Hardness of Approximation

Subhash Khot *

Abstract. This article accompanies the talk given by the author at the International Congress of Mathematicians, 2014. The article sketches some connections between approximability of NP-complete problems, analysis and geometry, and the role played by the Unique Games Conjecture in facilitating these connections. For a more extensive introduction to the topic, the reader is referred to survey articles [39, 40, 64].

Mathematics Subject Classification (2010). Primary 68Q17.

Keywords. NP-completeness, Approximation algorithms, Inapproximability, Probabilistically Checkable Proofs, Discrete Fourier analysis.

1. Introduction

The P $\neq$ NP hypothesis says that a large class of computational problems known as NP-complete problems do not have efficient algorithms. An algorithm is called efficient if it runs in time polynomial in the size of the input, typically denoted as n. A natural question is whether one can efficiently compute *approximate* solutions to NP-complete problems and how good an approximation one can achieve. We are interested in both upper and lower bounds: designing algorithms with a guarantee on the quality of approximation (upper bounds) as well as results showing that no efficient algorithm exists that achieves an approximation guarantee beyond a certain threshold (lower bounds). It is the latter question, namely the lower bounds, that is the focus of this article. Such results are known as *inapproximability* or *hardness of approximation* results, proved under a standard complexity theoretic hypothesis such as P $\neq$ NP.

Let us consider two problems, the Traveling Salesperson (TSP) and the Clique, as illustration. In the (2-dimensional Euclidean version of) TSP problem, we are given a set of n cities in a plane and the pairwise distances between them and the goal is to find a tour that visits all the cities and has minimum length. In the Clique problem, we are given an n-vertex graph and the goal is find a clique of maximum size where a clique is a subset of vertices such that all its vertices are pairwise connected by edges. Both the problems are NP-complete[1] and hence one does not hope to efficiently find optimal solutions. Now consider the question of how well

*This work is supported by NSF grants CCF-0832795, 1061938, 1422159, and Simons Collaboration on Algorithms and Geometry grant.

[1]There are some subtleties regarding the computational complexity of TSP that we omit here.

one can approximate them. For the TSP problem, for every constant $\varepsilon > 0$, Arora and Mitchell [1, 52] designed a polynomial time algorithm that computes a tour with length at most $1 + \varepsilon$ times the length of the minimum tour. For the Clique problem, Håstad [35] showed that it cannot be approximated at all. Specifically, for every constant $\varepsilon > 0$, assuming P $\neq$ NP, no polynomial time algorithm, given an n-vertex graph that has a clique of size at least $n^{1-\varepsilon}$, can find a clique of size even n^{ε}. Thus, we know the precise extent to which the TSP and the Clique problems are approximable: the former is approximable as well as one might hope for and the latter is not approximable at all. There are a few more problems for which also we know the precise extent of approximability. In particular, for the 3SAT and the Set Cover problems [36, 30], we know an approximation algorithm that achieves a reasonable (but not too close as TSP) approximation guarantee and we also know that achieving an approximation better than this *threshold* guarantee is an NP-complete problem itself. To emphasize, the last statement implies that an algorithm with approximation guarantee better than the threshold for these problems can then be used to find optimal solutions!

However, for a vast majority of the NP-complete problems of interest, there is (often a big) gap between the quality of the best known approximation algorithms and the known hardness of approximation results. Filling up these gaps, as well as understanding why different NP-complete problems seem to behave differently in terms of their approximability, is largely open. The Unique Games Conjecture was proposed towards making progress on this topic, and in particular towards showing optimal hardness of approximation results, i.e. results that match the quality of the best known approximation algorithms. As it turns out, showing hardness results is closely related to Fourier analysis of boolean functions on a boolean hypercube and to certain problems in geometry, especially related to isoperimetry. This article gives a sketch of some of these connections and cites a couple of open questions towards settling the Unique Games Conjecture. We anticipate that the intended audience of this article is not necessarily familiar with the language and techniques in computer science, so an attempt is made to keep the presentation as self-contained as possible.

2. The Unique Games Conjecture

The Unique Games Conjecture [38] states that a certain computational problem called the Unique Game is very hard to approximate. We do state the conjecture here, but we will not really use the statement in the rest of the article.

An instance $\mathcal{L}$ of the Unique Game problem is a system of linear equations over $\mathbb{Z}_p$ of a specific form. There are n variables $x_1, \ldots, x_n$ and m equations, where i^{th} equation is of the form $x_{i_1} - x_{i_2} = c_i$. The constants $c_i \in \mathbb{Z}_p$ may depend on the equation. The goal is to find an assignment to the variables that satisfies a *good* fraction of the equations. Let $\mathsf{OPT}(\mathcal{L})$ denote the maximum fraction of equations satisfied by any assignment. The Unique Games Conjecture states:

Conjecture 2.1. *For every constant $\delta > 0$, there is a large enough constant $p = p(\delta)$, such that there is no polynomial time algorithm that given an instance of* Unique Game *over $\mathbb{Z}_p$ that has an assignment satisfying $1 - \delta$ fraction of the equations, finds an assignment that satisfies (even) δ fraction of the equations.*[2]

A few comments are in order. The term *game* refers to the context of *2-prover-1-round games* where the problem was studied initially. Given an instance of the Unique Game as above, consider the following game between two provers and a verifier: the verifier picks an equation $x_{i_1} - x_{i_2} = c_i$ at random, sends the variable x_{i_1} to prover P_1 and the variable x_{i_2} to prover P_2. Each prover is supposed to answer with a value in $\mathbb{Z}_p$, and the verifier accepts if and only if $a_1 - a_2 = c_i$ where a_1 and a_2 are the answers of the two provers respectively. The strategies of the provers correspond to assignments $\sigma_1, \sigma_2 : \{x_1, \ldots, x_n\} \mapsto \mathbb{Z}_p$. The *value* of the game is the maximum over all prover strategies, the probability that the verifier accepts. It is not difficult to show that this value is between $\mathsf{OPT}(\mathcal{L})$ and $\max\{1, 4\mathsf{OPT}(\mathcal{L})\}$. Such games were initially motivated by the study of cryptographic protocols. The term *unique* refers to the property of the equations $x_{i_1} - x_{i_2} = c_i$ that for every assignment to one variable, there is a unique assignment to the other variable so that the equation is satisfied. Unique Games were studied before in literature, in particular by Feige and Lovász [31] in the context of *parallel repetition*.

The important feature of the Unique Games is that the equations are linear. If one allows equations of arbitrary degree, each equation still depending on two variables, the problem may be referred to as a Non-Unique Game. The statement analogous to Conjecture 2.1 is known to hold for Non-Unique Games (and is very useful). It follows from a combination of the PCP Theorem stated in the next section and the Parallel Repetition Theorem of Raz [61]. For Non-Unique Games, the statement holds even on instances that have an assignment that satisfies all equations, as opposed to only $1 - \delta$ fraction of the equations. Moreover, one only needs p to be polynomially large in $\frac{1}{\delta}$. For the Unique Games however, if there were an assignment that satisfies all equations, it can be efficiently found (an easy observation). Hence, it is essential in Conjecture 2.1 to consider only $1 - \delta$ satisfiable instances. Moreover, if the conjecture were correct, it is known that p would have to be at least exponentially large in $\frac{1}{\delta}$ [17].

[2]The original conjecture is stated in terms of a more general problem and strictly speaking the term Unique Game refers to the general problem. The problem presented here is referred to as Linear Unique Game. It is shown in [41] that the original conjecture is equivalent to the statement here. Also, the problem is conjectured to be NP-complete, rather than just that there is no polynomial time algorithm for it. It is widely believed that NP-complete problems do not have algorithms that run in time $2^{n^{o(1)}}$, rather than just in polynomial time. For much of the article, when we say "there is no polynomial time algorithm for a problem", we really mean "the problem is NP-complete".

3. The Max-Cut Problem, the PCP Theorem, the GW-Algorithm and its Optimality

The Unique Games Conjecture states that the Unique Game problem is hard to approximate. It has been shown that for several other optimization problems of interest, denoting a typical such problem by Π, there is a *reduction* from the Unique Game problem to the problem Π and as a consequence, the problem Π is hard to approximate as well. We sketch one such reduction below and refer the reader to [40] for a list of several reductions of this kind. We note that prior to formulation of the Unique Games Conjecture, researchers had already developed a general framework for similar reductions and techniques to analyze them [2, 61, 13, 35, 36, 27], with some remarkable successes such as Håstad's Clique result mentioned in the introduction. However these prior reductions were from the Non-Unique Game problem. For several problems Π of interest, we do not know how to reduce the Non-Unique Game problem to Π, but we do know how to reduce the Unique Game problem to Π. The Unique Game problem seems to strike a delicate balance: it has a simple enough structure that it is a convenient problem to reduce from and has a complex enough structure that it is plausibly a hard problem.

In this article, we focus on one specific optimization problem, namely the Max-Cut problem, and use it as an illustrative example throughout the article. In this problem, we are given a graph $G(V, E)$ and the goal is to find a cut, i.e. partition of the vertex set V into two disjoint sets V_1 and V_2, so as to maximize the number of edges cut. An edge is said to be cut if its one endpoint is in V_1 and the other endpoint is in V_2. The size of the cut is the fraction of edges cut. Let $\mathsf{OPT}(G)$ denote the maximum size of any cut. We will focus on a particular special case of the problem when the graph G is almost bipartite, i.e. it has a cut that cuts almost all the edges. Let $\varepsilon > 0$ be a small enough positive constant. The following problem will be the focus of the rest of the article.

Max-Cut **Problem:** Given a graph $G(V, E)$ such that $\mathsf{OPT}(G) = 1 - \varepsilon$. Find (efficiently) a cut of as large size as possible.

We will be interested in the computational complexity of this problem. A couple of observations are immediate. Firstly, the Max-Cut problem is NP-complete and hence one cannot hope to efficiently find a cut of the maximum size, i.e. of size $1-\varepsilon$.[3] Secondly, one can easily find a cut of size $\frac{1}{2}$. Simply take a uniformly random cut in the graph; it cuts a fraction $\frac{1}{2}$ of the edges in expectation and this randomized algorithm, if desired, can be turned into a deterministic algorithm as well. Till early 90's, this was all that was known regarding what is computationally infeasible and what is feasible. Two breakthrough results then led to a significant progress on this question: one from the hardness side, known as the PCP Theorem, and the other from the algorithmic side, namely the Goemans-Williamson's algorithm.

From the hardness side, the PCP Theorem [29, 7, 5] implies that it is not only hard to find a cut of the maximum size, but also hard to find a cut of near-maximum

[3] The standard NP-completeness reduction to the Max-Cut problem can be easily modified so that it holds on graphs with maximum cut of size $1 - \varepsilon$.

size. Specifically:[4]

The PCP Theorem: Assume P $\neq$ NP. Then there is an absolute constant $\beta > 1$ such that no polynomial time algorithm, given a graph that has a cut of size $1-\varepsilon$, can find a cut of size $1-\beta\varepsilon$.

The PCP Theorem is stated above as a hardness of approximation result. The acronym PCP stands for *Probabilistically Checkable Proofs* and indeed there is an equivalent formulation of the theorem in terms of *proof checking* (and this is what led to its discovery, as a culmination of much prior work on *interactive proofs*). The theorem states that every NP statement has a polynomial size proof that can be checked by a probabilistic polynomial time verifier by reading only a constant number of bits in the proof! The verifier has the completeness and the soundness property: every correct statement has a proof that is accepted with probability 1 and every proof of an incorrect statement is accepted with only a small probability, say at most 1%. The equivalence between the two viewpoints, namely the hardness viewpoint and the proof checking viewpoint, is simple but illuminating, and has influenced much of the work in this area. In this article, we restrict ourselves to the hardness viewpoint, i.e. the hardness result for the Max-Cut problem as stated above.

From the algorithmic side, Goemans and Williamson [33] designed an efficient algorithm that given a graph $G(V,E)$ with a cut of size $1-\varepsilon$, finds a cut of size $1-\frac{1}{\pi}\arccos(1-2\varepsilon)$. The latter quantity is approximated as $1-\frac{2}{\pi}\sqrt{\varepsilon}-O(\varepsilon^{3/2})$. We provide a high-level sketch of the Goemans-Williamson's algorithm.[5] The algorithm proceeds by computing an embedding $\phi : V \mapsto \mathbb{S}^{m-1}$ of the set of vertices onto a unit sphere in $\mathbb{R}^m$. The dimension m is unrestricted, but w.l.o.g. can be assumed to be at most $|V|$. The embedding is computed by solving a so-called *semi-definite programming (SDP) relaxation* of the problem instance. We omit the description of this step (see [34] for introduction to SDPs and their algorithmic applications), but state the crucial property of the embedding: for most of the edges (u,v) in the graph, the endpoints u,v are embedded as points $\phi(u),\phi(v)$ on the sphere that are nearly antipodal points. Once the embedding has been computed, the algorithm selects a hyperplane H in $\mathbb{R}^m$ passing through the origin, uniformly at random from the set of all such hyperplanes. The hyperplane H cuts the sphere into two parts, which in turn induces a partition of the set V into two parts, depending on which side of the hyperplane the point $\phi(v)$ lies, for a vertex $v \in V$. This yields the desired cut in the graph. The analysis of the algorithm then shows that the expected size of the cut is at least $1-\frac{1}{\pi}\arccos(1-2\varepsilon)$. Using the approximation cited before, this is at least $1-\sqrt{\varepsilon}$ for small enough ε.

In spite of the progress offered by the PCP Theorem and Goemans-Williamson's algorithm, there is still a gap between $1-\beta\varepsilon$ and $1-\sqrt{\varepsilon}$, regarding the size of the cut that is infeasible to compute and feasible to compute. Bridging this gap

[4] The PCP Theorem actually proves that the stated computational task is NP-complete.

[5] Often we are interested in quality of approximation measured as a multiplicative factor. The minimum value of the ratio, over all $\varepsilon \in (0,1)$, between $1-\frac{1}{\pi}\arccos(1-2\varepsilon)$ and $1-\varepsilon$ is ≈ 0.878 and the Goemans-Williamson's algorithm is often cited as a 0.878-approximation to Max-Cut.

turns out to be an interesting pursuit as we demonstrate in this article. In particular, one could ask whether the Goemans-Williamson algorithm is the best possible algorithm in terms of its approximation guarantee. To the best of author's information, when the Goemans-Williamson algorithm was discovered, it was viewed as somewhat unnatural and roundabout way of solving a combinatorial problem via a geometric method, and it was believed that a better algorithm would follow soon. However, rather surprisingly, assuming the Unique Games Conjecture, Goemans-Williamson's algorithm is indeed optimal [41]:

Theorem 3.1. *Assume the Unique Games Conjecture. Fix any $\varepsilon \in (0, \frac{1}{2})$ and let $\eta > 0$ be an arbitrarily small constant. Then there is no polynomial time algorithm that given a graph with a cut of size at least $1-\varepsilon$, finds a cut of size $1 - \frac{1}{\pi}\arccos(1-2\varepsilon) + \eta$.*

Approximating $1 - \frac{1}{\pi}\arccos(1-2\varepsilon)$ as before, it will be convenient to focus on a (slightly weaker) statement: assuming the Unique Games Conjecture, there is no polynomial time algorithm that given a graph with a cut of size at least $1-\varepsilon$, finds a cut of size $1 - \frac{1}{2}\sqrt{\varepsilon}$. Such a statement is proved by reducing the Unique Game problem to the Max-Cut problem. A reduction is a polynomial time procedure that starts with an instance $\mathcal{L}$ of the Unique Game problem (i.e. a system of linear equations over $\mathbb{Z}_p$ with two variables per equation) and builds an instance G of the Max-Cut problem (i.e. a graph) such that finding a large cut in G amounts to finding a good approximate solution to the system $\mathcal{L}$. Since the Unique Games Conjecture states that the latter task is computationally infeasible, so is the former. Specifically, the correctness of such a reduction consists of two statements, referred to as the completeness and the soundness statements: for given $\varepsilon > 0$, if $\delta > 0$ is small enough,

$$\begin{array}{lccccccc} (\textit{Completeness}) & \mathsf{OPT}(\mathcal{L}) & \geq & 1-\delta & \implies & \mathsf{OPT}(G) & \geq & 1-\varepsilon. \\ (\textit{Soundness}) & \mathsf{OPT}(\mathcal{L}) & \leq & \delta & \implies & \mathsf{OPT}(G) & \leq & 1-\frac{1}{2}\sqrt{\varepsilon}. \end{array}$$

Now, the Unique Games Conjecture states that there is no polynomial time algorithm that given a $(1-\delta)$-satisfiable system $\mathcal{L}$, finds a δ-satisfying assignment. If the conjecture is correct, it then follows[6] that there is no polynomial time algorithm that given a graph with a cut of size $1-\varepsilon$, finds a cut of size $1 - \frac{1}{2}\sqrt{\varepsilon}$.

We only provide a glimpse of the reduction here. Let $\mathcal{L}$ be the given linear system over $\mathbb{Z}_p$. The reduction constructs, for every variable x_i in the linear system, a group of 2^p vertices C_i labeled by boolean strings $\sigma \in \{-1,1\}^p$. For every equation $x_{i_1} - x_{i_2} = c_i$ in the linear system, there are edges between the group C_{i_1} and the group C_{i_2}. Roughly speaking, there is an edge between a vertex σ in group C_{i_1} and a vertex τ in group C_{i_2} (here both σ, τ are boolean strings of length p) if

$$\mid \{\ell \in \{1,\ldots,p\} \mid \sigma_{\ell+c_i} \neq \tau_\ell\} \mid \approx (1-\varepsilon)p.$$

[6]Strictly speaking, this implication is not immediate just from the completeness and the soundness statements. Formally, one shows that a cut of size at least $1 - \frac{1}{2}\sqrt{\varepsilon}$ in G can be used, in polynomial time, to find a δ-satisfying assignment to $\mathcal{L}$. Most reductions are constructive in this sense.

If one considers the special case when the equation is $x_{i_1} - x_{i_2} = 0$, then the last condition is same as saying that σ and τ have Hamming distance $\approx (1-\varepsilon)p$.

We omit the proofs of the completeness and the soundness properties. The proof of the completeness property is actually immediate from the construction. Proving the soundness property takes some work and though we omit its proof, we describe a key ingredient known as the Majority Is Stablest Theorem. This is a theorem about *noise-stability* of boolean functions on a boolean hypercube, i.e. of functions $f : \{-1,1\}^p \mapsto \{-1,1\}$. Any such function can be viewed as a cut in the set of vertices $\{-1,1\}^p$ and this is how one relates the theorem to the proof of the soundness of the reduction above. We present the Majority Is Stablest Theorem as well as a sketch of its proof, illustrating the connections to probability and Gaussian iso-perimetry.

4. Majority is Stablest and Gaussian Isoperimetry

Suppose $f : \{-1,1\}^n \mapsto \{-1,1\}$ is a boolean function. Every such function can be viewed as a pre-determined rule to decide outcome of an election, also referred to as a *voting scheme*: consider an election with n voters and two candidates labeled as $\{-1,1\}$. The n voters vote for either of these candidates, uniformly and independently at random. If $x_1, \ldots, x_n \in \{-1,1\}$ denotes the sequence of their votes, the winner of the election is declared to be $f(x_1, \ldots, x_n)$. We focus on a voting scheme f that is *balanced*, i.e. both the candidates have equal chance of winning the election, and *democratic*, i.e. no individual voter has significant influence on the outcome of the election (formalized below). One example is the majority function $\mathsf{MAJ}_n = \mathsf{sign}(x_1 + \ldots + x_n)$ that corresponds to taking majority vote (say n is odd). Another example is *majority of majorities* that roughly corresponds to the electoral college system. We desire a voting scheme that is noise-stable, i.e. if a small fraction of votes are corrupted at random, then the probability that the outcome of the election changes is small. The Majority Is Stablest Theorem states that among all balanced and democratic voting schemes, the majority function is the most noise stable (up to a negligible additive error).

Formally, let $f : \{-1,1\}^n \mapsto \{-1,1\}$ be a balanced boolean function, i.e. $\Pr_x[f(x) = 1] = \Pr_x[f(x) = -1] = \frac{1}{2}$, where the choice of input x is uniformly random over $\{-1,1\}^n$. For a co-ordinate $i \in \{1,2,\ldots,n\}$, let the influence of the i^{th} co-ordinate on the function f be defined as:

$$\mathrm{Infl}_i(f) := \Pr_x \left[f(x_1, \ldots, x_i, \ldots, x_n) \neq f(x_1, \ldots, -x_i, \ldots, x_n) \right].$$

This is the probability that the function changes its value when the i^{th} co-ordinate is flipped, starting with a uniformly chosen input. A function is democratic if the influence of every co-ordinate is small. Let $\varepsilon \in (0, \frac{1}{2})$ be a noise parameter. The ε-noise stability of the function f is defined as

$$\mathsf{Stab}_\varepsilon(f) := \Pr_{x,y \sim \mathsf{N}_\varepsilon(x)} \left[f(x) = f(y) \right], \tag{1}$$

where x is a uniformly chosen input and y is chosen from the distribution $\mathsf{N}_\varepsilon(x)$ obtained by flipping every co-ordinate of x independently with probability ε (thus y is a perturbed or noisy version of x). It is known that the noise stability of the majority function MAJ_n tends to $1 - \frac{1}{\pi}\arccos(1-2\varepsilon)$ as $n \to \infty$. The Majority Is Stablest Theorem, proved by Mossel, O'Donnell, and Oleszkiewicz [54] (and conjectured in [41]) states that the noise stability of any balanced, democratic function is at most that of the majority function up to a negligible additive error.

Theorem 4.1. *Let $\varepsilon \in (0, \frac{1}{2})$ be a noise parameter and $\delta > 0$ be an arbitrarily small error parameter. Then for a sufficiently small constant $\eta > 0$, any balanced function $f : \{-1,1\}^n \mapsto \{-1,1\}$ such that $\forall i \in \{1,2,\ldots,n\}$, $\mathrm{Infl}_i(f) \leq \eta$, satisfies:*

$$\mathsf{Stab}_\varepsilon(f) \leq 1 - \frac{1}{\pi}\arccos(1-2\varepsilon) + \delta.$$

We present a sketch of the proof as it demonstrates the connection to an isoperimetric problem in geometry and its solution by Borell [15]. The proof involves an application of the *invariance principle* [62, 54, 19, 53]. Before we state the invariance principle, we note a few well-known facts. Any function $f : \{-1,1\}^n \mapsto \mathbb{R}$ can be represented as a multi-linear polynomial (Fourier or Walsh representation):

$$f(x) = \sum_{S \subseteq \{1,\ldots,n\}} \widehat{f}(S) \prod_{i \in S} x_i,$$

where $\widehat{f}(S) \in \mathbb{R}$ are the Fourier coefficients. When f is a boolean function, by Parseval's identity, $\sum_S \widehat{f}(S)^2 = \mathbb{E}_x[f(x)^2] = 1$. It is easily proved that

$$\mathrm{Infl}_i(f) = \sum_{i \in S} \widehat{f}(S)^2 \qquad \text{and} \qquad \mathsf{Stab}_\varepsilon(f) = \frac{1}{2} + \frac{1}{2}\sum_S \widehat{f}(S)^2(1-2\varepsilon)^{|S|}. \tag{2}$$

Using these formulas, the notion of influence and noise-stability can be extended to all multi-linear polynomials (and not just those representing boolean functions). Here is a rough statement of the invariance principle:

Invariance Principle: *Suppose f is a low degree multi-linear polynomial in n variables and all its variables have small enough influence. Then the distribution of the values of f is nearly identical when the input is a uniform random point from $\{-1,1\}^n$ or a random point from $\mathbb{R}^n$ with the standard Gaussian measure.*

To motivate the invariance principle, one considers the case when $f = \sum_{i=1}^n a_i x_i$ is a linear polynomial. Assume w.l.o.g. that $\sum_{i=1}^n a_i^2 = 1$. The condition that all variables have small influence is equivalent to the condition that $|a_i|$ is small for all $i \in \{1,\ldots,n\}$. The invariance principle, in this case, states that the distribution of values of $f(x_1,\ldots,x_n)$ where x_i are i.i.d. $\{-1,1\}$ and the distribution of values of $f(x_1^*,\ldots,x_n^*)$ where x_i^* are i.i.d. standard Gaussian, are nearly identical. Indeed, by the Berry-Esseen Theorem [14, 28], the former distribution is nearly identical to a standard Gaussian and the latter distribution, being an appropriately weighted

sum of independent standard Gaussians, is a standard Gaussian itself. The invariance principle is now viewed as a generalization of this special case to *low degree* multi-linear polynomials, with the definition of influences as in Equation (2).

The invariance principle allows us to translate the noise stability problem on boolean hypercube to a similar problem in the Gaussian space and the latter problem has already been solved by Borell [15]! Towards this end, let f be a boolean function on n-dimensional hypercube that is balanced and has all influences small enough. We intend to upper bound its ε-noise stability. Consider the representation of f as a multi-linear polynomial:

$$f(x) = \sum_S \widehat{f}(S) \prod_{i \in S} x_i \qquad \forall x \in \{-1,1\}^n.$$

Let $f^* : \mathbb{R}^n \mapsto \mathbb{R}$ be a function that has the same representation as a multi-linear polynomial as f (with underlying standard Gaussian measure on $\mathbb{R}^n$):

$$f^*(x^*) = \sum_S \widehat{f}(S) \prod_{i \in S} x_i^* \qquad \forall x^* \in \mathbb{R}^n. \tag{3}$$

Assume for the moment that f has *low* degree. By the invariance principle, the distributions of $f(x)$ and $f^*(x^*)$ are nearly identical, and let's assume them to be identical for the sake of convenience. This implies that $\mathbb{E}[f^*] = \mathbb{E}[f] = 0$ and since f is boolean, so is f^*. In other words, f^* is a partition of $\mathbb{R}^n$ (with Gaussian measure) into two sets of equal measure. The next observation is that the ε-noise stability of f is same as the ε-"Gaussian noise stability" of $f^* : \mathbb{R}^n \mapsto \{-1,1\}$, defined as

$$\mathsf{Stab}_\varepsilon(f^*) := \Pr_{x^*, y^* \sim \mathsf{N}_\varepsilon(x^*)} \left[f^*(x^*) = f^*(y^*) \right]. \tag{4}$$

In the definition above, x^* is chosen from the standard n-dimensional Gaussian distribution and then y^* is chosen from the distribution $\mathsf{N}_\varepsilon(x^*)$, namely the perturbed or noisy version of x^*. Formally, $y^* = (1-2\varepsilon)x^* + \sqrt{1-(1-2\varepsilon)^2}z^*$ where z^* is a standard n-dimensional Gaussian independent of x^*. When f^* is a multi-linear polynomial as in Equation (3), it is easily proved that

$$\mathsf{Stab}_\varepsilon(f^*) = \frac{1}{2} + \frac{1}{2} \sum_S \widehat{f}(S)^2 (1-2\varepsilon)^{|S|}.$$

But this expression is same as the ε-noise stability of the boolean function f and thus $\mathsf{Stab}_\varepsilon(f) = \mathsf{Stab}_\varepsilon(f^*)$. It is important here that the co-ordinate-wise correlation between the boolean pair (x, y) is same as the co-ordinate-wise correlation between the Gaussian pair (x^*, y^*) in Equations (1), (4) defining the boolean and Gaussian noise stability respectively (both correlations equal $1-2\varepsilon$). Theorem 4.1 now follows from Borell's result that upper bounds $\mathsf{Stab}_\varepsilon(f^*)$.

Theorem 4.2. *If $g^* : \mathbb{R}^n \mapsto \{-1,1\}$ is a measurable function with $\mathbb{E}[g^*] = 0$, then*

$$\mathsf{Stab}_\varepsilon(g^*) \leq \mathsf{Stab}_\varepsilon(\textsf{HALF SPACE}) = 1 - \frac{1}{\pi} \arccos(1-2\varepsilon),$$

where HALF-SPACE *is the partition of $\mathbb{R}^n$ by a hyperplane through origin.*

We note that the error parameter δ in the statement of Theorem 4.1 accounts for additive errors involved at multiple places during the argument: firstly, the distributions $f(x)$ and $f^*(x^*)$ are only nearly identical. Secondly, f is not necessarily of low degree, and the invariance principle is not directly applicable. One gets around this issue by *smoothening* f that *kills* the high degree Fourier coefficients (which are then discarded) and only slightly affects the noise stability. This *truncated* version of f then has low degree and the invariance principle can be applied. We also note that the statement of Borell's Theorem holds for g^* that takes values in the interval $[-1, 1]$ and the noise stability is defined as in Equation (2).

To summarize, an iso-perimetric (type) result in the Gaussian space (e.g. Borell's Theorem) implies a Fourier analytic result on the hypercube (e.g. Majority Is Stablest), which in turn implies correctness of a reduction from the Unique Game problem to an optimization problem Π of interest (e.g. Max-Cut), showing that Π is hard to approximate. It turns out that this scheme applies to several optimization problems Π and not just for Max-Cut. In fact, for a class of problems known as constraint satisfaction problems, Max-Cut being one example, the three components, namely an iso-perimetric type result, a Fourier analytic result and a UGC-based hardness of approximation result, are formally equivalent [58]. The scheme also leads to new iso-perimetric type and Fourier analytic theorems and conjectures, motivated by applications to hardness of approximation (see [39] for examples).

5. Counter-examples to Proposed Algorithms

An interesting aspect of the Unique Games Conjecture is that it predicts the existence of *counter-examples* to proposed algorithms and answering whether such counter-examples indeed exist often turns out to be a challenging task with connections to geometry. We briefly explain this scheme and cite one example that leads to non-embeddability results for finite metrics.

Suppose there is a reduction from the Unique Game problem to a computational problem Π (similar to the reduction to the Max-Cut problem described earlier). Thus assuming the Unique Games Conjecture, the problem Π is hard to approximate. Nevertheless, one is free to propose an efficient algorithm $\mathcal{A}$ towards approximating Π and even a family of efficient algorithms $\{\mathcal{A}_i\}_{i=1,2,\ldots}$ that are increasingly more sophisticated. The Unique Games Conjecture predicts that Π is hard to approximate, and hence each of these proposed algorithms must fail. In particular, there must be a family of counter-examples (i.e. instances of the problem) $\{\mathcal{C}_i\}_{i=1,2,\ldots}$ demonstrating the failure of the corresponding algorithms. Moreover, the more sophisticated the proposed algorithms are, the more sophisticated the counter-examples would need to be. To emphasize, the counter-examples are concrete instances of the problem (e.g. graphs when the Max-Cut problem is considered) with a specific combinatorial or geometric structure.

When this scheme is applied to a problem called Sparsest Cut that is closely related to the Max-Cut problem, the Unique Games Conjecture predicts that there

are n-point finite metrics with non-trivial structural properties. In the Sparsest Cut problem, given a graph, the goal is to cut the graph into two roughly equal sized parts so as to minimize the fraction of edges cut.[7] There is a reduction from the Unique Game problem to the Sparsest Cut problem [20, 43, 66], so one predicts that the latter problem is hard to approximate.

Nevertheless, since mid-90s, researchers have proposed a family of increasingly sophisticated algorithms based on linear and semi-definite programming relaxation [49, 8, 51, 50, 6, 4]. These algorithms *relax* the Sparsest Cut problem to computing a metric on the set of vertices of the given graph that is *well-spread* and minimizes the average distance along the edges of the graph. It is possible to impose increasingly stringent restrictions on the type of metric allowed, leading to increasingly sophisticated algorithms. Cuts in an n-vertex graph are closely related to n-point ℓ_1 metrics and the approximation quality of the algorithm depends on how *close* the metric happens to be an ℓ_1 metric.

However, the Unique Games Conjecture predicts that all these algorithms must fail and hence corresponding counter-examples exist.[8] For some of these algorithms, researchers have already been able to construct such counter-examples (technically known as *integrality gap examples*), which amount to construction of n-point metrics with increasingly stringent structural properties. Before we state the known results, we introduce a notion of metric embedding.

A metric (X, d_X) consists of a set of points X and a distance function $d_X(\cdot,\cdot)$ on pairs of points that is non-negative, symmetric and satisfies the triangle inequality. An embedding of a metric space (X, d_X) into another metric space (Y, d_Y) is a map $\phi : X \mapsto Y$. The embedding is said to have distortion D if distances do not shrink and are not stretched by more than a factor D, i.e.

$$\forall a, b \in X, \quad d_X(a,b) \leq d_Y(\phi(a), \phi(b)) \leq D \cdot d_X(a,b).$$

An embedding with distortion $D = 1$ is said to be an isometric embedding. It is easily observed that if (X, d_X) is a metric, then so is $(X, \sqrt{d_X})$, i.e. when the new distances are square root of the original distances. A metric (X, d_X) is said to be of *negative type* if the metric $(X, \sqrt{d_X})$ embeds isometrically into ℓ_2. A sub-metric of a metric (X, d_X) is a subset $S \subseteq X$ with the same distances between points in S. We are now ready to state the result predicted by the Unique Games Conjecture and verified by researchers with explicit constructions (some of which precede the prediction).

Theorem 5.1. *There are functions $D(n), t(n) \to \infty$ as $n \to \infty$ and a family of n-point metrics (X, d_X) such that*

- *There is no embedding of (X, d_X) into ℓ_1 with distortion $D(n)$ [16, 8, 51].*
- *The metric (X, d_X) is of negative type [43, 45, 26, 48, 21, 22, 23, 24].*

[7] There are some subtleties regarding the so-called *uniform* and *non-uniform* versions of the problem that we omit here.

[8] Here we mean failure to approximate up to a constant multiplicative factor. If the approximation factor is allowed to depend on the size of the graph, the papers cited do indeed give a reasonable approximation.

- *Every sub-metric of* (X, d_X) *on* $t(n)$ *points embeds isometrically into* ℓ_1 *[59, 42].*

To state the theorem succinctly, there are negative-type metrics that embed isometrically into ℓ_1 locally, but do not embed well into ℓ_1 globally. The results cited hold for various quantitative settings of the parameters $D(n), t(n)$, but we omit these here and refer to [39, 55]. From the algorithmic side, it is possible to impose even more stringent restrictions on the metric (e.g. via the so-called Lasserre SDP relaxation), but then the existence of metrics with these restrictions (on top of those in Theorem 5.1) is open.

6. Open Problem: Power of Sum-of-Squares Refutation System

In this section and the next, we present two open problems towards settling the Unique Games Conjecture. The first one concerns the power of *refutation systems.* Suppose we have a correct, efficient algorithm for a computational problem Π (computing either exact or approximate solution). Suppose moreover that on some instance $\mathcal{I}$ of the problem, the algorithm does not find a solution. Since the algorithm is correct, the fact that it does not find a solution, is a proof that no solution exists, and often, a formal proof of infeasibility of a solution can be obtained by examining the execution of the algorithm on the instance $\mathcal{I}$. A proof of infeasibility of a solution is referred to as a *refutation.* More specifically, a refutation starts with a false hypothesis that a solution exists and then reaches a contradiction via a sequence of deductions. Naturally, for a refutation derived from the execution of an algorithm, the complexity of the refutation is related to the complexity of the algorithm. Turning this argument around, if on some instance $\mathcal{I}$ of the problem, if there is no *simple* refutation, this may be considered as evidence that the problem Π has no *simple* or *efficient* algorithm.

This motivates the study of *refutation systems* where a refutation conforms to a given set of rules for deducing successive statements, starting with a hypothesis to be refuted, e.g. a false hypothesis stating that a feasible solution exists when one doesn't. Depending on the kind of deduction rules allowed, one gets different refutation systems and their study is the subject of *proof complexity* (see [12, 11] for surveys). Here we focus on the Lovász-Schrijver, Sherali-Adams and the Lasserre systems. In these systems, each step of the refutation is an inequality and the system specifies how to derive new inequalities from the previous ones. There is a dual, algorithmic view of these systems and from that viewpoint, these systems correspond to LP/SDP relaxations (known as LP/SDP *hierarchies*) that we mentioned before. We refer to [65] for an introduction to and comparison between these systems (hierarchies).

As we said, we wish to show lower bounds for refutation systems, i.e. construct (infeasible) instances $\mathcal{I}$ such that there is no *simple* refutation within a given system. Showing such lower bounds then corresponds to constructing counter-

examples (i.e. *integrality gaps*) for the corresponding LP/SDP relaxation in the dual viewpoint, as discussed in Section 5.

Regarding the Max-Cut problem, reasonable lower bounds are known for the Lovász-Schrijver and Sherali-Adams systems (which are LP based) and also for some basic SDP-based systems [37, 32, 43, 25, 18, 59, 42]. However, showing lower bounds for the Lasserre system (which is SDP based) remains a major challenge and this is our first open problem. The Lasserre system is also known as Sum-of-Squares system and its variants have been studied independently by various authors including Shor, Parrilo, Nesterov, and Lasserre [63, 57, 56, 46]. It is closely related to the Hilbert's 17^{th} problem and we refer to [10, 47] for detailed expositions. Here we present the open problem in a self-contained manner.

Let us fix a graph $G(V = \{1, 2, \ldots, n\}, E)$ such that the maximum sized cut in the graph cuts exactly $(1-\varepsilon)|E|$ edges. We can write down an infeasible set of polynomial equalities and inequalities over reals, denoted $\mathcal{S}$, as follows:

$$\mathcal{S}: \quad \forall i \in \{1, \ldots, n\}, \quad x_i^2 - 1 = 0 \qquad (P_i(x) = 0).$$

$$\sum_{(i,j)\in E} \tfrac{1-x_i x_j}{2} - (1-\varepsilon)|E| - 1 \geq 0 \qquad (Q(x) \geq 0).$$

The set of equations is written as $P_i(x) = 0$ and the inequality is written as $Q(x) \geq 0$ where P_i, Q are polynomials in $\mathbb{R}[x_1, \ldots, x_n]$ as shown. Let's first see why this set of (in)equalities is infeasible. The equations $x_i^2 - 1 = 0$ force the variables x_i to take values in $\{-1, 1\}$. Any $\{-1, 1\}$-assignment to the variables is viewed as a cut in the graph and then the inequality $Q \geq 0$ states that the cut cuts at least $(1-\varepsilon)|E|+1$ edges, contradicting the assumption that the maximum sized cut cuts only $(1-\varepsilon)|E|$ edges. Indeed the expression $\frac{1-x_i x_j}{2}$ equals 1 or 0 depending on whether the edge (i, j) is cut or not and hence the sum $\sum_{(i,j)\in E} \frac{1-x_i x_j}{2}$ equals the number of edges cut.

How could one refute this infeasible set of (in)equalities? One possible way is to come up with polynomials $\{R_i\}_{i=1}^n, \{S_j, T_j\}_{j=1}^\ell \in \mathbb{R}[x_1, \ldots, x_n]$ such that the following polynomial identity holds:

$$\sum_{i=1}^{n} R_i P_i + (S_1^2 + \ldots + S_\ell^2)Q + (T_1^2 + \ldots + T_\ell^2) = -1.$$

This would be a contradiction, hence providing a valid refutation. Indeed, since $P_i = 0$ and $Q \geq 0$ and the polynomials S_j, T_j appear only in squared form, the left hand side of the identity is non-negative whereas the right hand side is -1. The refutation is called a Sum-of-Squares refutation.

It turns out that a Sum-of-Squares refutation always exists and the question is whether there is one that is *simple*. A natural measure of its complexity is the maximum degree of the polynomials $R_i P_i, S_j^2 Q, T_j^2$ involved, called the degree of the refutation. It is known that a degree d refutation, if one exists, can be found in time $n^{O(d)}$, i.e. in polynomial time for constant d. Thus it is desirable to have a refutation with constant degree (independent of the size of the graph). From a

lower bound perspective, it is known that there are n-vertex graphs for which any Sum-of-Square refutation requires degree $\Omega(n)$ (degree $O(n)$ always suffices).

What if we insist on having a constant degree refutation? One possibility is to start with a hypothesis that is *even more false*. In particular, one can consider the set of (in)equalities:

$$\mathcal{S}' : \quad \forall i \in \{1, \ldots, n\}, \quad x_i^2 - 1 = 0.$$

$$\textstyle\sum_{(i,j)\in E} \frac{1-x_i x_j}{2} - (1-\varepsilon^2)|E| \geq 0.$$

Note that the inequality hypothesizes that there is a cut that cuts at least $(1-\varepsilon^2)|E|$ edges. This hypothesis is much more false than the earlier hypothesis stating that there is a cut that cuts at least $(1-\varepsilon)|E|+1$ edges and thus is plausibly easier to refute. Indeed, for any graph (with maximum cut of size $(1-\varepsilon)|E|$), the set of (in)equalities $\mathcal{S}'$ has a Sum-of-Squares refutation of degree 2! Such a refutation can be obtained by taking a dual view of the Goemans-Williamson's SDP algorithm for the Max-Cut problem.

These considerations lead to our first open problem: what happens when we use a hypothesis stating that there is a cut that cuts a number of edges that is intermediate between $(1-\varepsilon)|E|+1$ and $(1-\varepsilon^2)|E|$? Is there always a constant degree refutation (noting that one needs degree $\Omega(n)$ for some graphs at first extreme and degree 2 always suffices at the second extreme)? Specifically, Let $\tilde{\varepsilon}$ be any constant such that $\varepsilon^2 \ll \tilde{\varepsilon} \ll \varepsilon$. The Unique Games Conjecture predicts, as discussed in Section 3, that no polynomial time algorithm, given a graph with maximum cut of size $1-\varepsilon$, finds a cut of size $1-\frac{1}{2}\sqrt{\varepsilon}$. This prediction, when translated to a prediction regarding lower bounds for the Sum-of-Squares refutation system, states:

Prediction: Let $\varepsilon^2 \ll \tilde{\varepsilon} \ll \varepsilon$. There are graphs $G(V = \{1, 2, \ldots, n\}, E)$ with the maximum cut of size exactly $(1-\varepsilon)|E|$ such that any Sum-of-Squares refutation of the set of (in)equalities:

$$\tilde{\mathcal{S}} : \quad \forall i \in \{1, \ldots, n\}, \quad x_i^2 - 1 = 0.$$

$$\textstyle\sum_{(i,j)\in E} \frac{1-x_i x_j}{2} - (1-\tilde{\varepsilon})|E| \geq 0.$$

requires a super-constant degree (i.e. tending to ∞ as $n \to \infty$).

Clearly, constructing graphs that require a super-constant degree refutation for some $\varepsilon^2 \ll \tilde{\varepsilon} \ll \varepsilon$ would support the Unique Games Conjecture whereas showing that there is always a constant degree refutation for some $\varepsilon^2 \ll \tilde{\varepsilon} \ll \varepsilon$ would disprove the Unique Games Conjecture.

7. Open Problem: Small Set Expander Graphs with Many Large Eigenvalues

The second open problem concerns the existence of *small set expander* graphs with many large eigenvalues. The problem is motivated by the Small Set Expansion Conjecture posed by Raghavendra and Steurer [60]. The conjecture concerns the computational complexity of the *small set expansion* problem which, given a graph, asks for a small (but still of linear size) subset of vertices that does not expand well. The conjecture states that this problem is hard to approximate; a formal statement appears below.

For a d-regular graph $G(V, E)$ and a set $S \subseteq V$, define the expansion of the set S as $\phi(S) := \frac{|E(S,V\setminus S)|}{d\cdot|S|}$, i.e. the fraction of edges incident on S that leave S. Raghavendra and Steurer pose:[9]

Conjecture 7.1. *For every constant $\varepsilon > 0$, there exists a constant $\gamma > 0$ such that no polynomial time algorithm, given a regular graph $G(V, E)$, can distinguish whether it is a YES Type graph or a NO Type graph as defined below:*

- *(YES Type:) There is a set $S \subseteq V, |S| = \gamma|V|$ such that $\phi(S) \leq \varepsilon$.*
- *(NO Type:) For every set $S \subseteq V, |S| \approx \gamma|V|$, $\phi(S) \geq \frac{1}{10}$.*

As a clarification, we note that a *distinguishing* algorithm takes a graph as input and in polynomial time outputs an answer that is YES if the graph is of YES Type and is NO if the graph is of NO Type. For graphs that are of neither type, the output of the algorithm can be arbitrary. Though the conjecture is phrased as above (as is customary in computer science), the reader may find it more convenient to consider the following version implied by it (computer scientists tend to view the two versions as morally the same):

Conjecture: There is no polynomial time algorithm that, given a graph of the YES Type, meaning one containing a small set (i.e. of relative size $\approx \gamma$) that is almost non-expanding (i.e. has expansion at most ε), finds a small set that is somewhat non-expanding (i.e. has expansion less than $\frac{1}{10}$).

Finding small non-expanding sets is a natural problem in itself and in addition, Raghavendra and Steurer show that this conjecture implies the Unique Games Conjecture. Therefore, it is worthwhile to explore this conjecture. As discussed in Section 5, for a computational problem that is predicted to be hard to approximate, the small set expansion problem in this case, one can propose an efficient algorithm and then try to find counter-examples to the proposed algorithm.

It is indeed possible to propose a natural algorithm to find small non-expanding sets [44, 3]. We briefly sketch the algorithm. Let $A(G)$ be the normalized adjacency matrix of a d-regular n-vertex graph $G(V, E)$. This is a $n \times n$ matrix with diagonal entries as 1 and an off-diagonal entry (i, j) is $\frac{1}{d}$ if (i, j) is an edge in the graph and zero otherwise. It is well-known that the eigenvalues of this matrix are in $[-1, 1]$

[9] Here $|S| \approx \gamma|V|$ means that, say, $|S|$ is between $\frac{\gamma}{2}|V|$ and $2\gamma|V|$.

and the largest eigenvalue equals 1. Let $v_1, \ldots, v_m \in \mathbb{R}^n$ be the *top eigenvectors*, i.e. those corresponding to eigenvalues that are at least $1 - O(\varepsilon)$. For a subset of vertices $S \subseteq V$, let $\mathbf{1}_S \in \mathbb{R}^n$ denote the indicator vector of the subset S, i.e. its i^{th} co-ordinate equals 1 if the i^{th} vertex is in S and zero otherwise. It is easily shown (e.g. [3, Theorem 2.2]) that if the graph has a subset $S \subseteq V, |S| = \gamma n$ such that $\phi(S) \leq \varepsilon$, then the indicator vector $\mathbf{1}_S$ is essentially contained in the linear span of the top eigenvectors $v_1, \ldots, v_m$. Thus the vector $\mathbf{1}_S$ and hence the set S (or rather, an approximation to them) can be found by searching over all the vectors in this m-dimensional linear span (up to a suitable discretization) and outputting a vector that resembles an indicator vector of a set of size $\approx \gamma n$. Let's refer to this algorithm as a *subspace search* algorithm; it runs in time roughly $2^{O(m)}$.

Now consider a proposed algorithm to distinguish between the YES Type and NO Type graphs as in the statement of Conjecture 7.1. Compute the eigenvalues and eigenvectors of the matrix $A(G)$. If the number of large eigenvalues m is at most $n^{o(1)}$, proceed further and otherwise answer YES. If m is at most $n^{o(1)}$, run the subspace search algorithm and answer YES or NO depending on whether it manages to find a set of size $\approx \gamma n$ with expansion $\ll \frac{1}{10}$. Note that the proposed algorithm always answers YES on a graph of the YES Type.

However, Conjecture 7.1 predicts that every polynomial time algorithm fails in distinguishing between the YES Type and the NO Type graphs. In fact, Raghavendra and Steurer state Conjecture 7.1 in a stronger form, predicting that the task of distinguishing between the YES Type and the NO Type graphs is NP-complete, and every algorithm that runs in time $2^{n^{o(1)}}$ time fails as well. Considering the proposed algorithm as above, the only reason for it to fail is that it mistakenly answers YES on some graph that is of the NO Type. Thus we are led to the following prediction:

Prediction: For every constant $\varepsilon > 0$, there exist constants $\gamma, \delta > 0$ and an infinite family of n-vertex graphs $G(V, E)$ of the NO Type, i.e. $\forall S \subseteq V, |S| \approx \gamma n$, $\phi(S) \geq \frac{1}{10}$, such that the number of its eigenvalues $\geq 1 - \varepsilon$ is at least n^{δ}.

The open question is whether such graphs exist (see [9] for some progress). It is possible that such graphs do not exist and the Small Set Expansion Conjecture is false (and the Unique Games Conjecture might still be true).

8. Conclusion

We have sketched some connections between the Unique Games Conjecture, geometry and analysis. Irrespective of whether the Unique Games Conjecture turns out to be true or false, exploring these connections further, and in particular making progress on the open questions cited, seems worthwhile.

References

[1] S. Arora. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *Journal of the ACM*, 45 (5): 753-782, 1998.

[2] S. Arora, L. Babai, J. Stern, and Z. Sweedyk. The hardness of approximate optima in lattices, codes, and systems of linear equations. *Journal of Computer and System Sciences*, Volume 54, Issue 2, pages 317–331, 1997.

[3] S. Arora, B. Barak, and D. Steurer. Subexponential Algorithms for Unique Games and Related Problems. In *Proc. 51th IEEE Symposium on Foundations of Computer Science*, 2010.

[4] S. Arora, J. Lee, and A. Naor. Euclidean distortion and the sparsest cut. In *Proc. 37th ACM Symposium on Theory of Computing*, pages 553–562, 2005.

[5] S. Arora, C. Lund, R. Motawani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM*, 45(3):501–555, 1998.

[6] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proc. 36th ACM Symposium on Theory of Computing*, pages 222–231, 2004.

[7] S. Arora and S. Safra. Probabilistic checking of proofs : A new characterization of NP. *Journal of the ACM*, 45(1):70–122, 1998.

[8] Y. Aumann and Y. Rabani. An O(log k) approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.*, 27(1):291-301, 1998.

[9] B. Barak, P. Gopalan, J. Håstad, R. Meka, P. Raghavendra, and D. Steurer. Making the Long Code Shorter. In *Proc. IEEE Symposium on Foundations of Computer Science*, 2012, 370-379.

[10] B. Barak and D. Steurer. Sum-of-Squares Proofs and the Quest toward Optimal Algorithms. In *Proc. of the Internatioanl Congress of Mathematicians*, 2014.

[11] P. Beame. Proof complexity. *Computational Complexity Theory*, volume 10 of IAS/Park City mathematics series, pages 199-246. American Mathematical Society, 2004.

[12] P. Beame and T. Pitassi. Propositional proof complexity: Past, present, and future. *Current trends in theoretical computer science: Entering the 21st century*, World Scientific Publishing, 2001, pp. 42-70.

[13] M. Bellare, O. Goldreich, and M. Sudan. Free Bits, PCPs, and Nonapproximability—Towards Tight Results. *SIAM Journal on Computing*, Volume 27, Issue 3, pages 804–915, 1998.

[14] A. Berry. The Accuracy of the Gaussian Approximation to the Sum of Independent Variates. *Transactions of the American Mathematical Society*, 49 (1): 122-136, 1941.

[15] C. Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[16] J. Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52:46-52, 1985.

[17] M. Charikar, K. Makarychev, and Y. Makarychev. Near-Optimal Algorithms for Unique Games. In *Proc. Annual ACM Symposium on Theory of Computing*, 2006, pp. 205-214.

[18] M. Charikar, K. Makarychev, and Y. Makarychev. Integrality Gaps for Sherali-Adams Relaxations. In *Proc. ACM Symposium on Theory of Computing*, 2009, pp. 283-292.

[19] S. Chatterjee. A simple invariance theorem. *arXiv:math/0508213v1*, 2005.

[20] S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. In *Proc. 20th IEEE Conference on Computational Complexity*, pages 144–153, 2005.

[21] J. Cheeger and B. Kleiner. Differentiating maps into L^1 and the geometry of BV functions. *Ann. Math.*, Second Series, Vol. 171, No. 2, 2010.

[22] J. Cheeger and B. Kleiner. On the differentiation of Lipschitz maps from metric measure spaces to Banach spaces. *Inspired by S.S. Chern, Volume 11 of Nankai Tracts. Math.*, pages 129–152, 2006.

[23] J. Cheeger, B. Kleiner, and A. Naor. Compression bounds for Lipschitz maps from the Heisenberg group to L_1. *Acta Mathematica*, volume 207, issue 2 (2011), pages 291–373.

[24] J. Cheeger, B. Kleiner, and A. Naor. A $(\log n)^{\Omega(1)}$ integrality gap for the sparsest cut SDP. In *Proc. 50th IEEE Symposium on Foundations of Computer Science*, 2009.

[25] W.F. de la Vega and C. Kenyon-Mathieu. Linear Programming Relaxations of Max-cut. In *Proc. 18th Symposium on Discrete Algorithms*, pp. 53-61, 2007.

[26] N. Devanur, S. Khot, R. Saket, and N. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proc. 38th ACM Symposium on Theory of Computing*, 2006.

[27] I. Dinur and S. Safra. The importance of being biased. In *Proc. 34th Annual ACM Symposium on Theory of Computing*, 2002.

[28] C.G. Esseen. On the Liapunoff limit of error in the theory of probability. *Arkiv for matematik, astronomi och fysik*, A28: 1-19, 1942.

[29] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy. Interactive proofs and the hardness of approximating cliques. *Journal of the ACM*, 43(2):268–292, 1996.

[30] U. Feige. Threshold of ln n for Approximating Set Cover. *Journal of the ACM*, 45(4), 634–652, 1998.

[31] U. Feige and L. Lovász. Two-prover one-round proof systems, their power and their problems. In *Proc. of the ACM Symposium on the Theory of Computing*, pages 733-744, 2002.

[32] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for max cut. *Random Struct. Algorithms*, 20(3):403-440, 2002.

[33] M. Goemans and D. Williamson. 0.878 approximation algorithms for MAX-CUT and MAX-2SAT. In *Proc. 26th ACM Symposium on Theory of Computing*, pages 422–431, 1994.

[34] M. X. Goemans. Semidefinite programming in combinatorial optimization. *Math. Program.*, 79:143–161, 1997.

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

[36] J. Hastad. Some optimal inapproximability results. *Journal of ACM*, 48:798–859, 2001.

[37] H. Karloff. How good is the Goemans-Williamson MAX CUT algorithm? In *Proc. of the twenty-eighth annual ACM Symposium on Theory of Computing*, pages 427-434, 1996.

[38] S. Khot. On the power of unique 2-prover 1-round games. In *Proc. 34th ACM Symposium on Theory of Computing*, 2002.

[39] S. Khot. Inapproximability of NP-complete problems, Discrete Fourier Analysis, and Geometry. In *Proc. of the International Congress of Mathematicians*, 2010.

[40] S. Khot. On the Unique Games Conjecture. In *Proc. IEEE Conference on Computational Complexity*, 2010.

[41] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *Proc. 45th IEEE Symposium on Foundations of Computer Science*, pages 146–154, 2004.

[42] S. Khot and R. Saket. SDP integrality gaps with local ℓ_1-embeddability. In *Proc. 50th IEEE Symposium on Foundations of Computer Science*, pages 565–574, 2009.

[43] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proc. 46th IEEE Symposium on Foundations of Computer Science*, 2005.

[44] A. Kolla. Spectral Algorithms for Unique Games. In *Proc. IEEE Conference on Computational Complexity*, 2010.

[45] R. Krauthgamer and Y. Rabani. Improved lower bounds for embeddings into l_1. In *Proc. ACM-SIAM Symposium on Discrete Algorithms*, 2006.

[46] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11 (2001), no. 3, 796-817.

[47] M. Laurent. A Comparison of the Sherali-Adams, Lovasz-Schrijver, and Lasserre Relaxations for 0-1 Programming. *Math. Oper. Res.*, 28(3):470-496, 2003.

[48] J. R. Lee and A. Naor. l_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *Proc. 47th IEEE Symposium on Foundations of Computer Science*, pages 99–108, 2006.

[49] T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, Volume 46, Issue 6, pages 787–832, 1999.

[50] N. Linial. Finite metric spaces-combinatorics, geometry and algorithms. In *Proc. International Congress of Mathematicians*, volume 3, pages 573–586, 2002.

[51] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215-245, 1995.

[52] J.S.B. Mitchell. Guillotine subdivisions, approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, k-MST, and related problems. *SIAM Journal on Computing*, 28 (4): 12981309, 1999.

[53] E. Mossel. Gaussian Bounds for Noise Correlation of Functions. *Geometric and Functional Analysis*, March 2010, Volume 19, Issue 6, pp 1713-1756.

[54] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *Proc. 46th IEEE Symposium on Foundations of Computer Science*, pages 21–30, 2005.

[55] A. Naor. L_1 embeddings of the Heisenberg group and fast estimation of graph isoperimetry. In *Proc. the International Congress of Mathematicians*, 2010.

[56] Y. Nesterov. Squared functional systems and optimization problems. *High performance optimization*, 13 (2000), 405-440.

[57] P. Parrilo. Semidefinite programming relaxations for semialge-braic problems. *Mathematical Programming*, 96 (2003), 293-320.

[58] P. Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proc. ACM Symposium on Theory of Computing* , 2008: 245-254.

[59] P. Raghavendra and D. Steurer. Integrality gaps for strong SDP relaxations of unique games. In *Proc. 50th IEEE Symposium on Foundations of Computer Science*, pages 575–585, 2009.

[60] P. Raghavendra and D. Steurer. Graph expansion and the unique games conjecture. In *Proc. 42nd ACM Symposium on Theory of Computing*, 2010.

[61] R. Raz. A Parallel Repetition Theorem. *SIAM Journal of Computing*, 27(3), pp. 763-803, 1998.

[62] V. Rotar'. Limit theorems for polylinear forms. *J. Multivariate Anal.*, 9(4):511–530, 1979.

[63] N.Z. Shor. An approach to obtaining global extremums in polynomial mathematical programming problems. *Cybernetics and Systems Analysis*, 23 (1987), no. 5, 695-700.

[64] L. Trevisan. On Khot's Unique Games Conjecture. *Bulletin of the AMS*, 49(1): 91-111, 2012.

[65] E. Chlamtac and M. Tulsiani. Convex Relaxations and Integrality Gaps. *Springer Handbook on Semidefinite, Conic and Polynomial Optimization*, 2012.

[66] M. Tulsiani, P. Raghavendra, and D. Steurer. Reductions between Expansion Problems. In *Proc. IEEE Conference on Computational Complexity*, 2012.

251 Mercer Street, New York, NY-10012, USA.

E-mail: khot@cs.nyu.edu