INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Sur la Conjecture des Jeux Uniques

Stéphane Pérennes — Ignasi Sau

N° 6691

September 2008

Thème COM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6691--FR+ENG

Sur la Conjecture des Jeux Uniques *

Stéphane Pérennes† , Ignasi Sau‡§

Thème COM — Systèmes communicants
Projet MASCOTTE

Rapport de recherche n° 6691 — September 2008 — 8 pages

Résumé : La plupart des problèmes d'optimisation combinatoire sont NP-difficiles, c'est-à-dire qu'ils ne peuvent être résolus en temps polynomial que si les classes P et NP sont identiques. Pour ces problèmes on peut espérer soit trouver des algorithmes d'approximation, soit prouver qu'ils ne peuvent pas être approximés de manière efficace.
En 2002 S. Khot formula la *Conjecture des Jeux Uniques* (UGC), qui géneralise le théorème PCP et impliquerait d'importants résultats d'innaproximabilité pour plusieurs problèmes d'optimisation combinatoire (par exemple MAX CUT ou VERTEX COVER). Intuitivement, la UGC dit que, pour une certaine classe de jeux, appelés *uniques*, il est NP-dur de décider si l'on peut trouver une solution *proche* de l'optimale, ou si toutes les solutions sont *loin* de l'optimale.
Cette conjecture est devenue un problème ouvert des plus importants dans la théorie de la complexité et de l'approximation.
Dans cet article nous étudions un problème très relié à la UGC: MAX-E2-LIN2 dans les graphes bipartis. Dans MAX-E2-LIN2 on a un graphe G ayant deux type d'arêtes, requérant soit la même soit différente couleur pour ses extrémités. Le but est de 2-colorer les sommets de G en maximisant le nombre d'arêtes satisfaites. Nous prouvons que ce problème est APX-complet dans les graphes bipartis et, en utilisant le *Théorème de Répétition Paralèlle*, nous discutons les conséquences de ce résultat dans le cadre des jeux uniques et la UGC.

Mots-clés : Computational Complexity, Unique Games, Parallel Repetition, Hardness of Approximation.

* This work has been partially supported by European project IST FET AEOLUS, PACA region of France, Ministerio de Educación y Ciencia of Spain, European Regional Development Fund under project TEC2005-03575, Catalan Research Council under project 2005SGR00256 and COST action 293 GRAAL, and has been done in the context of the CRC CORSO with France Telecom.

† Mascotte joint project - INRIA/CNRS-I3S/UNSA, Sophia-Antipolis, FRANCE. Stephane.Perennes@sophia.inria.fr

‡ Mascotte joint project - INRIA/CNRS-I3S/UNSA, Sophia-Antipolis, FRANCE. Ignasi.Sau@sophia.inria.fr

§ Graph Theory and Combinatorics group at Applied Mathematics IV Department of UPC, Barcelona, SPAIN.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

A Note on MAX-Ek-LIN2 and the Unique Games Conjecture ¶

Abstract: Most combinatorial optimization problems are NP-hard, i.e. they cannot be solved in polynomial time unless P = NP. Two typical approaches to deal with these problems are either to devise approximation algorithms with good performance, or to prove that such algorithms cannot exist. In 2002 S. Khot stated the *Unique Games Conjecture* (UGC), which generalizes the PCP Theorem. The UGC would imply significant hardness results for several optimization problems (e.g., MAX CUT or VERTEX COVER). Loosely speaking, the UGC states that for a class of games (called *unique*) in which the optimal solution takes values between 0 and 1, it is NP-hard to decide whether the optimal is *close* to 0 or *close* to 1.
This conjecture has become one of the most outstanding open problems in complexity and approximation theory. In this article we study a problem which turns out to be closely related to the UGC: MAX-E2-LIN2 in bipartite graphs. The input of MAX-E2-LIN2 is a graph G with two types of edges. Namely, each edge requires its end-vertices to be colored with either the same color or different colors. The objective is to 2-color the vertices of G maximizing the number of satisfied edges. The problem is known to be APX-complete in general graphs. In this article we prove that the problem remains APX-complete in bipartite graphs and, using the *Parallel Repetition Theorem*, we discuss the consequences that this result could have in the framework of unique games and the UGC.

Key-words: Computational Complexity, Unique Games, Parallel Repetition, Hardness of Approximation.

1 Introduction

A *unique game* consists of an undirected connected graph $G = (V, E)$, a color set C, and for each edge (i, j) with $i < j$ a permutation $\pi_{i,j} : C \to C$. A coloring of the graph $c : V \to C$ *fulfills* (or *satisfies*) an edge (i, j) if $\pi_{i,j}(c(i)) = c(j)$. Trevisan gives a nice introduction to unique games in [10]. If a coloring fulfills all the edges, then knowing the color at one vertex uniquely determines all the other colors. That is the reason why such a game is called *unique*. One can efficiently determine whether such a coloring exists by trying all possible colors at one node to see if any of the resulting coloring fulfills all the edges. However it might be difficult to determine whether one can fulfill some large fraction of the edges. In 2002, Subhash Khot defined the *Unique Games Conjecture* (UGC for short) :

Conjecture 1.1 (Unique Games Conjecture [4]) *For every constant $\delta > 0$ there is a fixed finite color class C such that it is* NP*-hard to distinguish the following 2 cases for any unique game with color class C :*

1. *There is some coloring that fulfills at least a $(1 - \delta)$-fraction of the edges.*
2. *Every coloring fulfills at most a δ-fraction of the edges.*

Since it was first formulated in 2002, the UGC has become one of the most challenging open problems in computational complexity [5, 6, 8, 10]. The UGC is a strengthening of the Probabilistically Checkable Proof (PCP) Theorem. If the UGC were true, it would have important consequences for the theory of complexity and approximation. For instance, it would imply that any improvement in approximating Max Cut below 0.878567 would force P = NP [6], and that Vertex Cover would be hard to approximate within 2-ε [7].

In this article we study a problem which seems to be closely related to a particular case of the UGC. Namely, in Section 2 we prove that Max-E2-Lin2 in bipartite graphs is APX-complete, by a reduction from 3-Max-Cut. In Section 3 we discuss the relation of this result with the UGC, in order to understand better the UGC, or even to provide an intermediate step towards a hypothetic proof of the UGC.

2 Max-E2-Lin2 in Bipartite Graphs is APX-complete

In this section we focus on the following classical NP-complete problem :

Max-Ek-Lin2 :
Input : a set of linear equations modulo 2 with exactly k variables per equation.
Output : an assignment of the variables maximizing the number of equations satisfied.

The case when $k = 2$ can be seen in a natural way as a graph optimization problem as follows : each variable corresponds to a vertex, and each linear equation (involving exactly 2 variables) corresponds to an edge. There are two types of edges, namely those that require their endpoints to have the *same* value (either 0 or 1), and those that require their endpoints to have *different* value. In what follows, we call such edges of type S and type D, respectively. Let also $|S|$ and $|D|$ be the number of edges of type S and D of the input graph G, respectively. In this context, an optimal solution is a 2-coloring of the vertices of G maximizing the number of satisfied edges. In the edge-weighted version, the objective function to be maximized is the sum of the weights of the satisfied edges.

Karpinski proved in [3] that MAX-E2-LIN2 is APX-complete in general graphs. It is natural to ask what happens to complexity when we restrict the input graph to being bipartite (MAX-E2-LIN2-BIP for short). We prove in Theorem 2.1 that MAX-E2-LIN2-BIP remains APX-complete. The main reason why we study this problem, besides the fact that the hardness result is interesting by itself, is its close relation with the UGC, as we shall see in Section 3. Before going through the proof of Theorem 2.1, we need to introduce another classical NP-complete problem :

d-MAX-CUT :
Input : un undirected graph $G = (V, E)$ of degree bounded by d.
Output : a partition of V into two groups so as to maximize the number of edges with exactly one endpoint in each group.

Observe that solving MAX-CUT in a graph G is equivalent to finding a bipartite subgraph with maximum number of edges, since any bipartite subgraph $H \subseteq G$ with maximum number of edges can be transformed into another (not necessarily induced) bipartite subgraph $H' \subseteq G$ containing H such that $|E(H')| \geq |E(H)|$ and $V(H') = V(G)$.

Karpinski proved in [3] that 3-MAX-CUT is APX-complete, providing an approximation algorithm with approximation ratio 1.0858, together with a hardness lower bound of 1.003.

Theorem 2.1 MAX-E2-LIN2-BIP *is* APX-*complete.*

Proof: To see that MAX-E2-LIN2-BIP belongs to APX, it is easy to satisfy $\max\{|S|, |D|\}p$ edges of the input bipartite graph $G = (A \cup B, E)$ just by first coloring all the vertices of G with the same color (fulfilling $|S|$ edges), then coloring the vertices in A with color 0 and the vertices in B with color 1 (fulfilling $|D|$ edges), and taking the best solution. This naïve algorithm clearly provides a 2-approximation.

We prove that MAX-E2-LIN2-BIP is APX-hard with a gap-preserving reduction from 3-MAX-CUT, proved to be APX-hard in [3]. We first deal with the edge-weighted version, and then we describe how to adapt the reduction to the unweighted case.

Given an input graph $G = (V, E)$ of 3-MAX-CUT with maximum degree at most 3, we construct an instance G' of MAX-E2-LIN2-BIP in the following way : we put two copies V_1 and V_2 of the vertex set V, i.e. there are two vertices $u_1 \in V_1$ and $u_2 \in V_2$ for each vertex $u \in V$. (With slight abuse of notation, we will suppose henceforth that to each pair $\{u_1, v_2\} \in V_1 \times V_2$ corresponds without ambiguity a pair $\{u, v\} \in V \times V$, with possibly $u = v$.) Then we add an edge of type D with weight 1 between $u_1 \in V_1$ and $v_2 \in V_2$ if and only if $(u, v) \in E$, and an edge of type S with weight 4 between $u_1 \in V_1$ and $v_2 \in V_2$ if and only if $u = v$. This completes the construction of $G' = (V_1 \cup V_2, E')$, which is illustrated with a small example in Fig. 1a. Let $n = |V(G)|$.

Claim 1 *Any optimal solution of* MAX-E2-LIN2-BIP *in G' contains the n edges of type S.*

Let us see that given any solution H of MAX-E2-LIN2-BIP in G', we can transform H into a better (in terms of the weight of the satisfied edges) solution H' containing the n edges

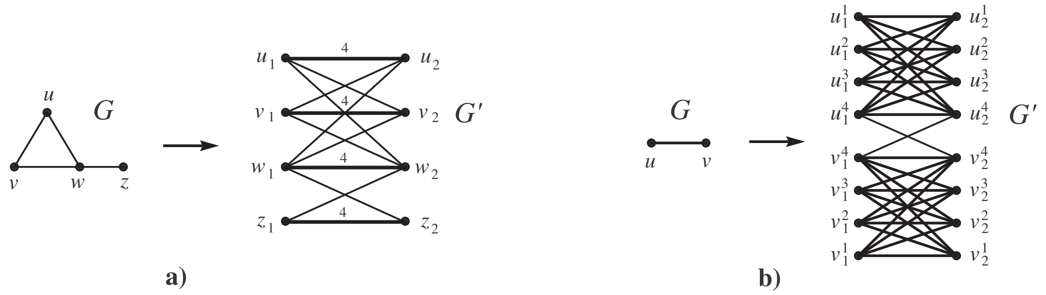

FIG. 1 – Reduction in the proof of Theorem 2.1 : **a)** for the weighted case ; and **b)** for the unweighted case. The thick edges are of type S, the others being of type D. If no weight is displayed beside an edge, it is assumed to be 1.

of type S. Indeed, if for a vertex u the edge (u_1, u_2) is not satisfied in H, then u_1 and u_2 have different color in H. Since the degree of u in G is at most 3, if we change the color of u_1 (let H' be this new subgraph) then at most 3 edges with weight 1 become unsatisfied in H', but the edge (u_1, u_2) with weight 4 is satisfied in H'. Thus, the weight of the satisfied edges of H' is strictly greater than the weight of those of H.

Let OPT_G be the optimal value of 3-MAX-CUT in G, and let $OPT_{G'}$ be the optimal value of the weighted version of MAX-E2-LIN2-BIP in G'.

Claim 2 $OPT_{G'} = 2 \cdot OPT_G + 4n$.

By Claim 1, in an optimal solution $H \subseteq G'$, for each vertex $u \in V(G)$, both vertices u_1 and u_2 are colored with the same color. So if the edge (u_1, v_2) is satisfied in H, then the edge (u_2, v_1) is also satisfied. Therefore, $OPT_{G'}$ is achieved by a 2-coloring of the vertices of G maximizing the number of edges with exactly one endpoint in each color class. This problem is exactly 3-MAX-CUT in G, so the claim follows.

By Claim 2, it is clear that the existence of a PTAS for MAX-E2-LIN2-BIP would imply the existence of a PTAS for 3-MAX-CUT, which is a contradiction.

For the unweighted case, we modify the construction of G' in the following way : for each vertex $u \in V(G)$, we put 8 copies $u_1^1, \ldots, u_1^4 \in V_1$ and $u_2^1, \ldots, u_2^4 \in V_2$. Then we add a complete bipartite graph with edges of type S between the vertices $u_1^1, \ldots, u_1^4 \in V_1$ and $u_2^1, \ldots, u_2^4 \in V_2$. Finally, we add an edge of type D between $u_1^4 \in V_1$ and $v_2^4 \in V_2$ if and only if $(u, v) \in E$ (see Fig. 1b). Once again, any optimal solution for G' contains all the $16n$ edges of type S, and $OPT_{G'} = 2 \cdot OPT_G + 16n$. The theorem follows. $\square$

3 Relation with the Unique Games Conjecture

In this section we discuss the relation of Theorem 2.1 with the UGC. First we need to introduce the *Parallel Repetition Theorem*, proved by Raz in [9] and further simplified by Holenstein in [2]. This result states that a parallel repetition of any *two-prover one-round proof system* (it can be seen as a unique game in a bipartite graph, see [1]) decreases the probability of error at an exponential rate.

Max-E2-Lin2-Bip can be naturally thought of as a unique game $\mathcal{G}$ with $|C| = 2$ colors, the permutation for each edge being either the identity or a transposition. The parallel repetition of k times $\mathcal{G}$ corresponds to a unique game with $|C| = 2^k$ colors. Indeed, let us describe the game when we repeat $\mathcal{G}$ once. Given the input graph G of Max-E2-Lin2-Bip, we consider the direct product $G \times G$, and then 4 types of edges can appear, namely SS, SD, DS, and DD. We have that $C = \mathbb{Z}_2^2$ and $|C| = 4$ (more generally, if we repeat k times $\mathcal{G}$ we have that $C = \mathbb{Z}_2^k$ and $|C| = 2^k$). This game is *abelian* in the sense that, for instance, an edge $((a, x), (b, y)) \in E(G \times G)$ of type SD is satisfied if and only if $c(a) + c(b) \equiv 0 \pmod 2$ **and** $c(x) + c(y) \equiv 1 \pmod 2$. Observe that in this particular game all the bijections are translations.

Using the notation of [2,9], let $v < 1$ be the optimal value of $\mathcal{G}$. This constant v can be seen as the maximum probability that both players can win, each player corresponding to one stable set of the bipartite graph G. This probability v corresponds to the percentage of satisfied edges of an optimal solution in G. The Parallel Repetition Theorem states that when we repeat k times $\mathcal{G}$ in parallel, then the probability that both players can win *all* games simultaneously is at most $\overline{v}^k$, where $\overline{v} < 1$ is a constant depending only on v.

On the one hand, the UGC states that for every fixed constant $\delta > 0$, there exists a color class C such that the *gap* (given by cases 1 and 2 in Conjecture 1.1) for any unique game with color class C is $[\delta, 1 - \delta]$.

On the other hand, Theorem 2.1 states that there exists a constant $0 < \alpha < 1$ such that it is impossible (unless P = NP) to find in polynomial time a solution *better* than $(1 - \alpha)v$, in terms of the percentage of satisfied edges. This APX-hardness result can be seen as a gap $[(1 - \alpha)v, v]$ for the game $\mathcal{G}$, in the sense that in this interval we have *no knowledge* about the optimal solution.

In order to get close to the UGC, we need to be able to say something more about this gap for $\mathcal{G}$. If one managed to prove that the gap for the game $\mathcal{G}$ corresponding to Max-E2-Lin2-Bip were, for instance, $[1 - 1/k^2, v^{1/k}]$, then the Parallel Repetition Theorem would imply that the gap of repeating k times $\mathcal{G}$ would be, asymptotically, $[1 - 1/k, v]$. Since for any fixed constant $\delta > 0$ there exists an integer k such that $1/k < \delta$, the interval $[1 - 1/k, v]$ would be contained in $[1 - \delta, 1]$, and the *upper* part of the gap given by the UGC would be covered. Finally, the same kind of argument could be adapted to cover the *lower* part of the gap.

Let $\hat{v}$ be a constant such that $1 - 1/k^2 < \hat{v} < 1$. Summarizing, the important question for $\mathcal{G}$ is the following : Provided that we know that a proportion at least $(1 - 1/k^2)$ of the edges of G can be satisfied, is it NP-hard to satisfy a proportion $\hat{v}$ of the edges ? In other words, the problem is whether Max-E2-Lin2-Bip can be efficiently approximated when one knows that in an optimal solution *almost* all edges are satisfied. Equivalently, is it possible to find a bipartition of a graph when it is *almost* bipartite ?

The UGC for this class of games would be true if we were able to have more information concerning the two cases of MAX-E2-LIN2-BIP, i.e. the case where any solution satisfies very few edges, and the case where there exists a solution satisfying almost all edges. We have seen in the proof of Theorem 2.1 that, modulo an additive factor, MAX-E2-LIN2-BIP looks like 3-MAX-CUT, so a better understanding of MAX-CUT could also provide important insights into the UGC.

In conclusion, we showed that the problem of MAX-E2-LIN2 is APX-complete in bipartite graphs, and we saw that the parallel repetition of the game corresponding to this problem naturally leads to unique games. Due to this strong relation, it would be very interesting to understand the hardness of MAX-E2-LIN2 better, in order to shed some light on the Unique Games Conjecture concerning unique games.

References

[1] M. Ben-Or, S. Goldwaser, J. Kilian, and A. Widgerson. Multi prover interactive proofs : How to remove intractability. In *STOC'88 : Proceedings of the 20th annual ACM Symposium on Theory of Computing*, pages 113–131, 1988.

[2] T. Holenstein. Parallel repetition : simplifications and the no-signaling case. In *STOC'07 : Proceedings of the 39th annual ACM Symposium on Theory of Computing*, pages 411–419, 2007.

[3] M. Karpinski. Approximating Bounded Degree Instances of NP-Hard Problems. In *FCT'01 : Proceedings of the 13th International Symposium on Fundamentals of Computation Theory*, pages 24–34, 2001.

[4] S. 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, 2002.

[5] S. Khot. On the Unique Games Conjecture. In *FOCS'05 : Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, page 3, 2005.

[6] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal Inapproximability Results for Max-Cut and Other 2-Variable CSPs ? In *FOCS'04 : Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 146–154, 2004.

[7] S. Khot and O. Regev. Vertex cover might be hard to approximate within 2-ε. *Journal of Computer and System Sciences*, 74(3) :335–349, 2008.

[8] S. Khot and N. Vishnoi. The Unique Games Conjecture, Integrality Gap for Cut Problems and Embeddability of Negative Type Metrics into ℓ_1. In *FOCS'05 : Proceedings of*

the 46th Annual IEEE Symposium on Foundations of Computer Science, pages 53–62, 2005.

[9] R. Raz. A Parallel Repetition Theorem. *SIAM Journal on Computing*, 27(3) :763–803, 1998.

[10] L. Trevisan. Approximation Algorithms for Unique Games. In *FOCS'05 : Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 197–205, 2005.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399