Electronic Colloquium on Computational Complexity, Report No. 112 (2011)

The Projection Games Conjecture and The NP-Hardness of $\ln n$-Approximating Set-Cover

Dana Moshkovitz *

August 7, 2011

Abstract

In this paper we put forward a conjecture: an instantiation of the Sliding Scale Conjecture of Bellare, Goldwasser, Lund and Russell to projection games. We refer to this conjecture as the Projection Games Conjecture.

We further suggest the research agenda of establishing new hardness of approximation results based on the conjecture. We pursue this line of research by establishing a tight $\mathcal{NP}$-hardness result for the Set-Cover problem. Specifically, we show that under the projection games conjecture (in fact, under a quantitative version of the conjecture that is only slightly beyond the reach of current techniques), it is $\mathcal{NP}$-hard to approximate Set-Cover on instances of size N to within $(1-\alpha)\ln N$ for arbitrarily small $\alpha > 0$. We do this by modifying Feige's reduction that gives a $(1-\alpha)\ln N$ inapproximability under the assumption $\mathcal{NP} \not\subseteq DTIME(N^{O(\log\log N)})$.

1 Introduction

1.1 Projection Games and The Projection Games Conjecture

Most of the $\mathcal{NP}$-hardness of approximation results known today (e.g., all of the results in Håstad's paper [Hås01]) are based on a PCP Theorem for *projection games* (also known as Label-Cover). The input to a projection game consists of: (i) a graph $G = (A, B, E)$; (ii) finite alphabets Σ_A, Σ_B; (iii) constraints (also called *projections*) $\pi_e : \Sigma_A \to \Sigma_B$ for every edge $e \in E$. The goal is to find assignments to the vertices $\varphi_A : A \to \Sigma_A$, $\varphi_B : B \to \Sigma_B$ that *satisfy* as many of the edges as possible. We say that an edge $e = (a, b) \in E$ is satisfied, if the projection constraint holds, i.e., $\pi_e(\varphi_A(a)) = \varphi_B(b)$. We denote the size of a projection game by $n = |A| + |B| + |E|$. A PCP Theorem for projection games with soundness error ε and alphabet size k (where ε and k may depend on n) states the following:

> *Given a projection game with alphabets of size k, it is $\mathcal{NP}$-hard to distinguish between the case where all edges can be satisfied and the case where at most ε fraction of the edges can be satisfied.*

We can refine this statement by saying that there is a reduction from (exact) Sat to projection games, and the reduction maps instances of Sat of size n to projection games of size $N = n^{1+o(1)} poly(1/\varepsilon)$. Such PCPs are referred to as "almost-linear size PCP" because of the exponent of n, although for small ε the blow-up may be super-linear. The refined version gives sharp time

*dmoshkov@csail.mit.edu. Department of Electrical Engineering and Computer Science, MIT.

ISSN 1433-8092

lower bounds for projection games based on the exponential time hypothesis, namely assuming that exact SAT requires $2^{\Omega(n)}$ time.

The state of the art today for PCP Theorems for projection games is the following:

Theorem 1 ([MR10])**.** *There exists $c > 0$, such that for every $\varepsilon \geq 1/n^c$, an almost-linear size PCP Theorem for projection games with soundness error ε and alphabet size $exp(1/\varepsilon)$ holds.*

The $1/n^c$ is tight, as ε that is lower than polynomially small in n requires super-logarithmic randomness, and hence a super-polynomial construction, not permitting $\mathcal{NP}$-hardness. The $exp(1/\varepsilon)$ is not tight. It can be shown that $|\Sigma| \geq 1/\varepsilon$, and we conjecture that an alphabet size of $poly(1/\varepsilon)$ could be achieved:

Conjecture 1 (Projection games conjecture, PGC)**.** *There exists $c > 0$, such that for every $\varepsilon \geq 1/n^c$, an almost-linear size PCP Theorem with soundness error ε and alphabet size $poly(1/\varepsilon)$ holds.*

In fact, in almost all applications, one wishes the alphabet size to be at most polynomial in n, and so Theorem 1 is useful only when $\varepsilon \geq 1/(\log n)^\beta$ for some constant $\beta > 0$.

The conjecture is consistent with existing approximation algorithms for projection games, giving $1/\varepsilon = O(\sqrt[3]{nk})$ (see [CHK09]; note that their formulation is slightly different than ours – they have a vertex per vertex and a possible assignment to it in our formulation).

If one replaces the $\mathcal{NP}$-hardness in a projection game PCP with intractability under the assumption $\mathcal{NP} \not\subseteq DTIME(n^{\Theta(\log(1/\varepsilon))})$, then Conjecture 1 follows from the Parallel Repetition Theorem of Raz [Raz98].

We remark that the projection games conjecture is in fact the Sliding Scale Conjecture of Bellare, Goldwasser, Lund and Russell [BGLR93] instantiated for projection games. By "sliding scale" we refer to the idea that the error can be decreased as we increase the alphabet size. Bellare et al conjectured that polynomially small error could be achieved simultaneously with polynomial alphabet, even for two queries. They did not formulate their conjecture for projection games – the importance of projection games was not fully recognized when they published their work in 1993.

We believe that the projection games conjecture provides a stable foundation on which many new hardness of approximation results can be based. Several concrete results we think should be achievable (but require further ideas) are: tight low order terms for the $\mathcal{NP}$-hardness of approximation of the problems appearing in Håstad's paper, such as 3LIN and 3SAT [Hås01]; tight lower bound for approximating CLIQUE assuming the exponential time hypothesis [Hås99, Kho01]; and n^α $\mathcal{NP}$-hardness of approximation for the SHORTEST-VECTOR-PROBLEM (SVP) in lattices [Kho05]. In this paper, we show a tight $\mathcal{NP}$-hardness of approximation $(1-\alpha)\ln n$ for SET-COVER.

In all the aforementioned examples, the existing reductions have super-polynomial blow-up, not only in order to achieve low error for a projection game, but also to facilitate the reduction. For instance, Håstad's reductions use the long code on top of a projection game. For low error ε, the long code incurs a large blow-up $2^{(1/\varepsilon)^{O(1)}}$. Basing the results on the PGC, would require reductions that do not resort to large blow-ups.

1.2 The Potential Influence of The PGC

The area of hardness of approximation already benefited greatly from the introduction of another conjecture: the unique games conjecture (UGC) by Khot [Kho02]. The UGC provides a basis on which researchers can prove (conditionally) new hardness results. The projection games conjecture has a similar flavor. One major difference is that while we do not know of any reduction from SAT to unique games (in the parameters setting considered by the UGC) with less than doubly-exponential blow-up, we do know a super-polynomial reduction from SAT to projection games (in the parameters setting considered by the PGC) - by the parallel repetition theorem [Raz98]. This fact means several things for the potential influence of the PGC. On the positive side, we have much better evidence that the PGC is true than we have that the UGC is true. Hence, we have much more assurance that results proved conditioned on the PGC indeed hold. On the negative side, one could claim that the PGC is not likely to revolutionize the area of hardness of approximation the way the UGC did: While the UGC allows hardness results that could not be conceived before, the PGC only gives rise to hardness results that could have been based on the assumption $\mathcal{NP} \not\subseteq DTIME(n^{\Theta(\log n)})$. As a reply, we wish to stress two points:

- *Super-polynomial reductions yield weak lower bounds.* Suppose that there is a reduction from Sat instances of size n to instances of size $N = n^a$ of some other problem Π. If the exponential time hypothesis holds and Sat requires time $2^{\Omega(n)}$, then Π requires time $2^{\Omega(N^{1/a})}$.

  The best reduction for projection games known today gives $a = 1 + o(1)$ [MR10], and hence a nearly-exponential lower bound $2^{\Omega(N^{1-o(1)})}$. A super-polynomial reduction with $a = \Theta(\log n)$ gives a much weaker, strictly sub-exponential, lower bound of $2^{\Omega(2^{\sqrt{\log N}})}$.

- *The PGC is likely to be proved in the foreseeable future.* While it may seem that all we do is replace one plausible assumption ($\mathcal{NP} \not\subseteq DTIME(n^{\Theta(\log n)})$) with another (the PGC), the two assumptions are, in fact, quite different qualitatively. The PGC is likely to be proved in the foreseeable future, while proving that $\mathcal{NP} \not\subseteq DTIME(n^{\Theta(\log n)})$, even assuming $\mathcal{P} \neq \mathcal{NP}$, is an open problem unlikely to be solved any time soon.

One can argue that the PGC was already known to experts in the area to be an interesting open problem. However, we do not know of any "official" posing of a conjecture (the author did define it in a mini-course she taught at Princeton in 2009). Moreover, to the best of our knowledge, the next logical step, that of proving hardness results based on the conjecture, was not taken. The purpose of the current paper is to fix these two deficiencies.

1.3 Set-Cover

In this paper we demonstrate one application of the projection games conjecture to the $\mathcal{NP}$-hardness of approximating Set-Cover. In fact, the quantitative version of the conjecture that we need is much weaker than the full conjecture, and it is the one that is just outside the reach of current techniques, requiring soundness error $c/\log^4 n$ for sufficiently small constant $c > 0$. Feige already showed a tight inapproximability result for Set-Cover, but under the assumption $\mathcal{NP} \not\subseteq DTIME(N^{O(\log\log N)})$ [Fei98].

The definition of Set-Cover is as follows:

Definition 2 (Set-Cover)**.** *The input to* Set-Cover *consists of a universe U, $|U| = n$ and subsets $S_1, \ldots S_m \subseteq U$, $\bigcup_{j=1}^m S_j = U$, $m \leq poly(n)$. The goal is to find as few sets $S_{i_1}, \ldots, S_{i_k}$ as possible that cover U, i.e., $\bigcup_{j=1}^k S_{i_j} = U$.*

The greedy algorithm is a $\ln n$-approximation for SET-COVER. Feige showed that this is where the threshold for Set-Cover lies, i.e., that SET-COVER cannot be approximated better than $\ln n$, assuming $\mathcal{NP} \not\subseteq DTIME(n^{O(\lg\lg n)})$ [Fei98]. The untraditional assumption (stronger than $\mathcal{P} \neq \mathcal{NP}$) comes from the use of the parallel repetition theorem. Parallel repetition is used by Feige not only to ensure very low error $1/(\log n)^{O(1)}$, but also for its unique structure. We show that one can do without the structure that comes with parallel repetition, and in fact, any projection game PCP with poly-logarithmically small error suffices. We prove the following theorem by adapting Feige's ideas to the general projection games framework:

Theorem 3. *For every $0 < \alpha < 1$, there is $c = c(\alpha)$, such that if the projection games conjecture (even without almost-linear size) holds with error $\varepsilon = \frac{c}{\lg^4 n}$, then it is $\mathcal{NP}$-hard to approximate* SET-COVER *better than $(1-\alpha)\ln n$.*

2 Preliminaries

For a set S and a natural number ℓ we denote by $\binom{S}{\ell}$ the family of all sets of ℓ elements from S.

We assume without loss of generality that the projection game in Conjecture 1 is bi-regular, i.e., all the A vertices have the same degree, which we call the A-degree, and all the B vertices have the same degree, which we call the B-degree. We note that any projection game can be converted to bi-regular using a technique developed in [MR10] ("right degree reduction – switching sides – right degree reduction"), and the cost in the soundness error and graph size are swallowed in the parameters stated in Conjecture 1.

3 The New Component

Feige uses the structure obtained from parallel repetition to achieve a projection game in which the soundness guarantee is that very few B vertices have any two of their neighbors agree on a value for them:

Definition 4 (Total disagreement)**.** *Assume a projection game $(G = (A, B, E), \Sigma_A, \Sigma_B, \Phi)$. Let $\varphi_A : A \to \Sigma_A$ be an assignment to the A vertices. We say that the A vertices* totally disagree *on a vertex $b \in B$ if there are no two neighbors $a_1, a_2 \in A$ of b, $e_1 = (a_1, b), e_2 = (a_2, b) \in E$, for which*

$$\pi_{e_1}(\varphi_A(a_1)) = \pi_{e_2}(\varphi_A(a_2)).$$

Definition 5 (Agreement soundness)**.** *Assume a projection game $(G = (A, B, E), \Sigma_A, \Sigma_B, \Phi)$ for deciding membership in a language L. We say that G has* agreement soundness error *ε, if for inputs $x \notin L$, for any assignment $\varphi_A : A \to \Sigma_A$, the A vertices are in total disagreement on at least $1 - \varepsilon$ fraction of the $b \in B$.*

Feige used parallel repetition to achieve agreement soundness. We show a different way to achieve agreement soundness. Our construction centers around the following combinatorial construction:

Lemma 3.1 (Combinatorial construction)**.** *For every n, $0 < \varepsilon < 1$, for infinitely many D, there is an explicit construction of a regular graph $H = (U, V, E)$ with $|V| = n$, V-degree D, and $|U| \leq n^{O(1)}$ that satisfies the following. For every partition $U_1, \ldots, U_l$ of U into sets, such that $|U_i| \leq \varepsilon |U|$ for $i = 1, \ldots, l$, the fraction of vertices $v \in V$ with more than one neighbor in a single set U_i, is at most εD^2.*

Note that the combinatorial property could be achieved by a randomized construction, or by a construction that has a V vertex per every possible set of D neighbors in U. However, the first construction is randomized and the second – too wasteful with a size of $\approx |U|^D$. The lemma can therefore be thought of as a *derandomization* of the randomized/full constructions.

Proof. (of Lemma 3.1) Associate U with a space $\mathbb{F}^m$ where $\mathbb{F}$ is a finite field of size $|\mathbb{F}| = D$, and m is a natural number. Let V be the set of all lines in $\mathbb{F}^m$. Hence, $|V| = \binom{|U|}{2}/\binom{|\mathbb{F}|}{2}$. We connect a line $v \in V$ with a point $u \in U$ if u lies in v.

Let us show this construction satisfies the desired property. Fix a partition $U_1, \ldots, U_l$ of U into tiny sets, $|U_i| \le \varepsilon |U|$ for $i = 1, \ldots, l$. For every $1 \le i \le l$, the number of V lines that have at least two neighbors in U_i is at most $\binom{|U_i|}{2}$. Thus the total number of V vertices with more than one neighbor in a single U_i is at most

$$
\begin{aligned}
\sum_{i=1}^{l} \binom{|U_i|}{2} &\le \sum_{i=1}^{l} \frac{|U_i|^2}{2} \\
&\le \max\{\, |U_i| \mid 1 \le i \le l \} \cdot \sum_{i=1}^{l} \frac{|U_i|}{2} \\
&\le \varepsilon |U| \cdot \frac{|U|}{2} \\
&\le \varepsilon |\mathbb{F}|^2 |V| .
\end{aligned}
$$

□

We show how to take a projection game with standard soundness and convert it to a projection game with total disagreement soundness, by combining it with the graph from Lemma 3.1.

Lemma 3.2. *Let $D \ge 2$, $\varepsilon > 0$. From a projection game with soundness error $\varepsilon^2 D^2$, we can construct a projection game with agreement soundness error $2\varepsilon D^2$ and B-degree D. The transformation preserves the alphabets of the game. The size is raised to a constant factor.*

Proof. Let $\mathcal{G} = (G = (A, B, E), \Sigma_A, \Sigma_B, \Phi)$ be the original projection game. Assume that the B-degree is $|U|$, and we use U to enumerate the neighbors of a B vertex, i.e., there is a function $E^{\leftarrow} : B \times U \to A$ that given a vertex $b \in B$ and $u \in U$, gives us the A vertex which is the u neighbor of b.

Let $H = (U, V, E_H)$ be the graph from Lemma 3.1. We create a new projection game $(G = (A, B \times V, E'), \Sigma_A, \Sigma_B, \Phi')$. The intended assignment to every vertex $a \in A$ is the same as its assignment in the original game. The intended assignment to a vertex $\langle b, v \rangle \in B \times V$ is the same as the assignment to b in the original game. We put an edge $e' = (a, \langle b, v \rangle)$ if $E^{\leftarrow}(b, u) = a$ and $e = (u, v) \in E_H$. We define $\pi_{e'} \equiv \pi_e$.

If there is an assignment to the original game that satisfies c fraction of its edges, then the corresponding assignment to the new game satisfies c fraction of its edges.

Suppose there is an assignment for the new game $\varphi_A : A \to \Sigma_A$ in which more than $2\varepsilon D^2$ fraction of the vertices in $B \times V$ do not have total disagreement.

Let us say that $b \in B$ is "good" if for more than εD^2 of the vertices in $\{b\} \times V$ the A vertices do not totally disagree. Note that the fraction of good $b \in B$ is at least εD^2.

Focus on a good $b \in B$. Consider the partition of U into $|\Sigma_B|$ sets, where the set corresponding to $\sigma \in \Sigma_B$ is:

$$U_\sigma = \{\, u \in U \mid a = E^{\leftarrow}(b, u) \wedge e = (a, b) \wedge \pi_e(\varphi_A(a)) = \sigma \} .$$

By the property of H, there must be $\sigma \in \Sigma_A$ such that $|U_\sigma| > \varepsilon |U|$. We call σ the "champion" for b.

We define an assignment $\varphi_B : B \to \Sigma_B$ that assigns good b's their champions, and other b's arbitrary values. The fraction of edges that φ_A, φ_B satisfy in the original game is at least $\varepsilon^2 D^2$. □

Next we consider a variant of projection games that is relevant for the reduction to Set-Cover. In this variant the prover is allowed to assign each vertex ℓ values, and an agreement is interpreted as agreement on *one* of the assignments in the list:

Definition 6 (List agreement)**.** *Assume a projection game* $(G = (A, B, E), \Sigma_A, \Sigma_B, \Phi)$*. Let* $\ell \geq 1$*. Let* $\hat{\varphi}_A : A \to \binom{\Sigma_A}{\ell}$ *be an assignment that assigns each A vertex l alphabet symbols. We say that the A vertices* totally disagree *on a vertex* $b \in B$ *if there are no two neighbors* $a_1, a_2 \in A$ *of* b*,* $e_1 = (a_1, b), e_2 = (a_2, b) \in E$*, for which there exist* $\sigma_1 \in \hat{\varphi}_A(a_1)$*,* $\sigma_2 \in \hat{\varphi}_A(a_2)$*, such that*

$$\pi_{e_1}(\sigma_1) = \pi_{e_2}(\sigma_2).$$

Definition 7 (List agreement soundness)**.** *Assume a projection game* $(G = (A, B, E), \Sigma_A, \Sigma_B, \Phi)$ *for deciding membership in a language* L*. We say that* G *has* agreement soundness error (ℓ, ε)*, if for inputs* $x \notin L$*, for any assignment* $\hat{\varphi}_A : A \to \binom{\Sigma_A}{\ell}$*, the A vertices are in total disagreement on at least* $1 - \varepsilon$ *fraction of the* $b \in B$*.*

If a projection game has low error ε, then even when the prover is allowed to assign each A vertex ℓ values, the game is still sound. This is argued in the next corollary:

Lemma 3.3 (Projection game with list agreement soundness)**.** *Let* $\ell \geq 1$*,* $0 < \varepsilon' < 1$*. A projection game with agreement soundness error* ε' *has agreement soundness error* $(\ell, \varepsilon' \ell^2)$*.*

Proof. Assume on way of contradiction that the projection game has an assignment $\hat{\varphi}_A : A \to \binom{\Sigma_A}{\ell}$ such that on more than $\varepsilon' \ell^2$ fraction of the B vertices, the A vertices do not totally disagree. Define an assignment $\varphi_A : A \to \Sigma_A$ by assigning every vertex $a \in A$ a symbol picked uniformly at random from the ℓ symbols in $\hat{\varphi}_A(a)$. If a vertex $b \in B$ has two neighbors $a_1, a_2 \in A$ that agree on b under the list assignment $\hat{\varphi}_A$, then the probability that they agree on b under the assignment φ_A is at least $1/\ell^2$. Thus, under φ_A, the expected fraction of the B vertices that have at least two neighbors that agree on them, is more than ε'. In particular, there exists an assignment to the A vertices, such that more than ε' fraction of the B vertices have two neighbors that agree on them. This contradicts the agreement soundness of the game. □

By applying Lemma 3.2 and then Lemma 3.3 on the game from Conjecture 1, we get:

Corollary 3.4. *Assuming Conjecture 1, for any* $\ell \geq 1$*, for infinitely many* D*, for any* $\varepsilon \geq 1/n^c$*, given a projection game with alphabet size* $poly(1/\varepsilon)$ *and B-degree* D*, it is* $\mathcal{NP}$*-hard to distinguish between the case where all edges can be satisfied, and the case where the agreement soundness error is* $(\ell, 2D\ell^2\sqrt{\varepsilon})$*.*

4 Following Feige's Reduction

In the remainder, we will show how to use Corollary 3.4 to obtain the desired hardness result for Set-Cover. The reduction is along the lines of Feige's original reduction.

For the reduction we rely on a combinatorial construction of a universe together with partitions of it. Each partition covers the universe, but any cover than takes at most one set out of each partition, is necessarily large:

Lemma 4.1 (Partition system, [NSS95])**.** *For natural numbers m, D, for $\alpha \leq 2/D$, there is an explicit construction of a universe U, $|U| \leq poly(D^{\log D}, \log m)$ and partitions $\mathcal{P}_1, \ldots, \mathcal{P}_m$ of U into D sets that satisfy the following: there is no cover of U with $\ell = D \ln |U| (1 - \alpha)$ sets $S_{i_1}, \ldots, S_{i_\ell}$, $1 \leq i_1 < \cdots < i_\ell \leq m$, such that set S_{i_j} belongs to partition $\mathcal{P}_{i_j}$.*

To see why $\ell = D \ln |U| (1 - \alpha)$ is to be expected (this later determines the hardness factor we get), think of the following randomized construction: each element in U corresponds to a vector in $[D]^m$, specifying for each of the m partitions, to which of its D sets it belongs. Consider a uniformly random choice of such a vector. Fix any $S_{i_1}, \ldots, S_{i_\ell}$. The probability that a random element is not covered by $S_{i_1}, \ldots, S_{i_\ell}$ is $(1 - 1/D)^\ell \approx e^{-\ell/D}$. When $\ell = D \ln |U| (1 - \alpha)$, we have $e^{-\ell/D} \geq 1/|U|$, and we expect one of the $|U|$ elements in U not to be covered by $S_{i_1}, \ldots, S_{i_\ell}$.

We now describe the reduction from a projection game $\mathcal{G}$ as in Corollary 3.4, to a Set-Cover instance $\mathcal{SC}_{\mathcal{G}}$.

Apply Lemma 4.1 for $m = |\Sigma_B|$ and D which is the B-degree of the projection game. Let U be the universe, and $\mathcal{P}_{\sigma_1}, \ldots, \mathcal{P}_{\sigma_m}$ be the partitions of U. We index the partitions by Σ_B symbols $\sigma_1, \ldots, \sigma_m$. The elements of the Set-Cover instance are $B \times U$.

For every vertex $a \in A$ and an assignment $\sigma \in \Sigma_A$ to a we have a set $S_{a,\sigma}$ in the Set-Cover instance. The intuition is that whether we take $S_{a,\sigma}$ to the cover would correspond to assigning σ to a. The set $S_{a,\sigma}$ is a union of subsets, one for every edge $e = (a, b)$ touching a. Suppose e is the i'th edge coming into b ($1 \leq i \leq D$), then the subset associated with e is the i'th subset of the partition $\mathcal{P}_{\varphi_e(\sigma)}$. Note that if we have an assignment to the A vertices, such that all of b's neighbors agree on one value for b, then the D subsets corresponding to those neighbors and their assignments form a partition that covers b's universe. On the other hand, if one uses only sets that correspond to totally disagreeing assignments to the neighbors, then by the definition of the partitions, covering U requires $\approx \ln |U|$ times more sets.

Claim 4.2. *The following hold:*

- *Completeness: If all the edges in $\mathcal{G}$ can be satisfied, then $\mathcal{SC}_{\mathcal{G}}$ has a set cover of size $|A|$.*
- *Soundness: Let $\ell \doteq D \ln |U| (1 - \alpha)$ be as in Lemma 4.1. If $\mathcal{G}$ has agreement soundness (ℓ, α), then every set cover of $\mathcal{SC}_{\mathcal{G}}$ is of size more than $|A| \ln |U| (1 - 2\alpha)$.*

Proof. Completeness follows from taking the set cover corresponding to each of the A vertices and its satisfying assignment.

Let us prove soundness. Assume on way of contradiction that there is a set cover C of $\mathcal{SC}_{\mathcal{G}}$ of size at most $|A| \ln |U| (1 - 2\alpha)$. For every $a \in A$ let s_a be the number of sets in C of the form $S_{a,\cdot}$. For every $b \in B$ let s_b be the number of sets in C that participate in covering $\{b\} \times U$. Then, denoting the A-degree of G by D_A,

$$\sum_{b \in B} s_b = \sum_{a \in A} s_a = |A| \ln |U| D_A (1 - 2\alpha) = |B| \ln |U| D (1 - 2\alpha).$$

In other words, on average over the $b \in B$, the universe $\{b\} \times U$ is covered by $D \ln |U| (1 - 2\alpha)$ sets. Therefore, by Markov's inequality, the fraction of $b \in B$ whose universe $\{b\} \times U$ is covered by at most $D \ln |U| (1 - \alpha)$ sets is at least α. By Lemma 4.1 and our construction, for such $b \in B$, there are two edges $e_1 = (a_1, b), e_2 = (a_2, b) \in E$ with $S_{a_1,\sigma_1}, S_{a_2,\sigma_2} \in C$ where $\pi_{e_1}(\sigma_1) = \pi_{e_2}(\sigma_2)$.

We define an assignment $\hat{\varphi}_A : A \to \binom{\Sigma_A}{\ell}$ to the A vertices as follows. For every $a \in A$ pick ℓ different symbols $\sigma \in \Sigma_A$ from those with $S_{a,\sigma} \in C$ (add arbitrary symbols if there are not enough). As we showed, for at least α fraction of the $b \in B$, the A vertices will not totally disagree. □

Fix a constant $0 < \alpha < 1$. The inapproximability ratio we get for Set-Cover from Claim 4.2 is $(1-2\alpha)\ln|U|$, assuming agreement soundness (ℓ, α). The latter is obtained from Corollary 3.4 for $\varepsilon = c/\log^4 n$ for a certain constant $c = c(\alpha)$. Let $N = |U|\,|B|$ be the number of elements in $\mathcal{SC}_{\mathcal{G}}$. We take $|U| = \Theta(|B|^{1/\alpha})$ (we might need to duplicate elements for that), so $\ln|N| = (1+\alpha)\ln|U|$, and the inapproximability ratio is at least $(1-3\alpha)\ln N$. Note that the reduction is polynomial in n. This proves Theorem 3.

5 Open Problems

The inapproximability result we showed for Set-Cover is not tight with respect to the low order terms. It is interesting to prove an $\mathcal{NP}$-hardness result that is tight assuming only the projection games conjecture. We remark that Slavík showed that the greedy algorithm actually obtains an approximation to within $\ln n - \ln\ln n + O(1)$ [Sla96].

As we discussed in the introduction, we believe that many more hardness of approximation results could be proved based on the projection games conjecture. Two concrete open problems are to prove results for Clique and SVP. In an ongoing research we attempt to prove optimal low order terms for 3Lin and other problems from Håstad's paper [Hås01].

Acknowledgments

The author got interested in the question of how to base the Set-Cover hardness of approximation result on projection games following discussions with Ran Raz. The author would also like to thank Scott Aaronson for commenting on an early version of the paper.

References

[BGLR93] M. Bellare, S. Goldwasser, C. Lund, and A. Russell. Efficient probabilistically checkable proofs and applications to approximations. In *Proc. 25th ACM Symp. on Theory of Computing*, pages 294–304, 1993.

[CHK09] M. Charikar, M. Hajiaghayi, and H. Karloff. Improved approximation algorithms for label cover problems. In *ESA*, pages 23–34, 2009.

[Fei98] U. Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4):634–652, 1998.

[Hås99] J. Håstad. Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Mathematica*, 182:105–142, 1999.

[Hås01] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

[Kho01] S. Khot. Improved inapproximability results for maxclique, chromatic number and approximate graph coloring. In *Proc. 42nd IEEE Symp. on Foundations of Computer Science*, pages 600–609, 2001.

[Kho02] S. Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th ACM Symp. on Theory of Computing*, pages 767–775, 2002.

[Kho05] S. Khot. Hardness of approximating the shortest vector problem in lattices. *Journal of the ACM*, 52(5):789–808, 2005.

[MR10] D. Moshkovitz and R. Raz. Two query PCP with sub-constant error. *Journal of the ACM*, 57(5), 2010.

[NSS95] M. Naor, L. J. Schulman, and A. Srinivasan. Splitters and near-optimal derandomization. In *Proc. 36th IEEE Symp. on Foundations of Computer Science*, pages 182–191, 1995.

[Raz98] R. Raz. A parallel repetition theorem. In *SIAM Journal on Computing*, volume 27, pages 763–803, 1998.

[Sla96] Petr Slavík. A tight analysis of the greedy algorithm for set cover. In *Proc. 28th ACM Symp. on Theory of Computing*, pages 435–441, 1996.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de