $k = O(\epsilon/\gamma)$, the inequality holds for D_Δ. Hence it holds for $D + D_\Delta$, the metric resulting from the $\mathbf{z}_i$ as well.

Proof: [of Theorem 2] By Lemma 1 we already know that the vectors satisfy all hypermetric inequalities on 3 points, namely, the triangle inequalities.

So we only need to show that the solution satisfies hypermetric inequalities on 4 or more points. This is an important point since the arguments we will use to handle hypermetric inequalities on at least 4 points fail for the triangle inequalities.

Consider the set of vectors $\{\mathbf{z}_i\}$, $i \geq 1$. As mentioned above, for each coordinate, the vectors in this subset take on at most 2 different values, and hence this subset is ℓ_1-embeddable. Therefore, any ℓ_1 inequality (and in particular any hypermetric inequality) not involving $\mathbf{z}_0$ must be satisfied.

Now let $B = \mathbf{b}\mathbf{b}^t \in \mathcal{C}^*$, where $\mathbf{b} \in \mathbb{Z}^{k+1}$ and $\sum_{i=0}^k b_i = 1$, be any hypermetric inequality supported on $r = k+1$ points. By the above discussion, it suffices to consider the case where $\mathbf{z}_0$ is one of the points, and we can assume that the points are $0, 1, \ldots, k$. Our goal is to show that $\sum_{i<j\leq k} B_{ij} \|\mathbf{z}_i - \mathbf{z}_j\|^2 \leq 0$. By definition, for $i, j \geq 1$,

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

and $\|\mathbf{z}_i - \mathbf{z}_0\|^2 = 2 - 4\epsilon$. Hence,

$$\sum_{0\leq i<j\leq k} B_{ij}\|\mathbf{z}_i - \mathbf{z}_j\|^2 = 2(1-2\epsilon)\sum_{i=1}^k B_{0i} + 2(1-4\epsilon^2)\sum_{0<i<j\leq k} B_{ij}(1 - P(u_i \cdot u_j)).$$

Therefore, we need to show

$$\sum_{i=1}^k B_{0i} + (1+2\epsilon)\sum_{0<i<j\leq k} B_{ij}(1 - P(u_i \cdot u_j)) \leq 0 \tag{3}$$

We require the following technical lemma. First some definitions. By homogeneity we may assume $b_0 < 0$ (and hence that $b_0 \leq -1$ since $b_0 \in \mathbb{Z}$). Let

$$S = \{i \in [k] : b_i > 0\},$$
$$T = \{i \in [k] : b_i < 0\}.$$

Next denote $H_{ij} = (\mathbf{u}_i \cdot \mathbf{u}_j + 1)(\mathbf{u}_i \cdot \mathbf{u}_j)^{\frac{2m}{\gamma}}$ and $M_{ij} = (\mathbf{u}_i \cdot \mathbf{u}_j)^{\frac{1}{\gamma}}$, and let Δ_{ij} be the Hamming distance between $\mathbf{u}_i$ and $\mathbf{u}_j$. With these definitions we can then write $P(\mathbf{u}_i \cdot \mathbf{u}_j) = \beta H_{ij} + \alpha M_{ij} + (1-\alpha-2\beta)(1 - \frac{2}{m}\Delta_{ij})$.

Lemma 2 *Assume that γ, ϵ and m satisfy the conditions in Lemma 1. Then,*

1. $\sum_{0<i<j\leq k} B_{ij} = \frac{1}{2}((1-b_0)^2 - \sum_{i=1}^{k} b_i^2)$
2. $\sum_{0<i<j\leq k} B_{ij}(-\beta H_{ij} - \alpha M_{ij}) \leq 15\gamma \sum_{i\in S, j\in T} b_i(-b_j)$
3. $\sum_{0<i<j\leq k} B_{ij}\Delta_{ij} \leq \frac{1}{4}m(1-b_0)^2$

Proof: The first equality is an immediate consequence of the fact that $\sum_{i=1}^{k} b_i = 1 - b_0$ and that $(\sum_{i=1}^{k} b_i)^2 = \sum_{i=1}^{k} b_i^2 + 2\sum_{0<i<j\leq k} b_i b_j$.

For the second inequality, note first that $\mathbf{u}_i \cdot \mathbf{u}_j \leq 1 - 1/m$. Hence, H_{ij} is negligible for all $i \neq j$. Moreover, since the $\mathbf{u}_i$ are unit vectors and $1/\gamma$ is even, it follows that $0 \leq M_{ij} \leq 1$. Hence, by the bounds for α and β given by Lemma 1 it follows that $\beta H_{ij} + \alpha M_{ij} \leq 15\gamma$ and the second inequality follows.

For the last inequality notice that since Δ_{ij} is the sum of m cut metrics (defined by the m coordinates), it is enough to show that for every subset $I \subset \{0, 1, \ldots, k\}$,

$$\sum_{0<i<j\leq k} B_{ij}\delta_I(i,j) \leq \frac{1}{4}(1-b_0)^2.$$

Indeed, using the fact that B is a hypermetric we have,

$$\sum_{0<i<j\leq k} B_{ij}\delta_I(i,j) = \sum_{i\in I, j\notin I} b_i b_j = \left(\sum_{i\in I} b_i\right)\cdot\left(1 - b_0 - \sum_{i\in I} b_i\right) \leq \left(\frac{1-b_0}{2}\right)^2.$$

□

We can now bound the left-hand-side of (3). To begin with, we have,

$$\begin{aligned}
&\sum_{i=1}^{k} B_{0i} + (1+2\epsilon)\sum_{0<i<j\leq k} B_{ij}(1 - P(u_i\cdot u_j)) \\
=&\sum_{i=1}^{k} B_{0i} + (1+2\epsilon)\sum_{0<i<j\leq k} B_{ij}(1 - \beta H_{ij} - \alpha M_{ij} - (1-\alpha-2\beta)(1 - \frac{2}{m}\Delta_{ij})) \\
=&\sum_{i=1}^{k} B_{0i} + (1+2\epsilon)\sum_{0<i<j\leq k} B_{ij}(-\beta H_{ij} - \alpha M_{ij} + \alpha + 2\beta + (1-\alpha-2\beta)\frac{2}{m}\Delta_{ij}).
\end{aligned}$$

Applying the inequalities from Lemma 2 it then follows that the above is upper-bounded by

$$
\begin{aligned}
& b_0(1-b_0) + (1+2\epsilon)\left(15\gamma \sum_{i\in S, j\in T} b_i(-b_j) + \frac{1}{2}(\alpha+2\beta)\left[(1-b_0)^2 - \sum_{i=1}^{k} b_i^2\right] + \frac{1}{2}(1-\alpha-2\beta)(1-b_0)^2\right) \\
&= \frac{1}{2}(1-b_0^2) + 2\epsilon\frac{1}{2}(1-b_0)^2 + (1+2\epsilon)\left(15\gamma \sum_{i\in S, j\in T} b_i(-b_j) - \frac{1}{2}(\alpha+2\beta)\sum_{i=1}^{k} b_i^2\right) \\
&< \frac{1}{2}(1-b_0^2) + 2\epsilon\frac{1}{2}(1-b_0)^2 + (1+2\epsilon)\left(15\gamma \sum_{i\in S, j\in T} b_i(-b_j) - \left[\frac{2\epsilon}{1+2\epsilon} - 2\gamma\right]\sum_{i=1}^{k} b_i^2\right) \\
&= \frac{1}{2}(1-b_0^2) + 2\epsilon\frac{1}{2}(1-b_0)^2 - 2\epsilon\sum_{i=1}^{k} b_i^2 + (1+2\epsilon)\left(15\gamma \sum_{i\in S, j\in T} b_i(-b_j) + 2\gamma\sum_{i=1}^{k} b_i^2\right) \\
&< \frac{1}{2}(1-b_0^2) - \epsilon(2\sum_{i=1}^{k} b_i^2 - (1-b_0)^2) + 15\gamma(1+2\epsilon)\left(\sum_{i\in S, j\in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2\right) \\
&< \frac{1}{2}(1-b_0^2) - \epsilon(2\sum_{i=1}^{k} b_i^2 - (1-b_0)^2) + 30\gamma\left(\sum_{i\in S, j\in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2\right).
\end{aligned}
$$

Note that since the hypermetric inequality we are considering is *not* a triangle inequality, it follows that we must have $\sum_{i>0} b_i^2 \geq 3$. But then, the following technical lemma can be used to show that the above is bounded by 0, and hence complete the proof of the theorem.

Lemma 3 *Let $k \leq \frac{2}{45}\frac{\epsilon}{\gamma} - 1$ and let $0 < \epsilon < \frac{1}{6}$ and $\gamma > 0$. Assume $b_0 \leq -1$, $\sum b_i^2 \geq 3$ and that $b_i \neq 0$ for all i. Then*

$$
\frac{1}{2}(1-b_0^2) - 2\epsilon\left(\sum_{i=1}^{k} b_i^2 - \frac{1}{2}(1-b_0)^2\right) + 30\gamma\left(\sum_{i\in S, j\in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2\right) < 0.
$$

Proof: It is not hard to to check that since $\epsilon < \frac{1}{6}$ and b_0 is a (strictly) negative integer, we have

$$
\frac{1}{2}(1-b_0^2) - 2\epsilon\left(\sum_{i=1}^{k} b_i^2 - \frac{1}{2}(1-b_0)^2\right) \leq -2\epsilon\left(\sum_{i=1}^{k} b_i^2 - 2\right) \leq -\frac{2\epsilon}{3}\sum_{i=1}^{k} b_i^2.
$$

It is important to note that it was critical to have $\sum_{i>0} b_i^2 \geq 3$ here, as only then can we claim that $\sum_{i=1}^{k} b_i^2 - 2$ is a positive constant. Indeed, the triangle inequality $(b_0 = -1, b_1 = b_2 = 1)$ (i.e., the only hypermetric inequality for which this doesn't hold), we cannot expect any method bounding the slack of the inequality to do any good: the VERTEX COVER edge constraints force the triangle inequality to be tight for edges!

It now suffices to prove that

$$
-\frac{2\epsilon}{3}\sum_{i=1}^{k} b_i^2 + 30\gamma\left(\sum_{i\in S, j\in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2\right) < 0 \tag{4}
$$

Let s, t be the cardinalities of S, T, respectively, and let $x = \sum_{i \in S} b_i$ and $y = \sum_{i \in T}(-b_i)$. Now, using the Cauchy-Schwartz inequality and the fact that $s, t \leq k$, we get

$$\frac{\sum_{i \in S, j \in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2}{\sum_{i=1}^{k} b_i^2} \leq 1 + \frac{xy}{s(x/s)^2 + t(y/t)^2} \leq 1 + k\frac{xy}{x^2 + y^2} \leq 1 + k/2.$$

(Note that if $y = t = 0$ the bound is trivial and we therefore ignored this case above.) Hence,

$$-\frac{2\epsilon}{3}\sum_{i=1}^{k} b_i^2 + 30\gamma\left(\sum_{i \in S, j \in T} b_i(-b_j) + \sum_{i=1}^{k} b_i^2\right) < \left(-\frac{2\epsilon}{3} + 30\gamma(1 + k/2)\right)\sum_{i=1}^{k} b_i^2,$$

and so (4) holds as long as $k \leq \frac{2}{45}\frac{\epsilon}{\gamma} - 1$. $\square$

Theorem 2 now follows. $\square$

4 Hypermetric inequalities vs. Lovász-Schrijver SDP lift-and-project

In this section we show that hypermetric inequalities need not be derived by Lovász and Schrijver's LS_+ lift-and-project system. Our plan of attack is as follows. After giving all appropriate definitions we will first show that no pure hypermetric inequalities are derived by LS_+ for the convex cone defined by the inequalities $0 \leq x_i \leq x_0$, $i = 1, \ldots, n$. We will then use this result to show the following for VERTEX COVER: Fix a graph G and an independent set S in G, and consider a VERTEX COVER SDP for G derived using LS_+ lift-and-project. Then the constraints defining this SDP do not imply *any* of the pure hypermetric constraint supported on S.

We begin by defining the Lovász-Schrijver LS_+ lift-and-project system [**?**]. In what follows all vectors will be indexed starting at 0. Recall that a set $C \subset \mathbb{R}^n$ is a convex cone if for every $\mathbf{y}, \mathbf{z} \in C$, and for every $\alpha, \beta \geq 0$, $\alpha\mathbf{y} + \beta\mathbf{z} \in C$. Given a convex cone $C \subset \mathbb{R}^{n+1}$ we denote its projection onto the hyperplane $x_0 = 1$ by $C|_{x_0=1}$. Let $\mathbf{e}_i$ denote the vector with 1 in coordinate i and 0 everywhere else. Let $Q_n \subset \mathbb{R}^{n+1}$ be the convex cone defined by the constraints $0 \leq x_i \leq x_0$ and fix a convex cone $C \subset Q_n$. The lifted cone $M_+(C) \subseteq \mathbb{R}^{(n+1)\times(n+1)}$ consists of all positive semidefinite matrices $(n+1) \times (n+1)$ matrices Y such that,

Property I. For all $i = 0, 1, \ldots, n$, $Y_{0i} = Y_{ii}$.

Property II. For all $i = 0, 1, \ldots, n$, $Y\mathbf{e}_i, Y\mathbf{e}_0 - Y\mathbf{e}_i \in C$.

The cone $M_+(C)$ is the LS_+ *positive semidefinite tightening* for C. This procedure can be iterated by projecting $M_+(C)$ back to $\mathbb{R}^{n+1}$ and then re-applying the M_+ operator to the projection. In particular, let $N_+(C) = \{Y\mathbf{e}_0 : Y \in M_+(C)\} \subseteq \mathbb{R}^{n+1}$. Define $N_+^k(C)$ inductively by setting $N_+^0(C) = C$ and $N_+^k(C) = N_+(N_+^{k-1}(C))$, and define $M_+^k(C)$ to be $M_+(N_+^{k-1}(C))$. Lovász and Schrijver show that $N_+^{k+1}(C) \subseteq N_+^k(C)$ and $M_+^{k+1}(C) \subseteq M_+^k(C)$ and that moreover these containment are proper if and only if $N_+^k(C)|_{x_0=1}$ is not the integral hull of $C|_{x_0=1}$. Moreover, they show that $N_+^n(C)|_{x_0=1}$ is equal to the integral hull of $C|_{x_0=1}$. It is useful to note, that $Y \in M_+^k(C) \subseteq \mathbb{R}^{(n+1)\times(n+1)}$ if and only if Y is PSD, satisfies Property I, and the following:

Property II′. For all $i = 0, 1, \ldots, n$, $Y\mathbf{e}_i, Y\mathbf{e}_0 - Y\mathbf{e}_i \in N_+^k(C)$.

With these definitions in hand, we can now begin by showing that $M_+(Q_n)$ does not satisfy any pure hypermetric constraint (recall that Q_n is the cone satisfying $0 \leq x_i \leq x_0$ for all $i = 1, \ldots, n$). As a warm up we examine the triangle inequality of SDP (1) for a three vertex graph with no edges.

Note that this SDP has no edge constraints. Moreover, any vector solution $\mathbf{v}_i$ can be mapped using the affine transformation $\mathbf{v}_i \to (\mathbf{v}_i + \mathbf{v}_0)/2$ to a set of vectors whose Gram matrix is in $M_+(Q_3)$, and vice versa. Now consider three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ that correspond to the three vertices of the graph. Geometrically it is possible to place these vectors such that the Gram matrix of $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ satisfies Properties I and II above for an LS_+ tightening, yet $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ violate triangle inequality. We can accomplish this by making $\mathbf{v}_1$ and $\mathbf{v}_2$ almost coincide and placing $\mathbf{v}_3$ between them.

In the the LS_+ world, the above intuition leads to the following matrix in $M_+(Q_3)$:

$$Y = \begin{pmatrix} 1 & \epsilon & \epsilon & \epsilon \\ \epsilon & \epsilon & 0 & \beta \cdot \epsilon \\ \epsilon & 0 & \epsilon & \beta \cdot \epsilon \\ \epsilon & \beta \cdot \epsilon & \beta \cdot \epsilon & \epsilon \end{pmatrix}$$

By having $\epsilon \in (0, 1/2)$ and $\beta \in [0, 1]$ we ensure Y satisfies Properties I and II. One can show that by setting ϵ arbitrarily close to 0 and β close to but bigger than $1/2$, we ensure that Y is PSD, while ensuring that its Cholesky decomposition violates the triangle inequality.

This construction can be extended to show that $M_+(Q_n)$ does not satisfy any inequality $\sum b_i b_j d_{ij} \leq 0$ where b is a vector of length $n = 2k + 1$, $\sum b_i = 1$, and for all i, $|b_i| = 1$. Indeed, consider an inequality on $2k + 1$ points defined by the the vector $(0, b_1, b_2, \ldots, b_{2k+1}) \in \mathbb{Z}_+^{2k+2}$ (note that $b_0 = 0$) where $b_i = 1$ for $i = 1, \ldots, k + 1$ and $b_i = -1$ for $i = k + 2, \ldots, 2k + 2$. In this way we naturally split the points into two clusters of size $k + 1$ and k points. The associated inequality requires that the sum of distances across the clusters dominates the sum of distances within the clusters. Define the distance within the clusters as 2ϵ, and the distance across the clusters as $2\epsilon(1 - \beta)$. We have $k(k + 1)$ cross pairs and $\binom{k}{2} + \binom{k+1}{2} = k^2$ inner pairs. Therefore in order to violate the inequality, we should have $2\epsilon(1 - \beta)k(k + 1) < 2\epsilon k^2$. In other words it suffices for β to be even slightly bigger than $\frac{1}{k+1}$ (this will be crucial later).

Define the matrix

$$Y^{(s,t)} = \begin{pmatrix} 1 & \epsilon \cdot J_{1,s} & \epsilon \cdot J_{1,t} \\ \epsilon \cdot J_{s,1} & \epsilon \cdot I_s & \epsilon \cdot \beta \cdot J_{s,t} \\ \epsilon \cdot J_{t,1} & \epsilon \cdot \beta \cdot J_{t,s} & \epsilon \cdot I_t \end{pmatrix},$$

where $J_{m,n}$ is the $m \times n$ all-1 matrix. The configuration we described above can be realized by the matrix $Y^{(k+1,k)}$ of order $(2k + 2)$. Similarly as in the case of the triangle inequality, $Y^{(s,t)}$ satisfies Properties I and II as long as $\epsilon \in (0, 1/2)$ and $b \in [0, 1]$.

Hence, $Y^{(k+1,k)}$ is in $M_+(Q_n)$ provided we can show that it is PSD. This is implied by the following technical lemma.

Lemma 4 *For all s, t, such that $s + t = 2k + 1$ there exist $\epsilon \in (0, 1/2)$ and $\beta > \frac{1}{k+1}$ such that the matrix $Y^{(s,t)} \in \mathbb{R}^{(2k+2)\times(2k+2)}$ is PSD.*

Proof: Let $\mathbf{0}_{n,m}$ be the $n \times m$ zero matrix and define the matrix

$$U^{(s,t)} = \begin{pmatrix} \mathbf{0}_{s,s} & J_{s,t} \\ J_{t,s} & \mathbf{0}_{t,t} \end{pmatrix}.$$

We decompose $Y^{(s,t)}$ as

$$Y^{(s,t)} = \begin{pmatrix} 1 & \epsilon \cdot J_{1,s+t} \\ \epsilon \cdot J_{s+t,1} & \mathbf{0}_{s,t} \end{pmatrix} + \epsilon \cdot \begin{pmatrix} 0 & \mathbf{0}_{1,s+t} \\ \mathbf{0}_{s+t,1} & I_{s+t,s+t} + \beta U^{(s,t)} \end{pmatrix}.$$

We prove that the resulting matrices are PSD. First, it is not difficult to see that the only non-zero eigenvalues of the first matrix are $\frac{1\pm\sqrt{1+4(s+t)\epsilon^2}}{2}$ which are strictly positive for ϵ arbitrarily close to 0.

It is also easy to see that the only non zero eigenvalues of $U^{(s,t)}$ are $\pm\sqrt{s \cdot t}$. Therefore the eigenvalues of the matrix $I_{s+t,s+t} + \beta U^{(s,t)}$ are $1 \pm \beta\sqrt{st}$, and the rest are 1. These eigenvalues remain positive as long as $\beta\sqrt{st} < 1$. Recall here that we required $\beta > \frac{1}{k+1}$ and so we can take β arbitrarily close to that bound. But then

$$\beta\sqrt{st} \leq \frac{1}{k+1}\frac{2k+1}{2} = \frac{2k+1}{2k+2} < 1$$

completing the proof. □

We are ready now to show that VERTEX COVER SDPs in the LS_+ hierarchy violate pure hypermetrics on any independent set. Fix an n-vertex graph $G = (V, E)$ and consider the convex cone $C \subset Q_n$ consisting of all vectors $x \in \mathbb{R}^{n+1}$ such that $x_i + x_j \geq x_0$. Then LS_+ lifting yields the following sequence of SDPs for G: $M_+(C), M_+^2(C), \ldots$. We will show that for all k, every independent set S in G, and all pure hypermetrics B supported on S, there exists $Y \in M_+^k(C)$ such that Y does not satisfy B.

To that end, fix k and S, and let $s = |S|$ be odd. Without loss of generality, assume that $S = \{1, 2, \ldots, s\}$. Fix a pure hypermetric B defined on the set S. By the discussion above we know that there exists $Y' \in M_+(Q_s)$ that violates the pure hypermetric B. Let $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_s$ be the Cholesky decomposition for Y'. Now let $Y \in \mathbb{R}^{(n+1)\times(n+1)}$ be the matrix with Cholesky decomposition $\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_s, \mathbf{v}'_{s+1}, \ldots, \mathbf{v}'_n$ where $\mathbf{v}'_j = \mathbf{v}_0$ for all $j \geq s+1$. By construction Y is PSD, satisfies Property I, and does not satisfy B on S. So it suffices to verify Property II$'$ in order to show that $Y \in M_+^k(C)$. Note that $Y\mathbf{e}_i$ is the all-1 vector for all $i \geq s+1$ and hence Property II$'$ holds for all $i \geq s+1$ since the all-1 vector is in the integral hull and hence in $N_+^k(C)$ for all k. Now consider a vector $Y\mathbf{e}_i$ where $1 \leq i \leq s$. Note that $Y_{00} = Y_{0j}$ for all $j \geq s+1$. But then, since S is independent, it follows that the projection of $Y\mathbf{e}_i$ onto the hyperplane $x_0 = 1$ is also in the integral hull and hence in $N_+^k(C)$. Similarly, it follows that $Y(\mathbf{e}_0 - \mathbf{e}_i)$ is also in $N_+^k(C)$ whenever $1 \leq i \leq s$. So Property II' holds for all i, and $Y \in M_+^k(C)$.

We end this section by remarking that the above arguments can be combined with those from [**?**] to show that there is a graph G for which $O(\sqrt{\log n/\log\log n})$ rounds of LS_+ produce an SDP which (a) does not satisfy the triangle inequality *and* (b) has integrality gap $2-o(1)$. The argument, which we do not have room to go into here, considers the Frankl-Rödl graph G_m^γ to which we append three isolated vertices. The idea is to not satisfy the triangle inequality on the isolated vertices while the remaining vertices will essentially employ the SDP solutions from [**?**].

5 Acknowledgements

We thank Toniann Pitassi for many helpful discussions. The first author would also like to thank Periklis Papakonstantinou for his help on some technical aspects of this work.