The Complexity of Approximating Vertex Expansion

Anand Louis
College of Computing
Georgia Tech
Atlanta, USA
anandl@gatech.edu

Prasad Raghavendra
EECS
UC Berkeley
Berkeley, USA
prasad@cs.berkeley.edu

Santosh Vempala
College of Computing
Georgia Tech
Atlanta, USA
vempala@gatech.edu

***Abstract*—We study the complexity of approximating the vertex expansion of graphs $G = (V, E)$, defined as $\phi^{\mathsf{V}} \stackrel{\text{def}}{=} \min_{S \subset V} n \cdot |N(S)| / (|S|\, |V \setminus S|)$. We give a simple polynomial-time algorithm for finding a subset with vertex expansion $\mathcal{O}\left(\sqrt{\phi^{\mathsf{V}} \log d}\right)$ where d is the maximum degree of the graph. Our main result is an asymptotically matching lower bound: under the Small Set Expansion (SSE) hypothesis, it is hard to find a subset with expansion less than $C\sqrt{\phi^{\mathsf{V}} \log d}$ for an absolute constant C. In particular, this implies for all constant $\varepsilon > 0$, it is SSE-hard to distinguish whether the vertex expansion $< \varepsilon$ or at least an absolute constant. The analogous threshold for edge expansion is $\sqrt{\phi}$ with no dependence on the degree (Here ϕ denotes the optimal edge expansion). Thus our results suggest that vertex expansion is harder to approximate than edge expansion. In particular, while Cheeger's algorithm can certify constant edge expansion, it is SSE-hard to certify constant vertex expansion in graphs.**

***Keywords*-Graph Partitioning, Vertex Expansion, Hardness of Approximation, Small Set Expansion.**

I. Introduction

Vertex expansion is an important parameter associated with a graph, one that has played a major role in both algorithms and complexity. Given a graph $G = (V, E)$, the vertex expansion of a set $S \subseteq V$ of vertices is defined as

$$\phi^{\mathsf{V}}(S) \stackrel{\text{def}}{=} |V| \cdot \frac{|N(S)|}{|S|\, |V \setminus S|}$$

Here $N(S)$ denotes the outer boundary of the set S, i.e. $N(S) = \{i \in V \backslash S | \exists u \in S \text{ such that } \{u, v\} \in E\}$. The vertex expansion of the graph is given by $\phi^{\mathsf{V}} \stackrel{\text{def}}{=} \min_{S \subset V} \phi^{\mathsf{V}}(S)$. The problem of computing ϕ^{V} is a major primitive for many graph algorithms specifically for those that are based on the divide and conquer paradigm [LR99]. It is NP-hard to compute the vertex expansion ϕ^{V} of a graph exactly. In this work, we study the approximability of vertex expansion ϕ^{V} of a graph.

A closely related notion to vertex expansion is that of edge expansion. The edge expansion of a set S is defined as $\phi(S) \stackrel{\text{def}}{=} \mu(E(S, \bar{S}))/\mu(S)$ and the edge expansion of the graph is $\phi = \min_{S \subset V} \phi(S)$. Graph expansion problems have received much attention over the past decades, with applications to many algorithmic problems, to the construction of pseudorandom objects and more recenlty due to their connection to the unique games conjecture.

The problem of approximating edge or vertex expansion can be studied at various regimes of parameters of interest. Perhaps the simplest possible version of the problem is to distinguish whether a given graph is an expander. Fix an absolute constant δ_0. A graph is a δ_0-vertex (edge) expander if its vertex (edge) expansion is at least δ_0. The problem of recognizing a vertex expander can be stated as follows:

Problem I.1. Given a graph G, distinguish between the following two cases: (i) *(Non-Expander)* the vertex expansion is $< \varepsilon$ and (ii) *(Expander)* the vertex expansion is $> \delta_0$ for some absolute constant δ_0. Similarly, one can define the problem of recognizing an edge expander graph.

Notice that if there is some sufficiently small absolute constant ε (depending on δ_0), for which the above problem is easy, then we could argue that it is easy to "recognize" a vertex expander. For the edge case, the Cheeger's inequality yields an algorithm to recognize an edge expander. In fact, it is possible to distinguish a δ_0 edge expander graph, from a graph whose edge expansion is $< \delta_0^2/2$, by just computing the second eigenvalue of the graph Laplacian.

It is natural to ask if there is an efficient algorithm with an analogous guarantee for vertex expansion. More precisely, is there some sufficiently small ε (an arbitrary function of δ_0), so that one can efficiently distinguish between a graph with vertex expansion $> \delta_0$ from one with vertex expansion $< \varepsilon$. In this work, we show a hardness result suggesting that there is no efficient algorithm to recognize vertex expanders. More precisely, our main result is a hardness for the problem of approximating vertex expansion in graphs of bounded degree d. The hardness result shows that the approximability of vertex expansion degrades with the degree, and therefore the problem of recognizing expanders is hard for sufficiently large degree. Furthermore, we exhibit an approximation algorithm for vertex expansion whose guarantee matches the hardness result up to constant factors.

Related Work.: The first approximation for conductance was obtained by discrete analogues of the Cheeger inequality shown by Alon-Milman [AM85] and Alon [Alo86]. Specifi-

0272-5428/13 $26.00 © 2013 IEEE
DOI 10.1109/FOCS.2013.46

IEEE computer society

cally, Cheeger's inequality relates the conductance ϕ to the second eigenvalue of the adjacency matrix of the graph – an efficiently computable quantity. This yields an approximation algorithm for ϕ, one that is used heavily in practice for graph partitioning. However, the approximation for ϕ obtained via Cheeger's inequality is poor in terms of a approximation ratio, especially when the value of ϕ is small. An $\mathcal{O}(\log n)$ approximation algorithm for ϕ was obtained by Leighton and Rao [LR99]. Later work by Linial et al. [LLR95] and Aumann and Rabani [AR98] established a strong connection between the SPARSEST CUT problem and the theory of metric spaces, in turn spurring a large and rich body of literature. The current best algorithm for the problem is an $O(\sqrt{\log n})$ approximation for due to Arora et al. [ARV04] using semidefinite programming techniques.

Ambühl, Mastrolilli and Svensson [AMS07] showed that ϕ^{V} and ϕ have no PTAS assuming that SAT does not have sub-exponential time algorithms. The current best approximation factor for ϕ^{V} is $\mathcal{O}\left(\sqrt{\log n}\right)$ obtained using a convex relaxation [FHL08]. Beyond this, the situation is much less clear for the approximability of vertex expansion. Applying Cheeger's method leads to a bound of $\mathcal{O}\left(\sqrt{d\,\mathrm{OPT}}\right)$ [Alo86] where d is the maximum degree of the input graph.

Small Set Expansion Hypothesis.: A more refined measure of the edge expansion of a graph is its expansion profile. Specifically, for a graph G the expansion profile is given by the curve $\phi(\delta) = \min_{\mu(S) \leqslant \delta} \phi(S)\ \forall \delta \in [0, 1/2]$. The problem of approximating the expansion profile has received much less attention, and is seemingly far less tractable. In summary, the current state-of-the-art algorithms for approximating the expansion profile of a graph are still far from satisfactory. Specifically, the following hypothesis is consistent with the known algorithms for approximating expansion profile.

Hypothesis (Small-Set Expansion Hypothesis, [RS10])**.** *For every constant $\eta > 0$, there exists sufficiently small $\delta > 0$ such that given a graph G it is NP-hard to distinguish the cases,* (YES) *there exists a vertex set S with volume $\mu(S) = \delta$ and expansion $\phi(S) \leqslant \eta$, and* NO *all vertex sets S with volume $\mu(S) = \delta$ have expansion $\phi(S) \geqslant 1 - \eta$.*

Apart from being a natural optimization problem, the SMALL-SET EXPANSION problem is closely tied to the Unique Games Conjecture. Recent work by Raghavendra-Steurer [RS10] established reduction from the SMALL-SET EXPANSION problem to the well known Unique Games problem, thereby showing that Small-Set Expansion Hypothesis implies the Unique Games Conjecture. This result suggests that the problem of approximating expansion of small sets lies at the combinatorial heart of the Unique Games problem.

The Unique Games Conjecture is not known to imply hardness results for problems closely tied to graph expansion such as BALANCED SEPARATOR . The reason being that the hard instances of these problems are required to have certain global structure namely expansion. Gadget reductions from a unique games instance preserve the global properties of the unique games instance such as lack of expansion. Therefore, showing hardness for graph expansion problems often required a stronger version of the EXPANDING UNIQUE GAMES , where the instance is guaranteed to have good expansion. To this end, several such variants of the conjecture for expanding graphs have been defined in literature, some of which turned out to be false [AKK$^+$08]. The Small-Set Expansion Hypothesis could possibly serve as a natural unified assumption that yields all the implications of expanding unique games and, in addition, also hardness results for other fundamental problems such as BALANCED SEPARATOR . In fact, Raghavendra, Steurer and Tulsiani [RST12] show that the the SSE hypothesis implies that the Cheeger's algorithm yields the best approximation for the balanced separator problem.

Formal Statement of Results.: Our first result is a simple polynomial-time algorithm to obtain a subset of vertices S whose vertex expansion is at most $\mathcal{O}\left(\sqrt{\phi^{\mathsf{V}} \log d}\right)$. Here d is the largest vertex degree of G. The algorithm is based on a Poincairé-type graph parameter called λ_∞ defined by Bobkov, Houdré and Tetali [BHT00], which approximates ϕ^{V}. While λ_∞ also appears to be hard to compute, its natural SDP relaxation gives a bound that is within $\mathcal{O}(\log d)$, as observed by Steurer and Tetali [ST12], which inspires our first Theorem.

Theorem I.2. *There exists a polynomial time algorithm which given a graph $G = (V, E)$ having vertex degrees at most d, outputs a set $S \subset V$, such that $\phi^{\mathsf{V}}(S) = \mathcal{O}\left(\sqrt{\phi^{\mathsf{V}}_G \log d}\right)$.*

It is natural to ask if one can prove better inapproximability results for vertex expansion than those that follow from the inapproximability results for edge expansion. Indeed, the best one could hope for would be a lower bound matching the upper bound in the above theorem. Our main result is a reduction from SSE to the problem of distinguishing between the case when vertex expansion of the graph is at most ε and the case when the vertex expansion is at least $\Omega(\sqrt{\varepsilon \log d})$. This immediately implies that it is SSE-hard to find a subset of vertex expansion less than $C\sqrt{\phi^{\mathsf{V}} \log d}$ for some constant C. To the best of our knowledge, our work is the first evidence that vertex expansion might be harder to approximate than edge expansion. More formally, we state our main theorem below.

Theorem I.3. *There exists absolute constants C, C_0 such that for every $\eta, \varepsilon > 0$ the following holds: Given a graph $G = (V, E)$ with maximum degree $d \geqslant C_0/\varepsilon$, it is SSE-hard to distinguish whether (i) There exists a set $S \subset V$ of size $|S| \leqslant |V|/2$ such that $\phi^{\mathsf{V}}(S) \leqslant \varepsilon$ and (ii) For all sets*

$S \subset V$, $\phi^{\mathsf{V}}(S) \geqslant \min\{10^{-10}, C\sqrt{\varepsilon \log d}\} - \eta$.

By a suitable choice of parameters in the above theorem, we obtain the following theorem.

Theorem I.4. *There exists an absolute constant $\delta_0 > 0$ such that for every constant $\varepsilon > 0$ the following holds: Given a graph $G = (V, E)$, it is* SSE*-hard to distinguish between the following two cases: (i)There exists a set $S \subset V$ of size $|S| = |V|/2$ such that $\phi^{\mathsf{V}}(S) \leqslant \varepsilon$ and (ii) (G is a vertex expander with constant expansion) For all sets $S \subset V$, $\phi^{\mathsf{V}}(S) \geqslant \delta_0$*

In particular, the above result implies that it is SSE-hard to certify that a graph is a vertex expander with constant expansion. This is in contrast to the case of edge expansion, where the Cheeger's inequality can be used to certify that a graph has constant edge expansion.

At the risk of being redundant, we note that our main theorem implies that any algorithm that outputs a set having vertex expansion less than $C\sqrt{\phi^{\mathsf{V}} \log d}$ will disprove the SSE hypothesis; alternatively, to improve on the bound of $\mathcal{O}\left(\sqrt{\phi^{\mathsf{V}} \log d}\right)$, one has to disprove the SSE hypothesis. From an algorithmic standpoint, we believe that Theorem I.4 exposes a clean algorithmic challenge of recognizing a vertex expander – a challenging problem that is not only interesting on its own right, but whose resolution would probably lead to a significant advance in approximation algorithms.

At a high level, the proof is as follows. We introduce the notion of BALANCED ANALYTIC VERTEX EXPANSION for Markov chains. This quantity can be thought of as a CSP on $(d+1)$-tuples of vertices. We show a reduction from BALANCED ANALYTIC VERTEX EXPANSION of a Markov chain, say H, to vertex expansion of a graph, say H_1 (Section VI-A). Our reduction is generic and works for any Markov chain H. Surprisingly, the CSP-like nature of BALANCED ANALYTIC VERTEX EXPANSION makes it amenable to a reduction from SMALL-SET EXPANSION (Section VI). We construct a gadget for this reduction and study its embedding into the Gaussian graph to analyze its soundness (Section IV and Section V). The gadget involves a sampling procedure to generate a bounded-degree graph.

II. PROOF OVERVIEW

BALANCED ANALYTIC VERTEX EXPANSION .: To exhibit a hardness result, we begin by defining a combinatorial optimization problem related to the problem of approximating vertex expansion in graphs having largest degree d. This problem referred to as BALANCED ANALYTIC VERTEX EXPANSION can be motivated as follows.

Fix a graph $G = (V, E)$ and a subset of vertices $S \subset V$. For any vertex $v \in V$, v is on the boundary of the set S if and only if $\max_{u \in N(v)} |\mathbb{I}_S[u] - \mathbb{I}_S[v]| = 1$, where $N(v)$ denotes the neighbourhood of vertex v. In particular, the fraction of vertices on the boundary of S is given by $\mathbb{E}_v \max_{u \in N(v)} |\mathbb{I}_S[u] - \mathbb{I}_S[v]|$. The *symmetric* vertex expansion of the set $S \subseteq V$ is given by,

$$n \cdot \frac{|N(S) \cup N(V \backslash S)|}{|S|\,|V \backslash S|} = \frac{\mathbb{E}_v \max_{u \in N(v)} |\mathbb{I}_S[u] - \mathbb{I}_S[v]|}{\mathbb{E}_{u,v} |\mathbb{I}_S[u] - \mathbb{I}_S[v]|}.$$

Note that for a degree d graph, each of the terms in the numerator is maximization over the d edges incident at the vertex. The formal definition of BALANCED ANALYTIC VERTEX EXPANSION is as shown below.

Definition II.1. An instance of BALANCED ANALYTIC VERTEX EXPANSION , denoted by $(V, \mathcal{P})$, consists of a set of variables V and a probability distribution $\mathcal{P}$ over $(d+1)$-tuples in V^{d+1}. The probability distribution $\mathcal{P}$ satisfies the condition that all its $d+1$ marginal distributions are the same (denoted by μ). The goal is to solve the following optimization problem

$$\Phi(V, \mathcal{P}) \stackrel{\text{def}}{=} \min_{F: V \to \{0,1\} \,|\, \mathbb{E}_{X,Y \sim \mu} |F(X) - F(Y)| \geqslant \frac{1}{100}} \frac{\mathbb{E}_{(X, Y_1, \ldots, Y_d) \sim \mathcal{P}} \max_i |F(Y_i) - F(X)|}{\mathbb{E}_{X,Y \sim \mu} |F(X) - F(Y)|}$$

For constant d, this could be thought of as a constraint satisfaction problem (CSP) of arity $d+1$. Every d-regular graph G has an associated instance of BALANCED ANALYTIC VERTEX EXPANSION whose value corresponds to the vertex expansion of G. Conversely, we exhibit a reduction from BALANCED ANALYTIC VERTEX EXPANSION to problem of approximating vertex expansion in a graph of degree $\mathrm{poly}(d)$ (Section VI-A for details).

Dictatorship Testing Gadget.: As with most hardness results obtained via the label cover or the unique games problem, central to our reduction is an appropriate dictatorship testing gadget.

Simply put, a dictatorship testing gadget for BALANCED ANALYTIC VERTEX EXPANSION is an instance $\mathcal{H}^R$ of the problem such that, on one hand there exists the so-called *dictator* assignments with value ε, while every assignment far from every dictator incurs a cost of at least $\Omega(\sqrt{\varepsilon \log d})$.

The construction of the dictatorship testing gadget is as follows. Let H be a Markov chain on vertices $\{s, t, t', s'\}$ connected to form a path of length three. The transition probabilities of the Markov chain $\mathcal{H}$ are so chosen to ensure that if μ_H is the stationary distribution of H then $\mu_H(t) = \mu_H(t') = \varepsilon/2$ and $\mu_H(s) = \mu_H(s') = (1-\varepsilon)/2$. In particular, H has a vertex separator $\{t, t'\}$ whose weight under the stationary distribution is only ε.

The dictatorship testing gadget is over the product Markov chain H^R for some large constant R. The constraints $\mathcal{P}$ of the dictatorship testing gadget H^R are given by the following sampling procedure,

– Sample $x \in H^R$ from the stationary distribution of the chain.

– Sample d-neighbours $y_1, \ldots, y_d \in H^R$ of x independently from the transition probabilities of the chain H^R. Output the tuple $(x, y_1, \ldots, y_d)$.

For every $i \in [R]$, the i^{th} dictator solution to the above described gadget is given by the following function,

$$F(x) = \begin{cases} 1 & \text{if } x_i \in \{s, t\} \\ 0 & \text{otherwise} \end{cases}$$

It is easy to see that for each constraint $(x, y_1, \ldots, y_d) \sim \mathcal{P}$, $\max_j |F(x) - F(y_j)| = 0$ unless $x_i = t$ or $x_i = t'$. Since x is sampled from the stationary distribution for μ_H, $x_i \in \{t, t'\}$ happens with probability ε. Therefore the expected cost incurred by the i^{th} dictator assignment is at most ε.

Soundness Analysis of the Gadget.: The soundness property desired of the dictatorship testing gadget can be stated in terms of influences. Specifically, given an assignment $F : V(H)^R \to [0, 1]$, the influence of the i^{th} coordinate is given by $\mathrm{Inf}_i[F] = \mathbb{E}_{x_{[R]\setminus i}} \mathsf{Var}_{x_i}[F(x)]$, i.e., the expected variance of the function after fixing all but the i^{th} coordinate randomly. Henceforth, we will refer to a function $F : H^R \to [0, 1]$ as *far from every dictator* if the influence of all of its coordinates are small (say $< \tau$).

We show that the dictatorship testing gadget H^R described above satisfies the following soundness – for every function F that is far from every dictator, the cost of F is at least $\Omega(\sqrt{\varepsilon \log d})$. To this end, we appeal to the invariance principle to translate the cost incurred to a corresponding isoperimetric problem on the Gaussian space. More precisely, given a function $F : H^R \to [0, 1]$, we express it as a polynomial in the eigenfunctions over H. We carefully construct a Gaussian ensemble with the same moments up to order two, as the eigenfunctions at the query points $(x, y_1, \ldots, y_d) \in \mathcal{P}$. By appealing to the invariance principle for low degree polynomials, this translates in to the following isoperimetric question over Gaussian space $\mathcal{G}$.,

Suppose we have a subset $S \subseteq \mathcal{G}$ of the n-dimensional Gaussian space. Consider the following experiment:

– Sample a point $z \in \mathcal{G}$ the Gaussian space.
– Pick d independent perturbations $z'_1, z'_2, \ldots, z'_d$ of the point z by ε-noise.
– Output 1 if at least one of the edges (z, z'_i) crosses the cut $(S, \bar{S})$ of the Gaussian space.

Among all subsets S of the Gaussian space with a given volume, which set has the least expected output in the above experiment? The answer to this isoperimetric question corresponds to the soundness of the dictatorship test. A halfspace of volume $\frac{1}{2}$ has an expected output of $\sqrt{\varepsilon \log d}$ in the above experiment. We show that among all subsets of constant volume, halfspaces acheive the least expected output value.

This isoperimetric theorem proven in Section IV yields the desired $\Omega(\sqrt{\varepsilon \log d})$ bound for the soundness of the dictatorship test constructed via the Markov chain H. Here the noise rate of ε arises from the fact that all the eigenfunctions of the Markov chain H have an eigenvalue smaller than $1 - \varepsilon$. The details of the argument based on invariance principle is presented in Section V

We show a $\Omega(\sqrt{\varepsilon \log d})$ lower bound for the isoperimetric problem on the Gaussian space. The proof of this isoperimetric inequality is included in Section IV

We would like to point out here that the traditional noisy cube gadget does not suffice for our application. This is because in the noisy cube gadget while the dictator solutions have an edge expansion of ε they have a vertex expansion of εd, yielding a much worse value than the soundness.

Reduction from SMALL-SET EXPANSION *problem.:* Gadget reductions from the UNIQUE GAMES problem cannot be used towards proving a hardness result for edge or vertex expansion problems. This is because if the underlying instance of UNIQUE GAMES has a small vertex separator, then the graph produced via a gadget reduction would also have small vertex expansion. Therefore, we appeal to a reduction from the SMALL-SET EXPANSION problem (Section VI for details).

Raghavendra et al. [RST12] show optimal inapproximability results for the Balanced separator problem using a reduction from the SMALL-SET EXPANSION problem. While the overall approach of our reduction is similar to theirs, the details are subtle.

Notation.: We use μ_G to denote a probability distribution on vertices of the graph G. We drop the subscript G, when the graph is clear from the context. For a set of vertices S, we define $\mu(S) = \int_{x \in S} \mu(x)$. We use $\mu_{|S}$ to denote the distribution μ restricted to the set $S \subset V(G)$. For the sake of simplicity, we sometimes say that vertex $v \in V(G)$ has weight $w(v)$, in which case we define $\mu(v) = w(v) / \sum_{u \in V} w(u)$. We denote the weight of a set $S \subseteq V$ by $w(S)$. We denote the degree of a vertex v by $\deg(v)$. For a random variable X, define the variance and ℓ_1-variance as follows,

$$\mathsf{Var}[X] = \mathop{\mathbb{E}}_{X_1, X_2}[(X_1 - X_2)^2] \qquad \mathsf{Var}_1[X] = \mathop{\mathbb{E}}_{X_1, X_2}[|X_1 - X_2|]$$

where X_1, X_2 are two independent samples of X.

III. PRELIMINARIES

Symmetric Vertex Expansion.: For our proofs, the notion of Symmetric Vertex Expansion is useful.

Definition III.1. Given a graph $G = (V, E)$, we define the the symmetric vertex expansion of a set $S \subset V$ as follows.

$$\Phi^{\mathsf{V}}_G(S) \stackrel{\text{def}}{=} n \cdot \frac{|N_G(S) \cup N_G(V \backslash S)|}{|S|\,|V \backslash S|}$$

Balanced Vertex Expansion.: We define the balanced vertex expansion of a graph as follows.

Definition III.2. Given a graph G and balance parameter b, we define the *b-balanced vertex expansion* of

G as $\phi_b^{\mathsf{V,bal}} \stackrel{\text{def}}{=} \min_{S:|S||V\setminus S|\geqslant bn^2} \phi^{\mathsf{V}}(S)$ and $\Phi_b^{\mathsf{V,bal}} \stackrel{\text{def}}{=} \min_{S:|S||V\setminus S|\geqslant bn^2} \Phi^{\mathsf{V}}(S)$. We define $\phi^{\mathsf{V,bal}} \stackrel{\text{def}}{=} \phi_{1/100}^{\mathsf{V,bal}}$ and $\Phi^{\mathsf{V,bal}} \stackrel{\text{def}}{=} \Phi_{1/100}^{\mathsf{V,bal}}$.

Analytic Vertex Expansion.: Our reduction from SSE to vertex expansion goes via an intermediate problem that we call d-BALANCED ANALYTIC VERTEX EXPANSION . We define the notion of d-BALANCED ANALYTIC VERTEX EXPANSION as follows.

Definition III.3. An instance of d-BALANCED ANALYTIC VERTEX EXPANSION , denoted by $(V, \mathcal{P})$, consists of a set of variables V and a probability distribution $\mathcal{P}$ over $(d+1)$-tuples in V^{d+1}. The probability distribution $\mathcal{P}$ satisfies the condition that all its $d+1$ marginal distributions are the same (denoted by μ). The d-BALANCED ANALYTIC VERTEX EXPANSION under a function $F : V \to \{0,1\}$ is defined as

$$\Phi(V,\mathcal{P})(F) \stackrel{\text{def}}{=} \frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i |F(Y_i) - F(X)|}{\mathbb{E}_{X,Y\sim\mu} |F(X) - F(Y)|}.$$

The d-BALANCED ANALYTIC VERTEX EXPANSION of $(V, \mathcal{P})$ is defined as

$$\Phi(V,\mathcal{P}) \stackrel{\text{def}}{=} \min_{F:V\to\{0,1\} \mid \mathbb{E}_{X,Y\sim\mu}|F(X)-F(Y)| \geqslant \frac{1}{100}} \Phi(V,\mathcal{P})(F).$$

When drop the degree d from the notation, when it is clear from the context.

For an instance $(V, \mathcal{P})$ of BALANCED ANALYTIC VERTEX EXPANSION and an assignment $F : V \to \{0,1\}$ define

$$\mathrm{val}_{\mathcal{P}}(F) = \mathop{\mathbb{E}}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}} \max_i |F(Y_i) - F(X)|.$$

Gaussian Graph.: Recall that two standard normal random variables X, Y are said to be α-correlated if there exists an independent standard normal random variable Z such that $Y = \alpha X + \sqrt{1-\alpha^2} Z$.

Definition III.4. The *Gaussian Graph* $\mathcal{G}_{\Lambda,\Sigma}$ is a complete weighted graph on the vertex set $V(\mathcal{G}_{\Lambda,\Sigma}) = \mathbb{R}^n$. The weights are given by the following probability density function:

$$w(\{u,v\}) = \mathbb{P}[X = u \text{ and } Y = v]$$

where $Y \sim \mathcal{N}(\Lambda X, \Sigma)$, where Λ is a diagonal matrix such that $\|\Lambda\| \leqslant 1$ and $\Sigma \succeq \varepsilon I$ is a diagonal matrix.

Definition III.5. We say that a family of graphs $\mathcal{G}_d$ is $\Theta(d)$-regular, if there exist absolute constants $c_1, c_2 \in \mathbb{R}^+$ such that for every $G \in \mathcal{G}_d$, all vertices $i \in V(G)$ have $c_1 d \leqslant \deg(i) \leqslant c_2 d$.

We now formalize our notion of hardness.

Definition III.6. A constrained minimization problem $\mathcal{A}$ with its optimal value denoted by $\mathrm{val}(\mathcal{A})$ is said to be c-vs-s hard if it is SSE-hard to distinguish between the following two cases: (i) $\mathrm{val}(\mathcal{A}) \leqslant c$ and (ii) $\mathrm{val}(\mathcal{A}) \geqslant s$.

Small-Set Expansion Hypothesis.:

Problem III.7 (SMALL-SET EXPANSION (γ, δ))**.** Given a regular graph $G = (V, E)$, distinguish between the following two cases:

YES: There exists a non-expanding set $S \subset V$ with $\mu(S) = \delta$ and $\Phi_G(S) \leqslant \gamma$.
NO: All sets $S \subset V$ with $\mu(S) = \delta$ are highly expanding having $\Phi_G(S) \geqslant 1 - \gamma$.

Hypothesis III.8 (Hardness of approximating SMALL-SET EXPANSION)**.** *For all $\gamma > 0$, there exists $\delta > 0$ such that the promise problem* SMALL-SET EXPANSION *(γ, δ) is* NP*-hard.*

For the proofs, it will be more convenient to use the following version of the SMALL-SET EXPANSION problem, in which we high expansion is guaranteed not only for sets of measure δ, but also within an arbitrary multiplicative factor of M.

Problem III.9 (SMALL-SET EXPANSION (γ, δ, M))**.** Given a regular graph $G = (V, E)$, distinguish between the following two cases:

YES: There exists a non-expanding set $S \subset V$ with $\mu(S) = \delta$ and $\Phi_G(S) \leqslant \gamma$.
NO: All sets $S \subset V$ with $\mu(S) \in \left(\frac{\delta}{M}, M\delta\right)$ have $\Phi_G(S) \geqslant 1 - \gamma$.

The following stronger hypothesis was shown to be equivalent to Small-Set Expansion Hypothesis in [RST12].

Hypothesis III.10 (Hardness of approximating SMALL-SET EXPANSION)**.** *For all $\gamma > 0$ and $M \geqslant 1$, there exists $\delta > 0$ such that the promise problem* SMALL-SET EXPANSION *(γ, δ, M) is* NP*-hard.*

IV. ISOPERIMETRY OF THE GAUSSIAN GRAPH

In this section we bound the BALANCED ANALYTIC VERTEX EXPANSION of the Gaussian graph. For the Gaussian Graph, we define the canonical probability distribution on V^{d+1} as follows. The marginal distribution along any component X or Y_i is the standard Gaussian distribution in $\mathbb{R}^n$, denoted here by $\mu = \mathcal{N}(0,1)^n$.

$$\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}(X, Y_1, \ldots, Y_d) = \frac{\Pi_{i=1}^d w(X, Y_i)}{\mu(X)^{d-1}} = \mu(X) \Pi_{i=1}^d \mathbb{P}[Y = Y_i].$$

Here, random variable Y is sampled from $\mathcal{N}(\Lambda X, \Sigma)$.

Theorem IV.1. *For any closed set $S \subset of V(\mathcal{G}_{\Lambda,\Sigma})$ with Λ a diagonal matrix satisfying $\|\Lambda\| \leqslant 1$, and Σ a diagonal matrix satisfying $\Sigma \succeq \varepsilon I$, we have*

$$\frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}} \max_i |\mathbb{I}_S[X] - \mathbb{I}_S[Y_i]|}{\mathbb{E}_{X,Y\sim\mu} |\mathbb{I}_S[X] - \mathbb{I}_S[Y]|} \geqslant c\sqrt{\varepsilon \log d}$$

for some absolute constant c.

Lemma IV.2. *Let $u, v \in \mathbb{R}^n$ satisfy $|u - v| \leqslant \sqrt{\varepsilon \log d}$. Let Λ be a diagonal matrix satisfying $\|\Lambda\| \leqslant 1$, and let Σ a diagonal matrix satisfying $\Sigma \succeq \varepsilon I$. Let P_u, P_v be the distributions $\mathcal{N}(\Lambda u, \Sigma)$ and $\mathcal{N}(\Lambda v, \Sigma)$ respectively. Then,*

$$d_{\mathsf{TV}}(P_u, P_v) \leqslant 1 - \frac{1}{d}.$$

The proof follows from from standard Gaussian tail bounds, and we defer it to the full version of the paper [LRV13].

Proof of Theorem IV.1.: Let μ_X denote the Gaussian distribution $\mathcal{N}(\Lambda X, \Sigma)$. Then the LHS is:

$$\int_{\mathbb{R}^n \setminus S} \left(1 - (1 - \mu_X(S))^d\right) \, d\mu(X) + \int_S \left(1 - (1 - \mu_X(\mathbb{R}^n \setminus S))^d\right) \, d\mu(X).$$

To bound this, we will restrict ourselves to points X for which the μ_X measure of the complementary set is at least $1/d$. Roughly speaking, these will be points near the boundary of S. Define:

$$S_1 = \{x \in S : \mu_X(\mathbb{R}^n \setminus S) < \frac{1}{2d}\}$$
$$S_2 = \{x \in \mathbb{R}^n \setminus S : \mu_X(S) < \frac{1}{2d}\}$$

and $S_3 = \mathbb{R}^n \setminus S_1 \setminus S_2$. Using Lemma IV.2 over points in $S_1 * S_2$ along with the Gaussian Isoperimetry Inequality, we get the theorem. ∎

The following corollary falls out as an easy consequence of the theorem.

Corollary IV.3 (Corollary to Theorem IV.1)**.** *Let $F : V(\mathcal{G}_{\Lambda,\Sigma}) \to [0, 1]$ be any function. Then, for some absolute constant c,*

$$\frac{\mathbb{E}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}} \max_i |F(X) - F(Y_i)|}{\mathbb{E}_{X,Y\sim\mu} |F(X) - F(Y)|} \geqslant c\sqrt{\varepsilon \log d}\,.$$

V. Dictatorship Testing Gadget

In this section we initiate the construction of the dictatorship testing gadget for reduction from SSE.

Overall, the dictatorship testing gadget is obtained by picking an appropriately chosen constant sized Markov-chain H, and considering the product Markov chain H^R. Formally, given a Markov chain H, define an instance of Balanced Analytic Vertex Expansion with vertices as V_H and the constraints given by the following canonical probability distribution over V_H^{d+1}.

- Sample $X \sim \mu_H$, the stationary distribution of the Markov chain V_H.
- Sample $Y_1, \ldots, Y_d$ independently from the neighbours of X in V_H

For our application, we use a specific Markov chain H on four vertices. Define a Markov chain H on $V_H = \{s, t, t', s'\}$ as follows,$p(s|s) = p(s'|s') = 1 - \frac{\varepsilon}{1-2\varepsilon}$, $p(t|s) = p(t'|s') = \frac{\varepsilon}{1-2\varepsilon}$, $p(s|t) = p(s'|t') = \frac{1}{2}$ and $p(t'|t) = p(t|t') = \frac{1}{2}$. It is easy to see that the stationary distribution of the Markov chain H over V_H is given by,

$$\mu_H(s) = \mu_H(s') = \frac{1}{2} - \varepsilon \qquad \mu_H(t) = \mu_H(t') = \varepsilon$$

From this Markov chain, construct a dictatorship testing gadget $(V_H^R, \mathcal{P}_H^R)$ as described above. We begin by showing that this dictatorship testing gadget has small vertex separators corresponding to dictator functions.

Proposition V.1 (Completeness)**.** *For each $i \in [R]$, the i^{th}-dictator set defined as $F(x) = 1$ if $x_i \in \{s, t\}$ and 0 otherwise satisfies,* $\mathsf{Var}_1[F] = \frac{1}{2}$ *and* $\mathrm{val}_{\mathcal{P}_{H^R}}(F) \leqslant 2\varepsilon$.

A. Soundness

We will show a general soundness claim that holds for dictatorship testing gadgets $(V(H^R), \mathcal{P}_{H^R})$ constructed out of arbitrary Markov chains H with a given spectral gap.

Polynomials over H^R.: Let $e_0, e_1, \ldots, e_n : V(H) \to \mathbb{R}$ be an orthonormal basis of eigenvectors of H and let $\lambda_0, \ldots, \lambda_n$ be the corresponding eigenvalues. It is easy to see that the eigenvectors of H^R are indexed by $\sigma \in [n]^R$ as $e_\sigma(x) = \prod_{i=1}^R e_{\sigma_i}(x_i)$. Every function $f : H^R \to \mathbb{R}$ can be written in this orthonormal basis $f(x) = \sum_{\sigma\in[n]^R} \hat{f}_\sigma e_\sigma(x)$. For a polynomial $Q = \sum_\sigma \hat{Q}_\sigma e_\sigma$, the polynomial $Q^{>p}$ denotes the projection on to degrees higher than p, i.e., $Q^{>p} = \sum_{\sigma,|\sigma|>p} \hat{Q}_\sigma e_\sigma$. The influences of a polynomial $Q = \sum_\sigma \hat{Q}_\sigma$ are defined as, $\mathrm{Inf}_i(Q) = \sum_{\sigma:\sigma_i\neq 0} \hat{Q}_\sigma^2$.

For a function $Q : V(H^R) \to \mathbb{R}$ (or equivalently a polynomial), $\mathsf{Var}[Q]$ denotes the variance of the random variable $Q(x)$ for a random x from stationary distribution of H^R. It is an easy computation to check that this is given by, $\mathsf{Var}[Q] = \sum_{\sigma:|\sigma|\neq 0} \hat{Q}_\sigma^2$.

We will make use of the following Invariance Principle due to Isaksson and Mossel [IM12].

Theorem V.2 ([IM12])**.** *Let $X = (X_1, \ldots, X_n)$ be an independent sequence of ensembles, such that $\mathbb{P}[X_i = x] \geqslant \alpha > 0, \forall i, x$. Let Q be a d-dimensional multilinear polynomial such that $\mathsf{Var}(Q_j(X)) \leqslant 1$, $\mathsf{Var}(Q_j^{>p}) \leqslant (1 - \varepsilon\eta)^{2p}$ and $\mathrm{Inf}_i(Q_j) \leqslant \tau$ where $p = \frac{1}{18} \log(1/\tau)/\log(1/\alpha)$. Finally, let $\psi : \mathbb{R}^k \to \mathbb{R}$ be Lipschitz continuous. Then,*

$$\Big|\mathbb{E}[\psi(Q(X))] - \mathbb{E}[\psi(Q(Z))]\Big| = \mathcal{O}\left(\tau^{\frac{\varepsilon\eta}{18}/\log\frac{1}{\alpha}}\right)$$

where Z is an independent sequence of Gaussian ensembles with the same covariance structure as X.

Noise Operator.: We define a noise operator $\Gamma_{1-\eta}$ on functions on the Markov chain H as follows :

$$\Gamma_{1-\eta} F(X) \stackrel{\text{def}}{=} (1 - \eta)F(X) + \eta \mathop{\mathbb{E}}_{Y\sim X} F(Y)$$

for every function $F : H \to \mathbb{R}$. Similarly, one can define the noise operator $\Gamma_{1-\eta}$ on functions over H^R.

Proposition V.3 (Soundness). *For all $\varepsilon, \eta, \alpha, \tau > 0$ the following holds. Let H be a finite Markov-chain with a spectral gap of at least ε, and the probability of every state under stationary distribution is $\geqslant \alpha$. Let $F : V(H^R) \to \{0,1\}$ be a function such that $\max_{i\in[R]} \operatorname{Inf}_i(\Gamma_{1-\eta}F) \leqslant \tau$. Then we have*

$$\mathop{\mathbb{E}}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}}[\max_i |F(Y_i) - F(X)|] \geqslant$$
$$\Omega(\sqrt{\varepsilon \log d}) \mathop{\mathbb{E}}_{X,Y\sim\mu_{H^R}} |F(X) - F(Y)| - \mathcal{O}(\eta) - \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))})$$

Proof: Let $Q = (Q_0, Q_1, \ldots, Q_d)$ be the multi-linear polynomial representation of the vector-valued function $(\Gamma_{1-\eta}F(X), \Gamma_{1-\eta}F(Y_1), \ldots, \Gamma_{1-\eta}F(Y_d))$.

Define a function $s : \mathbb{R} \to \mathbb{R}$ as follows $s(x) = 0$ if $x < 0$, $s(x) = x$ if $x \in [0,1]$ and $s(x) = 1$ if $x > 1$. Define a function $\Psi : \mathbb{R}^{d+1} \to \mathbb{R}$ as, $\Psi(x, y_1, \ldots, y_d) = \max_i |s(y_i) - s(x)|$. Clearly, Ψ is a Lipshitz function with a constant of 1.

Apply the invariance principle to the polynomial $Q = (\Gamma_{1-\eta}F, \Gamma_{1-\eta}F, \ldots, \Gamma_{1-\eta}F)$ and Lipshitz function Ψ. By invariance principle Theorem V.2, we get for some appropriate Λ,

$$\mathop{\mathbb{E}}_{(X,Y_1,\ldots,Y_d)\sim\mathcal{P}_{H^R}} \max_a |s(\Gamma_{1-\eta}F(X)) - s(\Gamma_{1-\eta}F(Y_a))| \geqslant$$
$$\mathop{\mathbb{E}}_{(Z_X,Z_{Y_1},\ldots,Z_{Y_d})\sim\mathcal{P}_{\mathcal{G}_{\Lambda,\Sigma}}} \max_a |s(\Gamma_{1-\eta}F(Z_X)) - s(\Gamma_{1-\eta}F(Z_{Y_a}))$$
$$- \tau^{\Omega(\varepsilon\eta/\log(1/\alpha))}$$

We can now use Corollary IV.3 to finish the proof. ■

VI. Hardness Reduction from SSE

Let $G = (V,E)$ be an instance of Small-Set Expansion (γ, δ, M). Starting with the instance $G = (V,E)$ of Small-Set Expansion(γ, δ, M), our reduction produces an instance $(\mathcal{V}', \mathcal{P}')$ of Balanced Analytic Vertex Expansion .

To describe our reduction, let us fix some notation. For a set A, let $A^{\{R\}}$ denote the set of all multisets with R elements from A. Let $G_\eta = (1-\eta)G + \eta K_V$ where K_V denotes the complete graph on the set of vertices V. For an integer R, define $G_\eta^{\otimes R}$ to be the product graph $G_\eta^{\otimes R} = (G_\eta)^R$.

Define a Markov chain H on $V_H = \{s, t, t', s'\}$ as follows,$p(s|s) = p(s'|s') = 1 - \frac{\varepsilon}{1-2\varepsilon}$, $p(t|s) = p(t'|s') = \frac{\varepsilon}{1-2\varepsilon}$, $p(s|t) = p(s'|t') = \frac{1}{2}$ and $p(t'|t) = p(t|t') = \frac{1}{2}$. It is easy to see that the stationary distribution of the Markov chain H over V_H is given by, $\mu_H(s) = \mu_H(s') = 1/2 - \varepsilon$ and $\mu_H(t) = \mu_H(t') = \varepsilon$.

The reduction consists of two steps. First, we construct an "unfolded" instance $(\mathcal{V}, \mathcal{P})$ of the Balanced Analytic Vertex Expansion , then we merge vertices of $(\mathcal{V}, \mathcal{P})$ to create the final output instance $(\mathcal{V}', \mathcal{P}')$. The details of the reduction are presented in Figure 1.

Reduction

Input: A graph $G = (V,E)$ - an instance of Small-Set Expansion(γ, δ, M).

Parameters: $R = \frac{1}{\delta}$, ε

Unfolded instance $(\mathcal{V}, \mathcal{P})$

Set $\mathcal{V} = (V \times V_H)^R$. The probability distribution μ on $\mathcal{V}$ is given by $(\mu_V \times \mu_H)^R$. The probability distribution $\mathcal{P}$ is given by the following sampling procedure.

1) Sample a random vertex $A \in V^R$.
2) Sample $d+1$ random neighbors $B, C_1, \ldots, C_d \sim G_\eta^{\otimes R}(A)$ of the vertex A in the tensor-product graph $G_\eta^{\otimes R}$.
3) Sample $x \in V_H^R$ from the product distribution μ^R.
4) Independently sample d neighbours $y^{(1)}, \ldots, y^{(d)}$ of x in the Markov chain H^R, i.e., $y^{(i)} \sim \mu_H^R(x)$.
5) Output $((B,x), (C_1, y_1), \ldots, (C_d, y_d))$

Folded Instance $(\mathcal{V}', \mathcal{P}')$

Fix $\mathcal{V}' = (V \times \{s,t\})^{\{R\}}$. Define a projection map $\Pi : \mathcal{V} \to \mathcal{V}'$ as follows:

$$\Pi(A,x) = \{(a_i, x_i) | x_i \in \{s,t\}\}$$

for each $(A,x) = ((a_1,x_1),(a_2,x_2),\ldots,(a_R,x_R))$ in $(V \times \{s,t\})^{\{R\}}$.

Let μ' be the probability distribution on $\mathcal{V}'$ obtained by projection of probability distribution μ on $\mathcal{V}$. Similarly, the probability distribution $\mathcal{P}'$ on $(\mathcal{V}')^{d+1}$ by applying the projection Π to the probability distribution $\mathcal{P}$.

Figure 1. Reduction from SSE to Vertex Expansion

Observe that each of the queries $\Pi(B,x)$ and $\{\Pi(C_i, y_i)\}_{i=1}^d$ are distributed according to μ' on $\mathcal{V}'$. Let $F' \colon \mathcal{V}' \to \{0,1\}$ denote the indicator function of a subset for the instance. Let us suppose that $\mathbb{E}_{X,Y\sim\mathcal{V}}[|F'(X) - F'(Y)|] \geqslant \frac{1}{10}$. We fix $\eta = \varepsilon/(100d)$. We restrict $\gamma < \varepsilon/(100d)$.

Theorem VI.1. *(Completeness) Suppose there exists a set $S \subset V$ such that $\operatorname{vol}(S) = \delta$ and $\Phi(S) \leqslant \gamma$ then there exists $F' : \mathcal{V}' \to \{0,1\}$ such that, $\mathbb{E}_{X,Y\sim\mathcal{V}'}[|F'(X) - F'(Y)|] \geqslant \frac{1}{10}$ and,*

$$\mathop{\mathbb{E}}_{X,Y_1,\ldots,Y_d\sim\mathcal{P}} \left[\max_i |F'(X) - F'(Y_i)|\right] \leqslant 2\varepsilon + \mathcal{O}(d(\eta+\gamma))$$

Proof: Define $F : \mathcal{V} \to \{0,1\}$ as follows:

$$F(A,x) = \begin{cases} 1 & \text{if } |\Pi(A,x) \cap (S \times \{s,t\})| = 1 \\ 0 & \text{otherwise} \end{cases}$$

Observe that by definition of F, the value of $F(A,x)$ only depends on $\Pi(A,x)$. So the function F naturally defines a map $F' : \mathcal{V}' \to \{0,1\}$. Therefore we can write,

$$\mathbb{P}[F(A,x) = 1] =$$
$$\sum_{i\in[R]} \mathbb{P}[x_i \in \{s,t\}] \, \mathbb{P}[\{a_1,\ldots,a_R\} \cap S = \{a_i\} | x_i \in \{s,t\}]$$

$$\geqslant R \cdot \frac{1}{2} \cdot \frac{1}{R} \cdot \left(1 - \frac{1}{R}\right)^{R-1} \geqslant \frac{1}{10}$$

and,

$$\mathbb{P}\left[F(A,x) = 1\right] = \mathbb{P}\left[|\Pi(A,x) \cap (S \times \{s,t\})| = 1\right]$$
$$\leqslant \mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}}\left[|\Pi(A,x) \cap (S \times \{s,t\})|\right] = R \cdot \frac{1}{2} \cdot \frac{|S|}{|V|} \leqslant \frac{1}{2}$$

The above bounds on $\mathbb{P}\left[F(A,x) = 1\right]$ along with the fact that F takes values only in $\{0,1\}$, we get that

$$\mathop{\mathbb{E}}_{X,Y\sim\mathcal{V}'}|F'(X) - F'(Y)| = \mathop{\mathbb{E}}_{(A,x)\ (B,y)\sim\mathcal{V}}|F(A,x) - F(B,y)| \geqslant \frac{1}{10}$$

Suppose we sample $A \in V^R$ and $B, C_1, \ldots, C_d$ independently from $G_\eta^{\otimes R}(A)$. Let us denote $A = (a_1, \ldots, a_R)$, $B = (b_1, \ldots, b_R)$, $C_i = (c_{i1}, \ldots, c_{iR})$ for all $i \in [d]$. Note that,

$$\mathbb{P}\left[\exists i \in [R] \text{ such that } |\{a_i, b_i\} \cap S| = 1\right]$$
$$\leqslant \sum_{i\in[R]} (1-\eta)\,\mathbb{P}\left[(a_i,b_i) \in E[S,\bar{S}]\right] + \eta\,\mathbb{P}\left[(a_i,b_i) \in S \times \bar{S}\right]$$
$$\leqslant R(\mathrm{vol}(S)\Phi(S) + 2\eta\,\mathrm{vol}(S)) \leqslant 2(\gamma + \eta)$$

Similarly, for each $j \in [d]$,

$$\mathbb{P}\left[\exists i \in [R] \middle| |\{a_i, c_{ji}\} \cap S| = 1\right] \leqslant \sum_{i\in[R]} \mathbb{P}\left[(a_i, c_{ji}) \in E[S,\bar{S}]\right]$$
$$\leqslant R\,\mathrm{vol}(S)\Phi(S) \leqslant 2(\gamma + \eta)\,.$$

By a union bound, with probability at least $1 - 2(d+1)(\gamma + \eta)$ we have that none of the edges $\{(a_i, b_i)\}_{i\in[R]}$ and $\{(a_i, c_{ji})\}_{j\in[d], i\in[R]}$ cross the cut $(S, \bar{S})$.

Conditioned on the above event, we claim that if $(B,x) \cap (S \times \{t,t'\}) = \emptyset$ then $\max_i |F(B,x) - F(C_i,y_i)| = 0$. First, if $(B,x) \cap (S \times \{t,t'\}) = \emptyset$ then for each $b_i \in S$ the corresponding $x_i \in \{s,s'\}$. In particular, this implies that for each $b_i \in S$, either all of the pairs $(b_i, x_i), \{(c_{ji}, y_{ji})\}_{j\in[d]}$ are either in $S \times \{s,t\}$ or $S \times \{s',t'\}$, thereby ensuring that $\max_i |F(B,x) - F(C_i,y_i)| = 0$. From the above discussion we conclude,

$$\mathop{\mathbb{E}}_{(B,x),(C_1,y_1),\ldots,(C_d,y_d)\sim\mathcal{P}}\left[\max_i |F(B,x) - F(C_i,y_i)|\right]$$
$$\leqslant \mathbb{P}\left[|(B,x) \cap (S \times \{t,t'\})\,| \geqslant 1\right] + 2(d+1)(\gamma+\eta)$$
$$\leqslant \mathbb{E}\left[|(B,x) \cap (S \times \{t,t'\})\,|\right] + 2(d+1)(\gamma+\eta)$$
$$= R \cdot \mathrm{vol}(S) \cdot \varepsilon + 2(d+1)(\gamma+\eta) \leqslant e\varepsilon$$

■

Let $F' : \mathcal{V}' \to \{0,1\}$ be a subset of the instance $(\mathcal{V}', \mathcal{P}')$. We define the functions $F : \mathcal{V} \to [0,1]$ and $f_A, g_A : V_H^R \to [0,1]$ for each $A \in V^R$ as $F(A,x) \stackrel{\text{def}}{=} F'(\Pi(A,x))$, $f_A(x) \stackrel{\text{def}}{=} F(A,x)$ and $g_A(x) \stackrel{\text{def}}{=} \mathbb{E}_{B\sim G_\eta^{\otimes R}(A)} F(B,x)$. We defer the proof of the following two lemmas to the full version.

Lemma VI.2. $\mathrm{val}_{\mathcal{P}'}(F') \geqslant \mathbb{E}_{A\in V^R}\,\mathrm{val}_{\mu_H^R}(g_A)$

Lemma VI.3.

$$\mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x\sim\mu_H^R} g_A(x)^2 \geqslant \mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F^2(A,x) - \mathrm{val}_{\mathcal{P}'}(F')$$

The following crucial lemma translates the fact that the set is balanced on the entire instance to a balance within the individual long codes. To this end, it uses the spectral properties of the graph produced in the hardness reduction.

Lemma VI.4. $\mathbb{E}_{A\sim V^R}\,\mathsf{Var}_1[g_A] \geqslant \frac{1}{2}(\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F')$

1 *Proof:* Since the function g_A is bounded in $[0,1]$ we can write

$$\mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x,y\in\mu_H^R}|g_A(x) - g_A(y)|$$
$$\geqslant \mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x,y\in\mu_H^R}(g_A(x) - g_A(y))^2$$
$$\geqslant \mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x\in\mu_H^R} g_A^2(x) - \mathop{\mathbb{E}}_{A}\mathop{\mathbb{E}}_{x,y\in\mu_H^R} g_A(x)g_A(y) \qquad \text{(VI.1)}$$

In the above expression there are two terms. From Lemma VI.3, we already know that

$$\mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x\in\mu_H^R} g_A^2(x) \geqslant \mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F^2(A,x) - \mathrm{val}_{\mathcal{P}'}(F') \qquad \text{(VI.2)}$$

Let us expand out the other term in the expression.

$$\mathop{\mathbb{E}}_{A}\mathop{\mathbb{E}}_{x,y\in\mu_H^R} g_A(x)g_A(y)$$
$$= \mathop{\mathbb{E}}_{A}\mathop{\mathbb{E}}_{B,C\sim G_\eta^{\otimes R}(A)}\mathop{\mathbb{E}}_{x,y\in\mu_H^R} F'(\Pi(B,x))F'(\Pi(C,y)) \qquad \text{(VI.3)}$$

Now consider the following graph $\mathcal{H}$ on $\mathcal{V}'$ defined by the following edge sampling procedure.

– Sample $A \in V^R$, and $x, y \in \mu_H^R$.
– Sample independently $B \sim G_\eta^{\otimes R}(A)$ and $C \sim G_\eta^{\otimes R}(A)$
– Output the edge $\Pi(B,x)$ and $\Pi(C,y)$

Let λ denote the second eigenvalue of the adjacency matrix of the graph $\mathcal{H}$.

$$\mathop{\mathbb{E}}_{A}\mathop{\mathbb{E}}_{B,C\sim G_\eta^{\otimes R}(A)}\mathop{\mathbb{E}}_{x,y\in\mu_H^R} F'(\Pi(B,x))F'(\Pi(C,y)) = \langle F', \mathcal{H}F'\rangle$$
$$\leqslant \left(\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F'(\Pi(A,x))\right)^2$$
$$+ \lambda\left(\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}}(F'(\Pi(A,x)))^2 - (\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F'(\Pi(A,x)))^2\right)$$
$$= \lambda \mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F(A,x)^2 + (1-\lambda)(\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F(A,x))^2$$

(because $F'(\Pi(A,x)) = F(A,x)$)

Using the above inequality with equations VI.1, VI.2, VI.3 we can derive the following,

$$\mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x,y\in\mu_H^R}|g_A(x) - g_A(y)|$$
$$\geqslant \mathop{\mathbb{E}}_{A\sim V^R}\mathop{\mathbb{E}}_{x\in\mu_H^R} g_A^2(x) - \mathop{\mathbb{E}}_{A}\mathop{\mathbb{E}}_{x,y\in\mu_H^R} g_A(x)g_A(y)$$

$$\geqslant (1-\lambda)\left[\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F^2(A,x) - (\mathop{\mathbb{E}}_{(A,x)\sim\mathcal{V}} F(A,x))^2\right] - \mathrm{val}_{\mathcal{P}'}(F\ )$$
$$\geqslant (1-\lambda)\mathsf{Var}[F] - \mathrm{val}_{\mathcal{P}'}(F')$$
$$\geqslant (1-\lambda)(\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F')$$

To finish the argument, we need to bound the second eigenvalue λ for the graph $\mathcal{H}$. Here we will present a simple argument showing that the second eigenvalue λ for the graph $\mathcal{H}$ is strictly less than $\frac{1}{2}$. Let us restate the procedure to sample edges from $\mathcal{H}$ slightly differently.

– Define a map $\mathcal{M} : V\times V_H \to (V\cup \perp)\times(V_H\cup\{\perp\})$ as follows, $\mathcal{M}(b,x) = (b,x)$ if $x \in \{s,t\}$ and $\mathcal{M}(b,x) = (\perp,\perp)$ otherwise. Let $\Pi' : ((V\cup \perp)\times(V_H\cup \perp))^R \to (V\times\{s,t\})^{\{R\}}$ denote the following map.

$$\Pi'(B',x') = \{(b'_i,x'_i) | x_i \in \{s,t\}\}$$

– Sample $A \in V^R$ and $x,y \in \mu_H^R$
– Sample independently $B = (b_1,\dots,b_R) \sim G_\eta^{\otimes R}(A)$ and $C = (c_1,\dots,c_R) \sim G_\eta^{\otimes R}(A)$.
– Let $\mathcal{M}(B,x), \mathcal{M}(C,y) \in ((V\cup\{\perp\})\times(V_H\cup\{\perp\}))^R$ be obtained by applying $\mathcal{M}$ to each coordinate of (B,x) and (C,y).
– Output an edge between $(\Pi'(\mathcal{M}(B,x)), \Pi'(\mathcal{M}(C,y)))$.

It is easy to see that the above procedure also samples the edges of $\mathcal{H}$ from the same distribution as earlier. Note that Π' is a projection from $((V\cup \perp)\times(V_H\cup \perp))^R$ to $(V\times\{s,t\})^{\{R\}}$. Therefore, the second eigenvalue of the graph $\mathcal{H}$ is upper bounded by the second eigenvalue of the graph on $((V\cup \perp)\times(V_H\cup\{\perp\}))^R$ defined by $\mathcal{M}(B,x)\sim\mathcal{M}(C,y)$. Let $\mathcal{H}_1$ denote the graph defined by the edges $\mathcal{M}(B,x)\sim\mathcal{M}(C,y)$. Observe that the coordinates of $\mathcal{H}_1$ are independent, i.e., $\mathcal{H}_1 = \mathcal{H}_2^R$ for a graph $\mathcal{H}_2$ corresponding to each coordinate of $\mathcal{M}(B,x)$ and $\mathcal{M}(C,y)$. Therefore, the second eigenvalue of $\mathcal{H}_1$ is at most the second eigenvalue of $\mathcal{H}_2$. The Markov chain $\mathcal{H}_2$ on $(V\cup\{\perp\})\times(V_H\cup\perp)$ is defined as follows,

– Sample $a \in V$ and two neighbors $b,c\sim G_\eta(a)$.
– Sample $x,y \in V_H$ independently from μ_H.
– Output an edge between $\mathcal{M}(b,x)$ $\mathcal{M}(c,y)$.

Notice that in the Markov chain $\mathcal{H}_2$, for every choice of $\mathcal{M}(b,x)$ in $(V\cup\{\perp\})\times(V_H\cup\perp)$, with probability at least $\frac{1}{2}$, the other endpoint $\mathcal{M}(c,y) = (\perp,\perp)$. Therefore, the second eigenvalue of $\mathcal{H}_2$ is at most $\frac{1}{2}$, giving a bound of $\frac{1}{2}$ on the second eigen value of $\mathcal{H}$. ∎

The following lemma asserts that if the graph G is a NO-instance of Small-Set Expansion (γ, δ,M) then for almost all $A \in V^R$ the functions have no influential coordinates (see [LRV13] for the proof).

Lemma VI.5. *Fix $\delta = 1/R$. Suppose for all sets $S\subset V$ with* $\mathrm{vol}(S)\in(\delta/M, M\delta)$, $\Phi(S)\geqslant 1-\gamma$ *then for all $\tau>0$,*

$$\mathop{\mathbb{P}}_{A\sim V^R}[\exists i \mid \mathrm{Inf}_i[\Gamma_{1-\eta} g_A] \geqslant \tau] \leqslant \frac{1000}{\tau^3\varepsilon^2\eta^2}\cdot\max(1/M,\gamma)$$

Theorem VI.6. *(Soundness) For all ε, d there exists choice of M and γ,η such that the following holds. Suppose for all sets $S\subset V$ with* $\mathrm{vol}(S)\in(\delta/M, M\delta)$, $\Phi(S)\geqslant 1-\eta$, *then for all $F' : \mathcal{V}'\to[0,1]$ such that $\mathsf{Var}_1[F']\geqslant\frac{1}{10}$, we have* $\mathrm{val}_{\mathcal{P}'}(F')\geqslant\Omega(\sqrt{\varepsilon\log d})$

Proof: We will choose τ to small enough so that the error term in the soundness of dictatorship test (Proposition V.3) is smaller than ε ($\tau = \varepsilon^{1/\varepsilon^3}$ would suffice).

First, we know that if G is a NO-instance of Small-Set Expansion (γ,δ,M) then for almost all $A\in V^R$, the function g_A has no influential coordinates. Formally, by Lemma VI.5, we will have

$$\mathop{\mathbb{P}}_{A\sim V^R}[\exists i \mid \mathrm{Inf}_i[\Gamma_{1-\eta} g_A] \geqslant \tau] \leqslant \frac{1000}{\tau^3\eta^2}\cdot\max(1/M,\gamma)\,.$$

For an appropriate choice of M,γ, the above inequality implies that for all but an ε-fraction of vertices $A\in V^R$, the function g_A will have no influential coordinates.

Without loss of generality, we may assume that $\mathrm{val}_{\mathcal{P}'}(F')\leqslant\sqrt{\varepsilon\log d}$, else we would be done. Applying Lemma VI.4, we get that $\mathbb{E}_{A\in V^R}\mathsf{Var}_1[g_A]\geqslant(\mathsf{Var}_1[F])^2 - \mathrm{val}_{\mathcal{P}'}(F')\geqslant\frac{1}{200}$. This implies that for at least a $\frac{1}{400}$ fraction of $A\in V^R$, $\mathsf{Var}_1[g_A]\geqslant 1/400$. Hence for at least an $1/400-\varepsilon$ fraction of vertices $A\in V^R$ we have,

$$\mathsf{Var}_1[g_A]\geqslant\frac{1}{400}\qquad\text{and}\qquad\max_i \mathrm{Inf}_i(\Gamma_{1-\eta}(g_A))\leqslant\tau$$

By appealing to the soundness of the gadget (Proposition V.3), for every such vertex $A\in V^R$, $\mathrm{val}_{\mu_H^R}(g_A)\geqslant\Omega(\sqrt{\varepsilon\log d}) - O(\varepsilon) = \Omega(\sqrt{\varepsilon\log d})$. Finally, by applying Lemma VI.2, we get the desired conclusion. ∎

A. Putting it together

In order to finish the proof of Theorem I.3, we will need a fairly standard fairly standard reduction from d-Balanced Analytic Vertex Expansion to balanced symmetric vertex expansion (see [LRV13]).

Theorem VI.7. *A* c*-vs-*s *hardness for d-*Balanced Analytic Vertex Expansion *implies a* 4 c*-vs-*s/16 *hardness for balanced symmetric-vertex expansion on graphs of degree at most D, where* $D = \max\{100d/s, 2\log(1/c)\}$.

Proof of Theorem I.3: Follows from Theorem VI.1, Theorem VI.6, Theorem VI.7 and standard reductions from Balanced vertex expansion to vertex expansion and the computational equivalence of symmetric vertex expansion and vertex expansion. ∎

VII. An Optimal Algorithm for vertex expansion

In this section, we present an algorithm for approximating symmetric vertex expansion. To approximate vertex expansion, we appeal to a fairly easy reduction from symmetric vertex expansion to it which implies the following (see [LRV13]).

Algorithm VII.3.

– *Input :* A graph $G = (V, E)$
– *Output :* A set S with vertex expansion at most $72\sqrt{\mathsf{SDPval}\log d}$ (with constant probability).

1) Solve SDP 2 for graph G.
2) Pick a random Gaussian vector $g \sim N(0,1)^n$.
3) For each $i \in [n]$, define $x_i \stackrel{\text{def}}{=} \langle v_i, g\rangle$.
4) Sort the x_i's in decreasing order $x_{i_1} \geqslant x_{i_2} \geqslant \ldots x_{i_n}$. Let S_j denote the set of the first j vertices appearing in the sorted order. Let l be the index such that $l = \mathsf{argmin}_{1 \leqslant j \leqslant n/2} \Phi^{\mathsf{V}}(S_j)$.
5) Output the set corresponding to S_l in G.

Figure 3. Rounding Scheme

Theorem VII.1. *Given a graph G, there exists a graph G' such that* $\max_{i \in V(G)} \deg(i) = \max_{i \in V(G')} \deg(i)$ *and* $\phi^{\mathsf{V}}(G) = \Theta(\Phi^{\mathsf{V}}(G'))$. *Moreover, such a G' can be computed in time polynomial in the size of G.*

Our starting point is a the work of Bobkov et al. [BHT00] who define a spectral relaxation for vertex expansion in an undirected graph G.

$$\lambda_\infty \stackrel{\text{def}}{=} \min_x \frac{\sum_i \max_{j \sim i}(x_i - x_j)^2}{\sum_i x_i^2 - \frac{1}{n}(\sum_i x_i)^2}$$

The same work shows the following

Theorem VII.2 ([BHT00])**.** *For any unweighted, undirected graph G, we have* $\frac{\lambda_\infty}{2} \leqslant \phi^{\mathsf{V}} \leqslant \sqrt{2\lambda_\infty}$

Consider the following SDP relaxation for λ_∞.

SDP 2.

$$\mathsf{SDPval} \stackrel{\text{def}}{=} \min \sum_{i \in} \alpha_i$$

subject to:

$$\begin{aligned} \|v_j - v_i\|^2 &\leqslant \alpha_i \qquad \forall i \in V \text{ and } \forall j \sim i \\ \sum_i \|v_i\|^2 - \frac{1}{n}\|\sum_i v_i\|^2 &= 1 \end{aligned}$$

We present a simple randomized rounding of this SDP in Figure 3 and show that with constant probability it outputs a set with vertex expansion at most $O(\sqrt{\phi^{\mathsf{V}} \log d})$. We defer the details of the proof of Theorem I.2 to the full version of the paper [LRV13].

Acknowledgements.: Anand Louis and Santosh Vempala were supported in part by NSF awards CCF-0915903 and CCF-1217793. Prasad Raghavendra was supported by NSF Career Award CCF-1343104 and the Alfred P. Sloan Fellowship. The authors would like to thank Prasad Tetali for useful discussions and pointing us to the work of Bobkov et al. [BHT00].

References

[AKK+08] Sanjeev Arora, Subhash Khot, Alexandra Kolla, David Steurer, Madhur Tulsiani, and Nisheeth K. Vishnoi, *Unique games on expanding constraint graphs are easy: extended abstract*, STOC (Richard E. Ladner and Cynthia Dwork, eds.), ACM, 2008, pp. 21–28.

[Alo86] Noga Alon, *Eigenvalues and expanders*, Combinatorica **6** (1986), no. 2, 83–96.

[AM85] Noga Alon and V. D. Milman, λ_1*, isoperimetric inequalities for graphs, and superconcentrators*, J. Comb. Theory, Ser. B **38** (1985), no. 1, 73–88.

[AMS07] Christoph Ambühl, Monaldo Mastrolilli, and Ola Svensson, *Inapproximability results for sparsest cut, optimal linear arrangement, and precedence constrained scheduling*, FOCS, IEEE Computer Society, 2007, pp. 329–337.

[AR98] Yonatan Aumann and Yuval Rabani, *An $\mathcal{O}(\log k)$ approximate min-cut max-flow theorem and approximation algorithm*, SIAM J. Comput. **27** (1998), no. 1, 291–301.

[ARV04] Sanjeev Arora, Satish Rao, and Umesh V. Vazirani, *Expander flows, geometric embeddings and graph partitioning*, STOC (László Babai, ed.), ACM, 2004, pp. 222–231.

[BHT00] Sergey Bobkov, Christian Houdré, and Prasad Tetali, *λ_∞ vertex isoperimetry and concentration*, Combinatorica **20** (2000), no. 2, 153–172.

[FHL08] Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee, *Improved approximation algorithms for minimum weight vertex separators*, SIAM J. Comput. **38** (2008), no. 2, 629–657.

[IM12] Marcus Isaksson and Elchanan Mossel, *New maximally stable gaussian partitions with discrete applications*, To Appear in Israel Journal of Mathematics, 2012.

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica **15** (1995), no. 2, 215–245.

[LR99] Frank Thomson Leighton and Satish Rao, *Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms*, J. ACM **46** (1999), no. 6, 787–832.

[LRV13] Anand Louis, Prasad Raghavendra, and Santosh Vempala, *The complexity of approximating vertex expansion*, CoRR **abs/1304.3139** (2013).

[RS10] Prasad Raghavendra and David Steurer, *Graph expansion and the unique games conjecture*, STOC (Leonard J. Schulman, ed.), ACM, 2010, pp. 755–764.

[RST12] Prasad Raghavendra, David Steurer, and Madhur Tulsiani, *Reductions between expansion problems*, IEEE Conference on Computational Complexity, IEEE, 2012, pp. 64–73.

[ST12] David Steurer and Prasad Tetali, *Personal communication*.