

Unique Games and Small Set Expansion

The *Unique Games Conjecture (UGC)* (Khot [2002]) states that for every $\epsilon > 0$ there is some finite Σ such that the $1 - \epsilon$ vs. ϵ problem for the unique games problem with alphabet Σ is NP hard. An instance to this problem can be viewed as a set of equations on variables $x_1, \dots, x_n \in \Sigma$ where each equation has the form $x_i = \pi_{i,j}(x_j)$ where $\pi_{i,j}$ is a permutation over Σ .

The order of quantifiers here is crucial. Clearly there is a propagation type algorithm that can find a perfectly satisfying assignment if one exist, and as we mentioned before, it turns out that there is also an algorithm nearly satisfying assignment (e.g., satisfying 0.99 fraction of constraints) if ϵ tends to zero independently of Σ . However, when the order is changed, the simple propagation type algorithms fail. That is, if we guess an assignment to the variable x_1 , and use the permutation constraints to propagate this guess to an assignment to x_1 's neighbors and so on, then once we take more than $O(1/\epsilon)$ steps we are virtually guaranteed to make a mistake, in which case our propagated values are essentially meaningless. One way to view the unique games conjecture is as stating this situation is inherent, and there is no way for us to patch up the local pairwise correlations that the constraints induce into a global picture. While the jury is still out on whether the conjecture is true as originally stated, we will see that the sos algorithm allows us in some circumstances to do just that. The *small set expansion (SSE) problem* (Raghavendra and Steurer [2010]) turned out to be a clean abstraction that seems to capture the ‘‘combinatorial heart’’ of the unique games problem, and has been helpful in both designing algorithms for unique games as well as constructing candidate hard instances for it. We now describe the SSE problem and outline the known relations between it and the unique games problem.

1. Definition (Small Set Expansion problem). For every undirected d -regular graph $G = (V, E)$, and $\delta > 0$ the *expansion profile of G at δ* is

$$\varphi_\delta(G) = \min_{S \subseteq V, |S| \leq \delta|V|} |E(S, \bar{S})| / (d|S|). \quad (1)$$

For $0 < c < s < 1$, the c vs s δ -SSE problem is to distinguish, given a regular undirected graph G , between the case that $\varphi_\delta(G) \leq c$ and the case that $\varphi_\delta(G) \geq s$.

The *small set expansion hypothesis (SSEH)* is that for every $\epsilon > 0$ there is some δ such that the ϵ vs $1 - \epsilon$ δ -SSE problem is NP hard.

We will see that every instance of unique games on alphabet Σ

gives rise naturally to an instance of $\frac{1}{|\Sigma|}$ -SSE. But the value of the resulting instance is not always the same as the original one, and hence this natural mapping does not yield a valid reduction from the unique games to the SSE problem. Indeed, no such reduction is known. However, [Raghavendra and Steurer \[2010\]](#) gave a highly non trivial reduction in the other direction (i.e., reducing SSE to UG) hence showing that the SSEH *implies* the UGC. Hence, assuming the SSEH yields all the hardness results that are known from the UGC, and there are also some other problems such as *sparsest cut* where we only know how to establish hardness based on the SSEH and not the UGC ([Raghavendra et al. \[2012\]](#)).

Embedding unique games into small set expansion

Given a unique games instance I on n variables and alphabet Σ , we can construct a graph $G(I)$ (known as the *label extended graph* corresponding to I) on $n|\Sigma|$ vertices where for every equation of the form $x_i = \pi_{i,j}(x_j)$ we connect the vertex (i, σ) with (j, τ) if $\pi_{i,j}(\tau) = \sigma$. Note that every assignment $x \in \Sigma^n$ is naturally mapped into a subset S of size n (i.e. $1/|\Sigma|$ fraction) of the vertices of $G(I)$ corresponding to the vertices $\{(i, x_i) : i \in [n]\}$. Let's assume for simplicity that in the unique games instance every variable participated in exactly d of the constraints (and hence there is a total of $nd/2$ constraints). Then the graph G will have degree d and every constraint $x_i \neq \pi_{i,j}(x_j)$ that is violated by x contributes exactly two edges to $E(S, \bar{S})$ (the edges $\{(i, x_i), (j, \pi_{i,j}^{-1}(x_i))\}$ and $\{(i, \pi_{i,j}(x_j)), (j, x_j)\}$). Hence the fraction of violated constraints by x is equal to $E(S, \bar{S}) / (d|S|)$.

However, this is not a valid reduction, since we cannot do this mapping in reverse. That is, the graph G might have a set S of n vertices that does not expand that does not correspond to any assignment for the unique games. It can be shown that this reduction is valid if the constraint graph of the unique games instance is a *small set expander* in the sense that every set of $o(1)$ fraction of the vertices has expansion $1 - o(1)$.

The known algorithmic results for small set expansion are identical to the ones known for unique games:

- The basic SDP algorithm can solve the $1 - \epsilon$ vs $1 - O(\sqrt{\epsilon \log(1/\delta)})$ δ -SSE problem ([Raghavendra et al. \[2010\]](#)) which matches the $1 - \epsilon$ vs $1 - O(\sqrt{\epsilon \log |\Sigma|})$ algorithm for the Σ -UG problem of [Charikar et al. \[2006\]](#) (which itself generalizes the $1 - \epsilon$ vs $1 - O(\sqrt{\epsilon})$ [Goemans and Williamson \[1995\]](#) for Max Cut).

- When $\delta \ll \epsilon$ (specially $\delta < \exp(\Omega(1/\epsilon))$) then no polynomial time algorithm is known. However, Arora et al. [2015] showed a $\exp(\text{poly}(1/\delta)n^{\text{poly}(1/\epsilon)})$ time algorithm for the $1 - \epsilon$ vs $1/2$ δ -SSE problem and similarly a $\exp(\text{poly}(|\Sigma|)n^{\text{poly}(1/\epsilon)})$ time algorithm for the $1 - \epsilon$ vs $1/2$ Σ -UG problem.

Thus it is consistent with current knowledge that the $1 - \epsilon$ vs ϵ Σ -UG computational problem has complexity equivalent (up to polynomial factors) to the $1 - \epsilon$ vs $\epsilon \frac{1}{|\Sigma|}$ -SSE problem. At the moment one can think of the small set expansion problem as the “easiest” “UG like” problem and the Max-Cut (or Max $2LIN(2)$) problem as the “hardest” “UG like” problem where by “UG like” we mean problems that in the perfect case can be solved by propagation like algorithms and that we suspect might have a subexponential time algorithm.

(Such algorithms are known for unique games and small set expansion but not for Max Cut.) Unique Games itself can be thought of as lying somewhere between these two problems. The following figure represents are current (very partial) knowledge of the landscape of complexity for CSPs and related graph problems such as small set expansion and sparsest cut

Small set expansion as finding “analytically sparse” vectors.

We have explained that the defining feature of “unique games like” problems is the fact that they exhibit *strong pairwise correlations*. One way to see it is that these problems have the form of trying to optimize a *quadratic polynomial* $x^\top Ax$ over some “structured” set x :

- The *maximum cut problem* can be thought as the task of maximizing $x^\top Lx$ over $x \in \{\pm 1\}^n$ where L is the *Laplacian matrix* of the graph.
- The *sparsest cut problem* can be thought of as the task of *minimizing* the same polynomial $x^\top Lx$ over unit vectors (in ℓ_2 norm) of the form $x \in \{0, \alpha\}^n$ for some $\alpha > 0$.
- The *small set expansion problem* can be thought of as the task of minimizing the same polynomial $x^\top Lx$ over unit vectors of the form $x \in \{0, \alpha\}^n$ subject to the condition that $\|x\|_1 / \sqrt{n} \leq \sqrt{\delta}$ (which corresponds to the condition that $\alpha \geq 1/\sqrt{n\delta}$).
- The *unique games problem* can be thought of as the task of minimizing the polynomial $x^\top Lx$ where L is the Laplacian matrix corresponding to the *label extended graph* of the game and x ranges

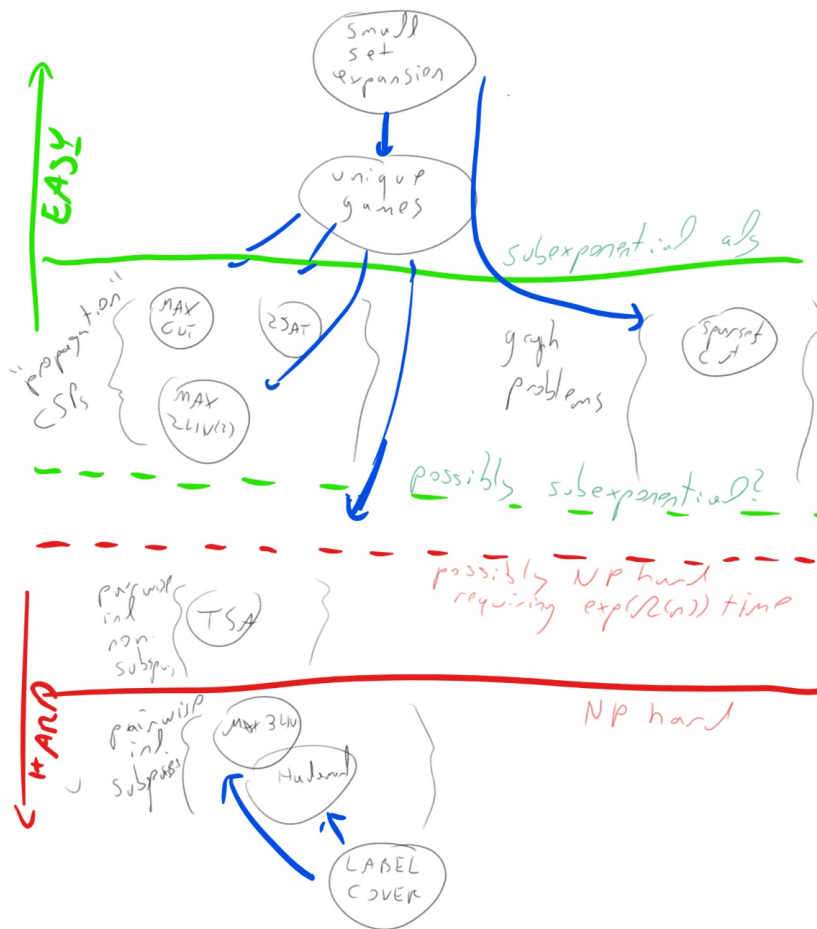


Figure 1: The “landscape” of the computational difficulty of computational problems. If the Unique Games conjecture is false, all the problems above the dotted green line might be solvable by sos of degree $\tilde{O}(1)$.

over all vectors in $\{0, 1\}^{n^{|\Sigma|}}$ such that for every $i \in [n]$ there exists exactly one $\sigma \in \Sigma$ such that $x_{i,\sigma} = 1$.

In all of these cases the computational question we are faced with amounts to what kind of *global information* can we extract about the unknown assignment x from the the *local pairwise information* we get by knowing the expected value of $(x_i - x_j)^2$ over an average edge $i \sim j$ in the graph.

Proxies for sparsity

We see that all these problems amount to optimizing some quadratic polynomial $q(x)$ over unit vector x 's that are in some *structured* set. By “guessing” the optimum value λ that $q(x)$ can achieve, we can move to a “dual” version of these problems where we want to maximize $p(x)$ over all $x \in \mathbb{R}^n$ such that $q(x) = \lambda$ where $p(x)$ is some “reward function” that is larger the closer that x is to our structured set (equivalently, one can look at $-p(x)$ as a function that *penalizes* x 's that are far from our structured set).

In fact, in our setting this λ will often be (close to) the minimum or maximum of $q(\cdot)$ which means that when x achieves $q(x) = \lambda$ then all the directional derivatives of the quadratic q (which are linear functions) vanish. Another way to say it is that the condition $q(x) = \lambda$ will correspond to x being (almost) inside some linear subspace $V \subseteq \mathbb{R}^n$ (which will be the eigenspace of q , looked at as a matrix, corresponding to the value λ and other nearby values). This means that we can reduce our problem to the task of approximating the value

$$\max_{x \in V, \|x\|=1} p(x) \quad (2)$$

where $p(\cdot)$ is a problem-specific reward function and V is a linear subspace of \mathbb{R}^n that is derived from the input. This formulation does not necessarily make the problem easier as $p(\cdot)$ can (and will) be a non convex function, but, when p is a low degree polynomial, it does make it potentially easier to analyze the performance of the sos algorithm. That being said, in none of these cases so far we have found the “right” reward function $p(\cdot)$ that would yield a tight analysis for the sos algorithm. But for the small set expansion question, there is a particularly natural and simple reward function that shows promise for better analysis, and this is what we will focus on next.

Choosing such a reward function might seem to go against the “sos philosophy” where we are not supposed to use creativity in

designing the algorithm, but rather formulate a problem such as small set expansion as an sos program in the most straightforward way, without having to guess a clever reward function. We believe that in the cases we discussed, it is possible to consign the choice of the reward function to the *analysis* of the standard sos program. That is, if one is able to show that the sos relaxation for Eq. (2) achieves certain guarantees for the original small set expansion problem, then it should be possible to show that the standard sos program for small set expansion would achieve the same guarantees.

Norm ratios as proxies for sparsity

In the small set expansion we are trying to minimize the quadratic $x^\top Lx$ subject to $x \in \{0, \alpha\}$ being *sparse*. It turns out that the *sparsity* condition is more significant than the Booleanity. For example, we can show the following result:

2. Theorem (Hard local Cheeger). *Suppose that $x \in \mathbb{R}^n$ has at most δn nonzero coordinates, and satisfies $x^\top Lx \leq \epsilon \|x\|^2$ where L is the Laplacian of an n -vertex d -regular graph G . Then there exists a set S of at most δn of G 's vertices such that $|E(S, \bar{S})| \leq O(\sqrt{\epsilon} \delta n)$.*

It turns out that this theorem can be proven in much the same way that Cheeger itself is proven, showing that a random level set of x (chosen such that it will have the support be a subset of the nonzero coordinates of x) will satisfy the needed condition. Note that if x has at most δn nonzero coordinates then $\|x\|_1 \leq \sqrt{n\delta} \|x\|_2$. It turns out this condition is sufficient for the above theorem:

3. Theorem (Local Cheeger). *Suppose that $x \in \mathbb{R}^n$ satisfies $\|x\|_1 \leq \sqrt{\delta n} \|x\|_2$, and satisfies $x^\top Lx \leq \epsilon \|x\|^2$ where L is the Laplacian of an n -vertex d -regular graph G . Then there exists a set S of at most $O(\sqrt{\delta n})$ of G 's vertices such that $|E(S, \bar{S})| \leq O(\sqrt{\epsilon} \delta n)$.*

Proof. If x satisfies the theorem's assumptions, then we can write $x = x' + x''$ where for every i , if $|x_i| > \|x\|_2 / (2\sqrt{n})$ then $x'_i = x_i$ and $x''_i = 0$ and else $x''_i = x_i$ and $x'_i = 0$. By this choice, $\|x''\|_2^2 \leq \text{norm}x^2 / 2$ and so (since x' and x'' are orthogonal) $\|x'\|^2 = \|x\|^2 - \|x''\|^2 \geq \|x\|^2 / 2$. By direct inspection, zeroing out coordinates will not increase the value of $x^\top Lx$ and so

$$x'^\top Lx' \leq x^\top Lx \leq \epsilon \|x\|^2 \leq 2\epsilon \|x'\|^2. \quad (3)$$

But the number of nonzero coordinates in x' is at most $2\sqrt{\delta n}$.

Indeed, if there were more than $2\sqrt{\delta n}$ coordinates in which $|x_i| \geq \text{norm}x_2 / (2\sqrt{n})$ then $\|x\|_1 \geq 2\sqrt{\delta n}\|x\|_2 / (2\sqrt{n}) = 2\sqrt{\delta n}\|x\|_2$. \square

The above suggests the following can serve as a smooth proxy for the small set expansion problem:

$$\max \|x\|_2^2 \quad (4)$$

over $x \in \mathbb{R}^n$ such that $\|x\|_1 = 1$ and $x^\top Lx \leq \epsilon \|x\|^2$. Indeed the ℓ_1 to ℓ_2 norm ratio is an excellent proxy for sparsity, but unfortunately one where we don't know how to prove algorithmic results for.

As we say in the sparsest vector problem, the ℓ_2 to ℓ_4 ratio (and more generally the ℓ_2 to ℓ_p ratio for $p > 2$) is a quantity we find easier to analyze, and indeed it turns out we can prove the following theorem:

4. Theorem. *Let G be a d regular and V_λ be the subspace of vectors corresponding to eigenvalues smaller than λ in the Laplacian.*

1. *If $\max_{\|x\|_2=1, x \in V_\lambda} \|x\|_p \leq Cn^{1/p-1/2}$ then for every subset S of at most δn vertices, $|E(S, \bar{S})| \geq \lambda(1 - C^2\delta^{1-2/p})$*
2. *If there is a vector $x \in V_\lambda$ with $\|x\|_p > C\|x\|_2$ then there exists a set S of at most $O(n/\sqrt{C})$ vertices such that $|E(S, \bar{S})| \leq 1 - (1 - \lambda)^{2p}2^{-O(p)}$ where the $O(\cdot)$ factor hide absolute factors independent of C .*

Statement 1. (a bound on the ℓ_p/ℓ_2 ratio implies expansion) has been somewhat of ‘‘folklore’’ (e.g., see Chapter 10 in [O'Donnell \[2014\]](#)). Statement 2. is due to ? .

The proof for 1. is not hard and uses Hölder's inequality in a straightforward matter. In particular (and this has been found useful) this proof can be embedded in the degree $O(p)$ sos program.

The proof for 2. is more involved. One way to get intuition for this is that if $x^\top Lx$ is small then one could perhaps hope that $y^\top Ly$ is also not too large where $y \in \mathbb{R}^n$ is defined as $(x_1^{p/2}, \dots, x_n^{p/2})$. That is, $x^\top Lx \leq \epsilon \|x\|^2$ means that on average over a random edge $\{i, j\}$, $(x_i - x_j)^2 \leq O(\epsilon(x_i^2 + x_j^2))$. So one can perhaps hope that for a ‘‘typical’’ edge i, j , $x_i \in (1 \pm O(\epsilon))x_j$ but if that holds then $y_i \in (1 \pm O(p\epsilon))y_j$ since $y_i = x_i^{p/2}$ and $y_j = x_j^{p/2}$. But if $\mathbb{E} x_i^p > C(\mathbb{E} x_i^2)^{p/2}$ we can show using Hölder that $\mathbb{E} x_i^p > C'(\mathbb{E} x_i^{p/2})^2$ for some C' that goes to infinity with C . Hence this means that y is sparse in the ℓ_2 vs ℓ_1 sense, and we can apply [Theorem 3](#). The actual proof breaks x into ‘‘level sets’’ that of values that equal one another up $1 \pm \epsilon$ multiplicative factors, and then analyzes the interaction between

these level sets. This turns out to be quite involved but ultimately doable.

We should note that there is a qualitative difference between the known proof of 2. and the proof of [Theorem 3](#) and it is that the known proof does *not* supply a simple procedure to “round” any vector x that satisfies $\|x\|_p \geq C\|x\|_2$ to a small set S that doesn’t expand, but rather transforms such a vector to either such a set or a different vector $\|x'\|_p \geq 2C\|x'\|_2$, which is a procedure that eventually converges.

Subexponential time algorithm for small set expansion

At the heart of the sub-exponential time algorithm for small set expansion is the following result:

5. Theorem (Structure theorem for small set expanders). *Suppose that G is a d regular graph and let A be the matrix corresponding to the lazy random walk on G (i.e., $A_{i,i} = 1/2$ and $A_{i,j} = 1/(2d)$ for $i \sim j$), with eigenvalues $1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. If $|\{i : \lambda_i \geq 1 - \epsilon\}| \geq n^\beta$ then there exists a vector $x \in \mathbb{R}^n$ such that $x^\top Ax \geq (1 - O(\epsilon/\beta))\|x\|^2$ and $\|x\|_1 \leq \|x\|_2 \sqrt{n^{1-\beta/2}}$.*

Proof. Let Π be the projector matrix to the span of the eigenvectors of A corresponding to eigenvalues over $1 - \epsilon$. Under our assumptions $\text{Tr}(\Pi) \geq n^\beta$ and so in particular there exists i such that $e_i^\top \Pi e_i = \|\Pi e_i\|^2 \geq n^{\beta-1}$. Define $v_j = A^j e_i$ to be the vector corresponding to the probability distribution of taking j steps according to A from the vertex i . Since the eigenvalues of A^j are simply the eigenvalues of A raised to the j^{th} power, and e_i has at least $n^{1-\beta}$ norm-squared projection to the span of eigenvalues of A larger than $1 - \epsilon$, $\|v_j\|_2^2 \geq n^{1-\beta}(1 - \epsilon)^j$ which will be at least $n^{1-\beta/2}$ if $j \leq \beta \log n / (3\epsilon)$. Since v_j is a probability distribution $\|v_j\|_1 = 1$ which means that v_j satisfies the ℓ_1 to ℓ_2 ratio condition of the conclusion. Hence all that is left to show is that there is *some* $j \leq \beta \log n / (3\epsilon)$ such that $v_j^\top A v_j \geq (1 - O(\epsilon/\beta))\|v_j\|^2$. Note that (**exercise!**) it enough to show this for A^2 instead of A (i.e. that $v_j^\top A^2 v_j \geq (1 - O(\epsilon/\beta))\|v_j\|^2$) and so, since $A v_j = v_{j+1}$ it is enough to show that $v_{j+1}^\top v_{j+1} = \|v_{j+1}\|^2 \geq (1 - O(\epsilon/\beta))\|v_j\|^2$.

But since $\|v_0\|_2^2 = 1$, if $\|v_{j+1}\|_2^2 \leq (1 - 10\epsilon/\beta)\|v_j\|^2$ for all such j , then we would get that $\|v_{\beta \log n / (3\epsilon)}\|^2$ would be less than $(1 - 10\epsilon/\beta)^{\beta \log n / (3\epsilon)} \leq 1/n^2$ which is a contradiction as $\|v\|_2^2 \geq 1/n$ for every n dimensional vector v with $\|v\|_1 = 1$. \square

[Theorem 5](#) leads to a subexponential time algorithm for sse quite directly. Given a graph G , if its top eigenspace has dimension at least n^β then it is *not* a small set expander by combining [Theorem 5](#) with the local Cheeger inequality. Otherwise we can do a “brute force search” in time $\exp(\tilde{O}(n^\beta))$ over this subspace to search for vector that is (close to) the characteristic vector of a sparse non expanding set. Using the results of “fixing a vector in a subspace” we saw before, we can also do this via a $\tilde{O}(n^\beta)$ degree sos. Getting a subexponential algorithm for unique games is more involved but uses the same ideas. Namely we obtain a refined structure theorem that states that for any graph G , by removing $o(1)$ fraction of the edges we can decompose G into a collection of disjoint subgraphs each with (essentially) at most n^β large eigenvalues. See chapter 5 of ([Steurer 2010](#)) for more details.

References

- Sanjeev Arora, Boaz Barak, and David Steurer. Subexponential algorithms for unique games and related problems. *J. ACM*, 62(5): 42:1–42:25, 2015.
- Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Near-optimal algorithms for unique games. In *STOC*, pages 205–214. ACM, 2006.
- Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. ACM*, 42(6):1115–1145, 1995.
- Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 767–775. ACM, New York, 2002. doi: 10.1145/509907.510017. URL <http://dx.doi.org/10.1145/509907.510017>.
- Ryan O’Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014.
- Prasad Raghavendra and David Steurer. Graph expansion and the unique games conjecture. In *STOC*, pages 755–764. ACM, 2010.
- Prasad Raghavendra, David Steurer, and Prasad Tetali. Approximations for the isoperimetric and spectral profile of graphs and related parameters. In *STOC*, pages 631–640. ACM, 2010.
- Prasad Raghavendra, David Steurer, and Madhur Tulsiani. Reductions between expansion problems. In *2012 IEEE 27th Conference on*

Computational Complexity—CCC 2012, pages 64–73. IEEE Computer Soc., Los Alamitos, CA, 2012. doi: 10.1109/CCC.2012.43. URL <http://dx.doi.org/10.1109/CCC.2012.43>.

David Steurer. *On the complexity of Unique Games and graph expansion*. PhD thesis, Princeton University, 2010.