

# Cuts in Cartesian Products of Graphs

Sushant Sachdeva \*

Madhur Tulsiani †

May 17, 2011

## Abstract

The  $k$ -fold Cartesian product of a graph  $G$  is defined as a graph on tuples  $(x_1, \dots, x_k)$  where two tuples are connected if they form an edge in one of the positions and are equal in the rest. Starting with  $G$  as a single edge gives  $G^{\square k}$  as a  $k$ -dimensional hypercube. We study the distributions of edges crossed by a cut in  $G^{\square k}$  across the copies of  $G$  in different positions. This is a generalization of the notion of *influences* for cuts on the hypercube.

We show the analogues of the statements of Kahn, Kalai and Linial [11] and that of Friedgut [8], for the setting of Cartesian products of arbitrary graphs. We also extend the work on studying isoperimetric constants for these graphs [3] to the value of semidefinite relaxations for expansion. We connect the optimal values of the relaxations for computing expansion, given by various semidefinite hierarchies, for  $G$  and  $G^{\square k}$ .

## 1 Introduction

The Cartesian product of two graphs  $G$  and  $H$  is defined as the graph  $G \square H$  on the vertex set  $V(G) \times V(H)$ , with the tuples  $\{(i_1, i_2), (j_1, j_2)\}$  forming an edge if  $\{i_1, j_1\}$  forms an edge in  $G$  and  $i_2 = j_2$ , or  $\{i_2, j_2\}$  forms an edge in  $H$  and  $i_1 = j_1$ . The  $k^{\text{th}}$  power of a graph according to this product is defined by associativity as  $G^{\square k} = G \square G^{\square(k-1)}$ . The notion is a well known and well-studied one in graph theory (see [10]) for example.

Certain special cases of this product are particularly interesting to consider. For example, when  $G$  is just an edge,  $G^{\square k}$  is the  $k$ -dimensional hypercube on  $2^k$  vertices, which is perhaps the best known example. Starting with  $G$  as the  $n$ -cycle,  $G^{\square k}$  gives the  $k$ -dimensional torus. Similarly, starting from a path yields a grid.

Isoperimetric questions for Cartesian products of graphs have been studied by various authors. Chung and Tetali [3] computed the conductance for the Cartesian products in terms of those of the starting graph (see references in [3] for previous work on special cases). Houdré and Tetali [9] compute various isoperimetric invariants for the extension of this notion to the case of Markov chains, where the product of two Markov chains is defined by a process which randomly selects one of the two chains and make a transition according to it.

We study the question of extending some of the isoperimetric inequalities known for the case of the hypercube in terms of the *influences* of variables, to the case of Cartesian products of graphs. For a Boolean function on the  $k$ -dimensional hypercube, the influence of the function in the  $i^{\text{th}}$

---

\*Department of Computer Science, Princeton University. [sachdeva@cs.princeton.edu](mailto:sachdeva@cs.princeton.edu)

†Department of Computer Science, Princeton University. [tulsiani@cs.princeton.edu](mailto:tulsiani@cs.princeton.edu)

coordinate is defined as the probability over a random input that changing the  $i^{\text{th}}$  bit changes the value of the function. The total influence of the function is the sum of the influences along all the  $k$  coordinates. Viewing a Boolean function as cut on the hypercube, the influence along the  $i^{\text{th}}$  coordinate is simply the fraction of the edges along the the  $i^{\text{th}}$  direction (which correspond to changing the  $i^{\text{th}}$  bit) that are crossed by the cut. The total influence corresponds (after scaling) to the total number of edges crossed. Stated in this way, both the above definitions have obvious extensions to the  $k$ -fold Cartesian product of an arbitrary graph.

We consider the theorems of Kahn, Kalai and Linial [11] and that of Friedgut [8], which are proved via Fourier analysis for the hypercube, and generalize them to the case of Cartesian products of arbitrary graphs. We also consider applications of these products to integrality gaps for linear and semidefinite programming relaxations.

**The KKL theorem** The theorem of Kahn, Kalai and Linial [11], which introduced various tools in discrete Fourier analysis to Computer Science, states that for the  $k$ -dimensional hypercube, a Boolean function with variance  $v$  has influence at least  $\Omega(v \log k/k)$  along some coordinate. This result has been generalized in many ways. In particular, Bourgain et al. [2] considered the case when the function is not Boolean but takes values in  $[0, 1]$ . Also, O’Donnell and Wimmer [12] obtained a KKL theorem for some special classes of Schreier graphs, which may not necessarily have a product structure.

We consider a different generalization when we still have product graph  $G^{\square k}$ , but the influence in the  $i^{\text{th}}$  coordinate is defined as the probability the function changes (i.e. one lands on the other side of the of the cut) when taking a random step according to  $G$  in the  $i^{\text{th}}$  direction. We show that for a Boolean function with variance  $v$ , there is at least one coordinate with influence  $\Omega(\alpha \cdot v \cdot (\log k)/k)$ . Here,  $\alpha$  is the log-Sobolev constant of the graph  $G$ , which is a certain isoperimetric constant related to the mixing time for a random walk on the graph. We also discuss the tightness of these results in Appendix A. A similar KKL theorem for this class of graphs was also obtained independently by Cordero-Erausquin and Ledoux [4].

**Friedgut’s junta theorem** Friedgut’s junta theorem states that if a Boolean function  $f$  with variance  $\Omega(1)$  has total influence  $\mathbb{I}$ , then there exists a function  $g$  depending only on  $\exp(O(\mathbb{I}/\epsilon))$  coordinates such that  $\|f - g\|^2 \leq \epsilon$ . We show an analogous statement for the case of Cartesian product of a graph  $G$ , where for an  $f$  as above, we obtain a  $g$  depending only on  $\exp((\mathbb{I}/\alpha\epsilon))$  coordinates. As before,  $\alpha$  is the log-Sobolev constant of  $G$ .

For both the KKL theorem and Friedgut’s theorem, the original proofs involve a use of the hypercontractive inequality. We consider (and extend) different proofs for these theorem given by Rossignol [13] and by Falik and Samorodnitsky [7]. The proofs are based on the log-Sobolev inequality for the hypercube.

**Applications to integrality gaps** It follows from the results of Chung and Tetali [3] that if the starting graph  $G$  has expansion  $h$ , then the product  $G^{\square k}$  has expansion  $h/k$ . The same also holds for the spectral gap of  $G$  and  $G^{\square k}$ , which is also the optimum of the basic semidefinite program (SDP) for Sparsest Cut. This immediately implies that if one has a finite instance with integrality gap  $K$  for the basic SDP, then using Cartesian products it gives an infinite family of arbitrarily large instances with the same gap.

We show that above is also the case for various hierarchies of linear and semidefinite relaxations for

**Sparsest Cut.** In particular, if the optimum of such a relaxation obtained by  $r$  levels is  $\text{Opt}$  for  $G$ , then it is  $\text{Opt}/k$  for  $G^{\square k}$ . Most ways of increasing the size of a graph seem to alter the expansion of the graph. However, because of the above observation, Cartesian products provide the right way of “padding” integrality gap instances to arbitrarily large size while preserving the gap.

## 2 Preliminaries and Notation

**Definition 2.1 (Cartesian product)** *Given two graphs  $G$  and  $H$ , their Cartesian product  $G \square H$  is defined as a graph with the vertex set  $V(G) \times V(H)$  and the following set of edges,*

$$E(G \square H) = \left\{ ((i_1, i_2), (j_1, j_2)) \mid [(i_1, j_1) \in E(G) \wedge (i_2 = j_2)] \vee [(i_1 = j_1) \wedge ((i_2, j_2) \in E(H))] \right\} .$$

For a graph  $G$ , we define  $G^{\square 1} = G$  and  $G^{\square k} = G^{\square(k-1)} \square G$ .

The vertices of the graph  $G^{\square k}$  are  $V(G)^k$ . For a vertex  $x \in V(G^{\square k})$ , we will use the notation  $x = (x_1, \dots, x_k)$ . Note that we can think of each edge in  $G^{\square k}$  to be along a *coordinate*  $j$ , e.g. if  $(x, y)$  is an edge,  $x_i = y_i$  for all  $i \neq j$ , we say  $(x, y)$  is an edge along coordinate  $j$ . We denote the set of all edges along coordinate  $j$  as  $E_j$ .

We will mostly work with the case when  $G$ , and hence also,  $G^{\square k}$  is regular for all  $k$ . The results in the paper can be extended to the case of general undirected graphs by defining the notions more carefully. We provide the details in Appendix C.

Let  $\pi : V(G) \rightarrow \mathbb{R}$  be the uniform distribution on  $V(G)$ . We define the inner product and norms for the space of function  $V(G) \rightarrow \mathbb{R}$  as usual:

$$\langle f, g \rangle = \mathbf{E}_{x \sim \pi} [f(x)g(x)] , \quad \|f\|_2^2 = \langle f, f \rangle = \mathbf{E}_x [f(x)^2] \quad \text{and} \quad \|f\|_1 = \mathbf{E}_x [|f(x)|] .$$

The variance of function  $f : V(G) \rightarrow \mathbb{R}$  is defined with respect to the same distribution, as  $\text{Var}(f) = \mathbf{E}_{x \sim \pi} [f(x)^2] - (\mathbf{E}_{x \sim \pi} [f(x)])^2$ . Let the (non-negative) weights on the edges,  $w_e$  for  $e \in E(G)$ , sum up to 1. Let  $\mathbf{L}_G$  be the normalized Laplacian for the graph  $G$  ( $(\mathbf{L}_G)_{ii} = \pi(i)$  and  $(\mathbf{L}_G)_{ij} = -w_{ij}/2$ ). We have,

$$\langle f, \mathbf{L}_G f \rangle = \frac{1}{2} \mathbf{E}_{(x,y) \in E(G)} (f(x) - f(y))^2 .$$

For the graph  $G^{\square k}$ , we also define the directional Laplacian  $\mathbf{L}_j$  which only considers edges along the  $j^{\text{th}}$  coordinate,

$$\mathbf{L}_j = \mathbf{I} \otimes \mathbf{I} \otimes \dots \otimes \mathbf{L}_G \otimes \dots \otimes \mathbf{I},$$

where the matrix  $\mathbf{L}_G$  is in the  $j^{\text{th}}$  position. It is easy to check that the Laplacian for  $G^{\square k}$  is  $\mathbf{L}_{G^{\square k}} = \frac{1}{k} \sum_j \mathbf{L}_j = \mathbf{E}_j \mathbf{L}_j$ . The directional Laplacian also gives the definition of influence for Cartesian products.

**Definition 2.2 (Influence)** *For a Boolean function  $f : V(G^{\square k}) \rightarrow \{0, 1\}$ , we define its influence along the  $j^{\text{th}}$  coordinate as the quantity  $\langle f, \mathbf{L}_j f \rangle$ .*

We also define the variance of the function along the  $j^{\text{th}}$  coordinate as below. Here  $x \setminus \{x_j\}$  denotes the tuple  $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k)$ .

**Definition 2.3** For a function  $f : V(G^{\square k}) \rightarrow \mathbb{R}$ , its variance along the  $j^{\text{th}}$  coordinate is defined as  $\mathbf{Var}_j(f) = \mathbf{E}_{x \setminus \{x_j\}} [\mathbf{E}_{x_j} [f(x)^2] - (\mathbf{E}_{x_j} [f(x)])^2]$ .

Notice that for the hypercube the two definitions are the same. However, for a general graph  $G$ , the variance does not depend of the structure of the graph, while our notion of influence does.

## 2.1 Isoperimetric Constants of a graph

**Conductance** Given a set  $S \subseteq V(G)$ , we define the volume of the set,  $\text{Vol}(S)$  to be the fraction of the vertices contained in  $S$  i.e.  $\text{Vol}(S) = |S|/|V(G)|$ . We define the Conductance of a graph  $\Phi(G)$ , as follows

$$\Phi(G) = \min_S \frac{1}{4} \frac{|E(S, \bar{S})|}{|E|} \frac{1}{\text{vol}(S)\text{vol}(\bar{S})}$$

The factor of  $1/4$  ensures that  $\Phi(G) \leq 1$ . If we consider the the  $\{-1, 1\}$ -valued indicator function of a set  $S$ , we get an equivalent definition of  $\Phi(G)$  as follows,

$$\Phi(G) = \min_{f:V(G) \rightarrow \{-1,1\}} \frac{\langle f, \mathbf{L}_G f \rangle}{2 \mathbf{Var}(f)}.$$

This implies that if  $f$  is any  $\{-1, 1\}$ -valued function,

$$\langle f, \mathbf{L}_G f \rangle \geq 2 \cdot \Phi(G) \cdot \mathbf{Var}(f).$$

Note that this is also true if  $f$  is a  $\{0, 1\}$ -valued indicator function for a set.

**Eigenfunctions and Eigenvalues** Let  $v_0 = \mathbf{1}, v_1, \dots, v_{n-1}$  be the eigenfunctions for  $\mathbf{L}_G$  corresponding to eigenvalues  $\lambda_0 = 0 \leq \lambda_1 \leq \dots \leq \lambda_{n-1}$ . It is well known (and easy to verify) that the eigenfunctions of  $G^{\square k}$  are tensor products of the eigenfunctions of  $G$ .

**Claim 2.4** The eigenfunctions for  $G^{\square k}$  are  $v_{i_1} \otimes v_{i_2} \otimes \dots \otimes v_{i_k}$  with each  $i_j \in \{0, \dots, n-1\}$ .

Letting  $(i)$  denote the sequence  $(i_1, \dots, i_k)$ , we will denote the eigenfunction  $v_{i_1} \otimes v_{i_2} \otimes \dots \otimes v_{i_k}$  by  $v_{(i)}$ . We always think of eigenfunctions as having unit norm i.e.  $\|v_i\|^2 = 1$ . The following claim is easy to verify using the fact that  $\mathbf{L}_{G^{\square k}} = \mathbf{E}_j \mathbf{L}_j$ .

**Claim 2.5 (Eigenvalues, Eigenfunctions)**  $v_{(i)}$  is an eigenfunctions of  $\mathbf{L}_{G^{\square k}}$  with eigenvalue  $\mathbf{E}_j \lambda_{i_j}$ .

**Log-Sobolev Constant** For a function  $f : V(G) \rightarrow \mathbb{R}$ , we can define the entropy of the function as follows,

$$\begin{aligned} \mathbf{Ent}(f^2) &= \mathbf{E}[f(x)^2 \log f(x)^2] - (\mathbf{E}[f(x)^2]) \log \mathbf{E}[f(x)^2] \\ &= \mathbf{E}[f(x)^2 \log f(x)^2] - \|f\|_2^2 \log \|f\|_2^2 \end{aligned}$$

where  $\log$  is the natural logarithm.

**Definition 2.6 (Log-Sobolev Constant)** The Log-Sobolev constant of a graph  $G$  is defined to be the largest constant  $\alpha$  such that the following inequality holds for all functions  $f : V(G) \rightarrow \mathbb{R}$ ,

$$\langle f, \mathbf{L}_G f \rangle = \frac{1}{2} \mathbf{E}_{(x,y) \in E(G)} (f(x) - f(y))^2 \geq \frac{\alpha}{2} \cdot \mathbf{Ent}(f^2)$$

The following claim relates the log-Sobolev constant of  $G^{\square k}$  to that of  $G$ .

**Claim 2.7 (Diaconis et al. [6])** *Let  $\alpha$  be the Log-Sobolev constant for a graph  $G$ , then the Log-Sobolev constant for  $G^{\square k}$  is  $\alpha/k$ .*

It is known that the isoperimetric constants defined above satisfy the following inequalities between them (see [6] for example).

$$\alpha \leq \lambda_1 \leq 2\Phi$$

### 3 KKL Theorem and Friedgut's Junta Theorem

In this section, we shall prove the analogues of the theorems of Kahn, Kalai and Linial [11] and Friedgut [8] for Cartesian products of graphs. Both these theorems analyze cuts in the hypercube which is simply the Cartesian product of an edge. The proofs of both theorems proceed by using hypercontractivity of the Bonami-Beckner noise operator on the hypercube.

While the noise operator can be easily generalized to the setting of Cartesian products, the hypercontractivity based proofs do not seem to extend easily to the setting of Cartesian products of general graphs. Instead we develop on Rossignol's proof of the KKL theorem [13], which is based on the log-Sobolev inequality. For the KKL and Friedgut theorems on the hypercube, proofs using the log-Sobolev inequality were also given by Falik and Samorodnitsky [7].

To prove both the theorems, we shall need some preparatory lemmas. We develop these below. The manipulations are similar to those in [13].

Let  $f : V(G^{\square k}) \rightarrow \{-1, 1\}$  define a cut in the graph  $G^{\square k}$ . We first define an operator that measures the variance of  $f$  in the  $j^{\text{th}}$  direction. Let

$$\mathbf{K}_j = \mathbf{I} \otimes \mathbf{I} \otimes \dots \otimes \left( \mathbf{I} - \frac{1}{n} \mathbf{J} \right) \otimes \dots \otimes \mathbf{I}$$

where  $\mathbf{J}$  is the  $n \times n$  matrix with all ones and the matrix  $\mathbf{I} - \frac{1}{n} \mathbf{J}$  is in the  $j^{\text{th}}$  position. One can then write the variance of the function along the  $j^{\text{th}}$  coordinate as  $\mathbf{Var}_j(f) = \langle f, \mathbf{K}_j f \rangle$ .

Let  $f$  be decomposed in terms of the eigenfunctions of  $G^{\square k}$  as  $f = \sum_{(i)} \hat{f}_{(i)} v_{(i)}$  where  $\hat{f}_{(i)} = \langle f, v_{(i)} \rangle$ . Define the functions  $f_j$  as follows:

$$f_j = \sum_{(i): i_j \neq 0, i_l = 0 \forall l > j} \hat{f}_{(i)} v_{(i)} = \mathbf{E}_{x_{j+1}, \dots, x_k} [\mathbf{K}_j f] .$$

Observe that the functions  $f_j$  are orthogonal, i.e.  $\langle f_{j_1}, f_{j_2} \rangle = 0$  for  $j_1 \neq j_2$ .

We first bound the  $\ell_2$  norm of the functions  $f_j$ .

**Claim 3.1** *The  $\ell_2$  norm of the function  $f_j$  is bounded by  $\mathbf{Var}_j(f)$ .*

$$\|f_j\|_2^2 \leq \mathbf{Var}_j(f) = \langle f, \mathbf{K}_j f \rangle .$$

We now prove a simple lemma which bounds the  $\ell_1$  norm of the functions  $f_j$ .

**Lemma 3.2**

$$\|f_j\|_1 = \mathbf{E}_x [|f_j(x)|] \leq \mathbf{Var}_j(f) .$$

**Proof:** We first use triangle inequality to get an upper bound on  $|f_j(x)|$

$$\begin{aligned}
\mathbf{E}_x [|f_j(x)|] &= \mathbf{E}_x \left[ \left| \mathbf{E}_{x_{j+1}, \dots, x_k} [\mathbf{K}_j f(x)] \right| \right] \leq \mathbf{E}_x [|\mathbf{K}_j f(x)|] \\
&= \mathbf{E}_x \left[ \left| f(x) - \mathbf{E}_{y : y_i = x_i \forall i \neq j} [f(y)] \right| \right] \\
&= \mathbf{E}_x \left[ f(x) \cdot \left( f(x) - \mathbf{E}_{y : y_i = x_i \forall i \neq j} [f(y)] \right) \right] \\
&= \langle f, \mathbf{K}_j f \rangle = \mathbf{Var}_j(f)
\end{aligned}$$

where we used the observation that the the sign of  $f(x) - \mathbf{E}_y f(y)$  is the same as  $f(x)$ , since  $f(x) \in \{-1, 1\}$ .  $\blacksquare$

We shall apply the Log-Sobolev inequality to the functions  $f_j$  defined above. However, the entropy of these functions is somewhat difficult to work with. The following lemma gives a different estimate in terms of the  $\ell_1$  and  $\ell_2$  norms of the functions we are applying the Log-Sobolev inequality to.

**Lemma 3.3** *For any  $t \leq 1/e^2$  and  $h : V(G^{\square k}) \rightarrow \mathbb{R}$ ,*

$$\langle h, \mathbf{L}_{G^{\square k}} h \rangle \geq \frac{\alpha}{2k} \cdot \left( \sqrt{t} \log t \cdot \|h\|_1 + \log t \cdot \|h\|_2^2 - \|h\|_2^2 \log \|h\|_2^2 \right)$$

**Proof:** Applying the log-Sobolev inequality for  $G^{\square k}$  to  $h$ ,

$$\langle h, L_{G^{\square k}} h \rangle \geq \frac{\alpha}{2k} \mathbf{Ent}(h^2) = \frac{\alpha}{2k} \cdot \left( \mathbf{E}_x [h(x)^2 \log h(x)^2] - \|h\|_2^2 \log \|h\|_2^2 \right)$$

Observe that the function  $\sqrt{z} \log z$  is decreasing in  $[0, t]$  if  $t \leq 1/e^2$ . We use this to bound the first term as below.

$$\begin{aligned}
\mathbf{E}_x [h^2(x) \log h(x)^2] &= \mathbf{E}_x [h^2(x) \log h(x)^2 \cdot \mathbf{1}_{h^2 \leq t}] + \mathbf{E}_x [h^2(x) \log h(x)^2 \cdot \mathbf{1}_{h^2 > t}] \\
&\geq \mathbf{E}_x [ |h(x)| \cdot \sqrt{h(x)^2} \log |h(x)|^2 \cdot \mathbf{1}_{h^2 \leq t}] + \mathbf{E}_x [h^2(x) \cdot \log t \cdot \mathbf{1}_{h^2 > t}] \\
&\geq \mathbf{E}_x [ |h(x)| \cdot \sqrt{t} \log t \cdot \mathbf{1}_{h^2 \leq t}] + \mathbf{E}_x [h^2(x) \cdot \log t \cdot \mathbf{1}_{h^2 > t}] \\
&\geq \sqrt{t} \log t \cdot \mathbf{E}_x [|h(x)|] + \log t \cdot \mathbf{E}_x [h^2(x)] \\
&= \sqrt{t} \log t \cdot \|h\|_1 + \log t \cdot \|h\|_2^2
\end{aligned}$$

where the last two inequalities used the fact that  $\log t < 0$  for  $t \leq 1/e^2$ . Plugging the above bound in the log-Sobolev inequality proves the claim.  $\blacksquare$

Combining the above two lemmas gives the following corollary, which shall be useful in the proofs of both the theorems.

**Corollary 3.4** *For all  $t \leq \frac{1}{e^2}$  and all  $j \in \{0, \dots, k-1\}$ ,*

$$\langle f_j, \mathbf{L}_{G^{\square k}} f_j \rangle \geq \frac{\alpha}{2k} \cdot \left( \sqrt{t} \log t \cdot \mathbf{Var}_j(f) + \log t \cdot \|f_j\|_2^2 - \|f_j\|_2^2 \log \|f_j\|_2^2 \right)$$

### 3.1 Deriving the KKL Theorem for Cartesian products

We now prove an analogue of the KKL theorem, which says that a cut with high variance, must have a somewhat large number of edges crossing in at least one direction. Note that our bound has an extra factor of  $\alpha$ , which is 2 for the case of the hypercube when  $G$  is an edge.

**Theorem 3.5 (KKL Theorem)** *Given  $f : V(G^{\square k}) \rightarrow \{-1, 1\}$ , we have*

$$\max_j \langle f, \mathbf{L}_j f \rangle \geq \Omega \left( \frac{\log k}{k} \cdot \alpha(G) \cdot \mathbf{Var}(f) \right)$$

**Proof:** We add the result from Corollary 3.4 for all  $j$  and observe that  $\sum_j \|f_j\|_2^2 = \mathbf{Var}(f)$  and  $\sum_j \langle f_j, \mathbf{L}_{G^{\square k}} f_j \rangle = \langle f, \mathbf{L}_{G^{\square k}}(f) \rangle$ . Thus, for any  $t \leq \frac{1}{e^2}$ , we get

$$\langle f, \mathbf{L}_{G^{\square k}} f \rangle \geq \frac{\alpha}{2k} \left( \sqrt{t} \log t \cdot \sum_j \mathbf{Var}_j(f) + \log t \cdot \mathbf{Var}(f) - \sum_j \|f_j\|_2^2 \log \|f_j\|_2^2 \right)$$

Let  $V = \max_j \mathbf{Var}_j(f)$ . By subadditivity of variance,  $V \leq \mathbf{Var}(f) \leq kV$ . We can upper bound  $\log \|f_j\|_2^2$  by  $\log V$  and  $\sum_j \mathbf{Var}_j(f)$  by  $kV$  (since  $\log t$  is negative).

$$\langle f, \mathbf{L}_{G^{\square k}} f \rangle \geq \frac{\alpha}{2k} \left( \sqrt{t} \log t \cdot kV + \log t \cdot \mathbf{Var}(f) - \mathbf{Var}(f) \log V \right)$$

In order to balance the above two expressions and still have  $t \leq \frac{1}{e^2}$ , we can pick  $t = \left( \frac{\mathbf{Var}(f)}{ekV} \right)^2$ . This gives,

$$\langle f, \mathbf{L}_{G^{\square k}} f \rangle \geq \frac{\alpha \mathbf{Var}(f)}{2k} \cdot \left( 2 \left( 1 + \frac{1}{e} \right) \log \left( \frac{\mathbf{Var}(f)}{ekV} \right) + \log \frac{1}{V} \right)$$

Suppose that  $V \geq \frac{\log k}{k} \cdot \mathbf{Var}(f)$ . Then using the fact that  $f(x_1, \dots, x_{j-1}, \cdot, x_{j+1}, \dots, x_k)$  is a Boolean function on  $G$ , we get that

$$\max_j \langle f, \mathbf{L}_j f \rangle \geq 2\Phi(G) \cdot \max_j \langle f, \mathbf{K}_j f \rangle = \Omega \left( \frac{\log k}{k} \cdot \Phi(G) \cdot \mathbf{Var} f \right).$$

Also, for the case when  $V \leq \frac{\log k}{k} \cdot \mathbf{Var} f$ , we have

$$\begin{aligned} \langle f, \mathbf{L}_{G^{\square k}} f \rangle &\geq \frac{\alpha \mathbf{Var}(f)}{2k} \cdot \left( 2 \left( 1 + \frac{1}{e} \right) \log \left( \frac{1}{e \log k} \right) + \log \left( \frac{k}{\mathbf{Var}(f) \log k} \right) \right) \\ &= \Omega \left( \frac{\log k}{k} \cdot \alpha \cdot \mathbf{Var}(f) \right) \end{aligned}$$

Combining the above with the facts that  $2\Phi \geq \alpha$  and  $\max_j \langle f, \mathbf{L}_j f \rangle \geq \mathbf{E}_j \langle f, \mathbf{L}_j f \rangle$  then proves the result.  $\blacksquare$

### 3.2 Friedgut's theorem for Cartesian products of graphs

We now prove Friedgut's theorem which says that a cut which is crossed by very few edges is close to a cut that depends only on few coordinates. As before, substituting  $\alpha = 2$  gives the version for the hypercube.

**Theorem 3.6 (Friedgut's Junta Theorem)** *Given  $f : V(G^{\square k}) \rightarrow \{-1, 1\}$ ,  $f$  is  $\epsilon$ -close to a boolean function  $g : V(G^{\square k}) \rightarrow \{-1, 1\}$ , i.e.  $\|f - g\|_2^2 \leq \epsilon$ , which is determined only by the value of  $l$  coordinates, where,*

$$l \leq \exp\left(\frac{50k}{\epsilon\alpha} \cdot \langle f, \mathbf{L}_{G^{\square k}} f \rangle\right)$$

One interesting way to interpret the generalization of Friedgut's Junta theorem is as follows: If for the graph  $G$ ,  $\Phi$  and  $\alpha$  are within a constant factor (as a function of  $k$ ), then a balanced cut in  $G^{\square k}$  (described by  $f$ ) that is close to the sparsest cut in the graph  $(\Phi/k)$ , must be  $\epsilon$ -close to a cut that is determined only by a constant number of coordinates.

**Proof:** We order the coordinates  $j$  are so that  $\mathbf{Var}_j(f)$  is non-increasing. Let  $J = \{1, \dots, l\}$  be the subset of all coordinates  $j$  that have  $\mathbf{Var}_j(f)$  at least  $V$  (where  $V$  is some threshold we will pick later). Let  $f = \sum_{(i)} \hat{f}_{(i)} v_{(i)}$ . We define the function  $g : V(G^{\square k}) \rightarrow \mathbb{R}$  as follows,

$$g(x) = \sum_{(i): i_j=0 \forall j \notin J} \hat{f}_{(i)} v_{(i)}$$

Clearly,  $g$  only depends on coordinates in  $J$ . We shall show that for an appropriate choice of  $V$ ,  $\|f - g\|_2^2$  is small i.e.  $\|f - g\|_2^2 \leq \epsilon$ . It will also follow from our choice of  $V$  that the number of coordinates in  $J$  is as claimed.

Adding the result from Corollary 3.4 for all  $j \notin J$  and observing  $\sum_{j \notin J} \langle f_j, f_j \rangle = \|f - g\|_2^2$  gives the following for any  $t \leq \frac{1}{e^2}$ .

$$\sum_{j \notin J} \langle f_j, \mathbf{L}_{G^{\square k}} f_j \rangle \geq \frac{\alpha}{2k} \cdot \left( \sqrt{t} \log t \sum_{j \notin J} \mathbf{Var}_j(f) + \log t \cdot \|f - g\|_2^2 - \sum_{j \notin J} \|f_j\|_2^2 \log \|f_j\|_2^2 \right) \quad (1)$$

Note that the functions  $f_j$  partition  $f$  into orthogonal eigenspaces of  $\mathbf{L}_{G^{\square k}}$ . Using this and the fact that  $\mathbf{L}_{G^{\square k}}$  is positive semidefinite, we can bound the LHS as

$$\sum_{j \notin J} \langle f_j, \mathbf{L}_{G^{\square k}} f_j \rangle \leq \langle f, \mathbf{L}_{G^{\square k}} f \rangle .$$

Once again, we shall use that fixing a value for all other coordinates,  $f$  is a Boolean function on the  $j^{\text{th}}$  input, and hence  $\langle f, \mathbf{L}_j f \rangle \geq 2\Phi \cdot \mathbf{Var}_j(f)$ . This gives a bound on  $\sum_{j \notin J} \mathbf{Var}_j(f)$  in terms of  $\langle f, \mathbf{L}_{G^{\square k}} f \rangle$ .

$$\sum_{j \notin J} \mathbf{Var}_j(f) \leq \sum_j \mathbf{Var}_j(f) \leq \frac{k}{2\Phi} \cdot \left\langle f, \sum_j \mathbf{L}_j f \right\rangle = \frac{k}{2\Phi} \cdot \langle f, \mathbf{L}_{G^{\square k}} f \rangle$$

Finally, we can bound the last term in equation (1) using the fact that  $\mathbf{Var}_j(f) \leq V$  for each  $j \notin J$ .

$$\sum_{j \notin J} \|f_j\|_2^2 \log \|f_j\|_2^2 \leq \sum_{j \notin J} \|f_j\|_2^2 \log \mathbf{Var}_j(f) \leq \|f - g\|_2^2 \log V$$

We now plug in the above bounds into equation (1). Remember that  $\log t < 0$ . Thus, we get for any  $t \leq \frac{1}{e^2}$ ,

$$\langle f, \mathbf{L}_{G^{\square k}} f \rangle \geq \frac{\alpha}{2k} \cdot \left( \frac{k}{2\Phi} \sqrt{t} \log t \cdot \langle f, \mathbf{L}_{G^{\square k}} f \rangle + \log t \cdot \|f - g\|_2^2 - \|f - g\|_2^2 \log V \right) .$$

Again, in order to balance the terms and still have  $t \leq \frac{1}{e^2}$ , we choose  $t = \left( \frac{2\Phi \cdot \|f-g\|^2}{ek \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle} \right)^2$ . For this value of  $t$ , the above bound simplifies to

$$\begin{aligned} \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle &\geq \frac{\alpha}{2k} \cdot \|f-g\|^2 \cdot \left( 2 \left( 1 + \frac{1}{e} \right) \log \left( \frac{2\Phi \cdot \|f-g\|^2}{ek \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle} \right) - \log V \right) \\ \implies \log V &\geq 2 \left( 1 + \frac{1}{e} \right) \log \left( \frac{2\Phi \cdot \|f-g\|^2}{ek \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle} \right) - \frac{2k \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{\alpha \cdot \|f-g\|^2}. \end{aligned}$$

Consider the RHS as a function in  $\|f-g\|^2$ , say  $F(\|f-g\|^2)$  and notice that  $F$  is an increasing function. Thus, choosing  $\log V = F(\epsilon)$  and hence  $V = \exp(F(\epsilon))$  would imply that  $\|f-g\|^2 \leq \epsilon$ . It only remains to show that the size of the set  $J$  is bounded as claimed for this choice of  $V$ .

Since  $J$  was defined to be the set of coordinates  $j$  with  $\mathbf{Var}_j(f) \geq V$ , the size of  $J$  is at most  $(\sum_j \mathbf{Var}_j(f))/V = (\sum_j \mathbf{Var}_j(f)) \cdot \exp(-F(\epsilon))$ . Using the bounds  $\sum_j \mathbf{Var}_j(f) \leq (k/2\Phi) \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle$ , and  $\alpha \leq 2\Phi$  we can bound this as

$$\begin{aligned} |J| &\leq \left( \sum_j \mathbf{Var}_j(f) \right) \exp \left( \frac{2k}{\epsilon\alpha} \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle + 2 \left( 1 + \frac{1}{e} \right) \cdot \log \left( \frac{ek \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{2\Phi\epsilon} \right) \right) \\ &\leq \exp \left( \log \left( \frac{2k \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{\Phi} \right) + \frac{2k}{\epsilon\alpha} \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle + 2 \left( 1 + \frac{1}{e} \right) \cdot \log \left( \frac{ek \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{2\Phi\epsilon} \right) \right) \\ &\leq \exp \left( \frac{2k \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{\epsilon\alpha} + \left( 3 + \frac{2}{e} \right) \log \left( \frac{2k \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle}{\Phi\epsilon} \right) \right) \\ &\leq \exp \left( \frac{12k}{\epsilon\alpha} \cdot \langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle \right). \end{aligned}$$

Thus we have a function  $g : V(G^{\square k}) \rightarrow \mathbb{R}$  that depends only on a few coordinates and is close to  $f$ . Now, we prove that the boolean function obtained by taking the sign of  $g$ , denoted  $\tilde{g} = \text{sgn}(g) : V(G^{\square k}) \rightarrow \{-1, 1\}$  also satisfies  $\|f-g\|^2 = O(\epsilon)$ . To see this, observe that whenever  $f(x) \neq \tilde{g}(x)$ ,  $|f(x) - \tilde{g}(x)|^2 = 4$  but  $|f(x) - g(x)|^2$  must be at least 1. Hence  $\|f - \tilde{g}\|^2 \leq 4\epsilon$ . Replacing  $\epsilon$  by  $\epsilon/4$  proves the theorem.  $\blacksquare$

## Acknowledgements

We thank Ryan O'Donnell for helpful comments and for pointing us to the manuscript [4].

## References

- [1] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56:5:1–5:37, April 2009.
- [2] Jean Bourgain, Jeff Kahn, Gil Kalai, Yitzhak Katznelson, and Nathan Linial. The influence of variables in product spaces. *Israel Journal of Mathematics*, 77:55–64, 1992. 10.1007/BF02808010.
- [3] F. R. K. Chung and Prasad Tetali. Isoperimetric inequalities for cartesian products of graphs. *Comb. Probab. Comput.*, 7:141–148, June 1998.

- [4] D. Cordero-Erausquin and M. Ledoux. Hypercontractive measures, Talagrand’s inequality and influences. Manuscript, <http://www.math.univ-toulouse.fr/~ledoux/influence.pdf>, 2011.
- [5] Nikhil R. Devanur, Subhash A. Khot, Rishi Saket, and Nisheeth K. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, STOC ’06, pages 537–546, New York, NY, USA, 2006. ACM.
- [6] P. Diaconis and L. Saloff-Coste. Logarithmic sobolev inequalities for finite markov chains. *The Annals of Applied Probability*, 6(3):pp. 695–750, 1996.
- [7] Dvir Falik and Alex Samorodnitsky. Edge-isoperimetric inequalities and influences. *Combinatorics, Probability and Computing*, 16(05):693–712, 2007.
- [8] Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18:27–35, 1998. 10.1007/PL00009809.
- [9] C. Houdré and P. Tetali. Isoperimetric invariants for product markov chains and graph products. *Combinatorica*, 24:359–388, July 2004.
- [10] Wilfried Imrich, Sandi Klavzar, and Douglas F. Rall. *The Cartesian Product of Graphs*. A K Peters, Wellesley, MA, 2008.
- [11] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science*, pages 68–80, Washington, DC, USA, 1988. IEEE Computer Society.
- [12] Ryan O’Donnell and Karl Wimmer. Kkl, kruskal-katona, and monotone nets. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS ’09, pages 725–734, Washington, DC, USA, 2009. IEEE Computer Society.
- [13] Raphaël Rossignol. Threshold for monotone symmetric properties through a logarithmic sobolev inequality. *The Annals of Probability*, 34(5):pp. 1707–1725, 2006.
- [14] J. Michael Steele. An Efron-Stein inequality for nonsymmetric statistics. *Ann. Statist.*, 14(2):753–758, 1986.

## A Tightness of KKL theorem for Cartesian Products of Graphs

Our generalization of the KKL theorem for Cartesian products of graphs in particular implies one for the  $q$ -ary hypercube  $[q]^k$ . It is easy to see that the  $q$ -ary hypercube is exactly the graph  $K_q^{\square k}$  where  $K_q$  denotes the complete graph. We need the log-Sobolev constant of  $K_q$ , which is known to be  $\Theta(\frac{1}{\log q})$  for  $q \geq 3$  [6]. This gives the following as an easy corollary.

**Corollary A.1** *Given a function  $f : V(K_q^{\otimes k}) \rightarrow \{-1, 1\}$ ,*

$$\max_j \langle f, \mathbf{K}_j f \rangle \geq \Omega \left( \frac{\log k}{k \log q} \cdot \mathbf{Var} f \right)$$

Assuming this result, there is a simple way to obtain a variant of the KKL with different bounds. We can apply the conductance bound for graph  $G$  to conclude,

**Corollary A.2** Given a function  $f : V(G^{\square k}) \rightarrow \{-1, 1\}$ ,

$$\max_j \langle f, \mathbf{L}_j f \rangle \geq \Omega \left( \frac{\log k}{k \log n} \cdot \Phi(G) \cdot \mathbf{Var} f \right),$$

where  $n = |V(G)|$ .

In general, the above corollary is incomparable to Theorem 3.5. We get a quantitatively better bound from Theorem 3.5 if  $\alpha \gg \frac{\Phi}{\log n}$ .

An example where this is true is the following: Consider the  $R$ -dimensional hypercube  $H_R$  and consider  $H_R^{\square k}$ . Though this graph is isomorphic to the  $kR$ -dimensional hypercube, our notion of *influence* now translates to the number of edges cut along one of the hypercubes  $H_R$  i.e. the number of edges along one of the  $k$  blocks of  $n$  coordinates each.

Assume we have a boolean function  $f$  with variance  $\Omega(1)$  on this graph. Applying Theorem 3.5 to this instance, we conclude that there must be a block of  $R$  coordinates, along which the fraction of edges cut is at least  $\frac{\log k}{kR}$ . Whereas, the bound that Corollary A.2 gives is  $\frac{\log k}{kR^2}$ .

**An example with maximum influence**  $o\left(\Phi \cdot \frac{\log k}{k}\right)$  For the above example when  $G$  is the  $R$ -dimensional hypercube, the bound given by Theorem 3.5 is  $\frac{1}{R} \cdot \frac{\log k}{k}$ , which is also  $\Omega\left(\Phi \cdot \frac{\log k}{k}\right)$  for the  $R$ -dimensional hypercube since  $\Phi = \alpha/2 = 1/R$  for this graph. We give an example to show that in general the bound of Theorem 3.5 *cannot* be improved to  $\Omega(\Phi \cdot \frac{\log k}{k})$ .

Consider the  $R$ -dimensional hypercube in which we identify any two vertices which are the same after a cyclic permutation of the coordinates. Formally, the vertex set is  $(\{0, 1\}^R \setminus \{0^R, 1^R\}) / \mathcal{C}$ , where  $\mathcal{C}$  is the group of cyclic permutations on  $\{1, \dots, R\}$ . Each vertex of the new graph is then an equivalence class. There is an edge between two classes  $C_1$  and  $C_2$  if there exist  $u \in C_1$  and  $v \in C_2$  such that  $(u, v)$  is an edge in the hypercube. We take this to be our graph  $G$ .

It follows from the KKL theorem for the hypercube that for the above graph,  $\Phi = \Omega(\log R/R)$ . It also follows from the results of Devanur et al. [5] that for this graph,  $\lambda = O(1/R)$  and hence  $\alpha = O(1/R)$  (it is in fact easy to check that  $\alpha = \Theta(1/R)$  since collapsing vertices into equivalence classes can only increase  $\alpha$ , which was  $2/R$  for the hypercube). The bound given by Theorem 3.5 for maximum influence of a coordinate in  $G^{\square k}$  is again  $\Omega\left(\frac{1}{R} \cdot \frac{\log k}{k}\right)$ . We give below a Boolean function with variance  $\Omega(1)$  and maximum influence  $\Omega\left(\frac{1}{R} \cdot \frac{\log k}{k}\right)$ , which shows that the bound cannot be improved to  $\Omega(\Phi \cdot \frac{\log k}{k})$ .

A vertex  $x \in V(G^{\square k})$  is of the form  $(C_1, \dots, C_k)$  where  $C_1, \dots, C_k$  are equivalence classes in  $\{0, 1\}^R$  as described above. We define  $f$  as

$$f(C_1, \dots, C_k) = 1 \quad \text{iff} \quad \exists i \in k \text{ and } u \in C_i \text{ such that } u \text{ has } \log(kR) \text{ consecutive 1s.}$$

It is easy to check that  $\Pr[f = 0]$  is  $(1 - (R/kR))^k \approx 1/e$  since each  $C_i$  has  $\log(kR)$  consecutive 1s with probability  $R/kR$ . Thus,  $\mathbf{Var}(f) = \Omega(1)$ . Also, the definition of  $f$  is symmetric in all coordinates  $i$  and hence all influences are equal. It remains to compute the influence of a coordinate.

We estimate the probability that a random edge along the first coordinate (say) is crossed by the cut that  $f$  gives. Let  $f(C_1, C_2, \dots, C_k) = 1$  and  $f(C'_1, C_2, \dots, C_k) = 0$  where  $(C_1, C'_1)$  is an edge in  $G$ . Then  $C_1$  must have exactly one substring with  $\log(kR)$  1s, which happens with probability  $O(1/k)$ . Also,  $C'_1$  must differ from  $C_1$  in one of these  $\log(kR)$  1s. This happens with probability  $\log(kR)/R$ . Hence the fraction of edges crossed by the cut is  $O(\log(kR)/kR)$ . Choosing  $R = k^{O(1)}$  shows that this is  $O(\alpha \cdot (\log k/k))$ .

## B Applications to integrality gaps for Sparsest Cut

In this section, we show that the Cartesian product is the right method of *padding* Sparsest Cut integrality gap instances. We recall that the Sparsest Cut value of a graph  $G$  is determined by the Conductance of  $G$  up to a factor of 2. In this section, we will work with Conductance.

Firstly, we show that, given a graph  $G$  that has a conductance value of  $\Phi$ ,  $G^{\square k}$  has a conductance value of  $\frac{1}{k}\Phi$ . We also show a similar statement for the optimum value for several families of SDP relaxations of Sparsest Cut. Informally, we show that, given a graph and an SDP value of  $\text{Opt}$ , and an SDP value of  $\frac{1}{k}\text{Opt}$ , where the SDP is any one of several common classes of SDPs. In particular, the ratio of the two values is preserved.

We first prove the theorem about conductance. This theorem is similar to the one proved by Chung and Tetali[3] and we include a proof for completeness.

**Theorem B.1 (Conductance Value)** *Given a graph  $G$ , we can relate the Conductance of  $G^{\square k}$  and  $G$  as follows,*

$$\Phi(G^{\square k}) = \frac{1}{k}\Phi(G).$$

**Proof:** Let us first prove the simple direction  $\Phi(G^{\square k}) \leq \frac{1}{k}\Phi(G)$ . In order to prove this, fix a function  $f : V(G) \rightarrow \{-1, 1\}$  that achieves

$$\frac{\langle f, \mathbf{L}_G f \rangle}{2 \mathbf{Var} f} = \Phi(G).$$

Now, define  $g : V(G^{\square k}) \rightarrow \mathbb{R}$  as  $g(v_1, \dots, v_k) = f(v_1)$ . It is easy to see that  $\mathbf{Var} g = \mathbf{Var} f$ . Now, let us compute  $\langle g, \mathbf{L}_{G^{\square k}} g \rangle$ .

$$\begin{aligned} \langle g, \mathbf{L}_{G^{\square k}} g \rangle &= \frac{1}{2} \mathbf{E}_{(x,y) \in E(G^{\square k})} (g(x) - g(y))^2 \\ &= \frac{1}{2} \mathbf{E}_{(x,y) \in E(G^{\square k})} (f(x_1) - f(y_1))^2 \\ &= \frac{1}{2k} \mathbf{E}_{x_2, \dots, x_k} \mathbf{E}_{(x_1, y_1) \in E(G)} (f(x_1) - f(y_1))^2 \end{aligned}$$

(for all other edges the contribution is 0)

$$\begin{aligned} &= \frac{1}{2k} \mathbf{E}_{x_2, \dots, x_k} 2 \langle f, \mathbf{L}_G f \rangle \\ &= \frac{2\Phi(G)}{k} \mathbf{Var} f = \frac{2\Phi(G)}{k} \mathbf{Var} g \end{aligned}$$

Thus,  $\Phi(G^{\square k}) = \frac{1}{k}\Phi(G)$ .

Now, let us prove that  $\Phi(G^{\square k}) \geq \frac{1}{k}\Phi(G)$ . Fix a set  $S \subseteq V(G^{\square k})$ . Let  $f$  denote the  $\{-1, 1\}$ -valued indicator function for  $S$ . If we fix  $\hat{v} = (v_2, \dots, v_k) \in V(G)^{k-1}$  and consider all  $x_1 \in V(G)$ , we get a copy of  $G$  when we look at the vertices  $x = (x_1, \hat{v})$ . Denote the restriction of  $f$  to these vertices as  $f_{\hat{v}}$ . On this copy of the graph, we know that

$$\langle f_{\hat{v}}, \mathbf{L}_G f_{\hat{v}} \rangle \geq 2 \Phi(G) \mathbf{Var}_{x_1} f_{\hat{v}}.$$

Averaging the above equation over all  $\hat{v}$ , we get that,

$$\langle f, \mathbf{L}_1 f \rangle \geq 2 \Phi(G) \mathbf{E}_{\hat{v}} \mathbf{Var}_{x_1} f_{\hat{v}} = 2 \Phi(G) \mathbf{E}_x \mathbf{Var}_{x_1} f.$$

Similarly, for all  $j$ , we get that,

$$\langle f, \mathbf{L}_j f \rangle \geq 2 \Phi(G) \mathbf{E}_x \mathbf{Var}_{x_j} f .$$

Now, averaging over  $j$ , we get that

$$\langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle \geq \frac{2\Phi(G)}{k} \sum_j \mathbf{E}_x \mathbf{Var}_{x_j} f .$$

In the next lemma, we show that  $\sum_j \mathbf{E}_x \mathbf{Var}_{x_j} f \geq \mathbf{Var} f$ . This is the well-known Efron-Stein inequality (see [14] for example). We give a proof below for completeness. Assuming the lemma, we can complete the proof by concluding  $\langle f, \mathbf{L}_{\mathbf{G}^{\square k}} f \rangle \geq \frac{2\Phi(G)}{k} \mathbf{Var} f$  and hence  $\Phi(G^{\square k}) \geq \frac{1}{k} \Phi(G)$ . ■

**Lemma B.2 (Sub-additivity of Variance)** *For any function  $f : V(G^{\square k}) \rightarrow \mathbb{R}$ ,*

$$\sum_j \mathbf{E}_x \mathbf{Var}_{x_j} f \geq \mathbf{Var}_x f$$

**Proof:** We will prove that  $\mathbf{E}_{x_2} \mathbf{Var}_{x_1} f + \mathbf{E}_{x_1} \mathbf{Var}_{x_2} f \geq \mathbf{Var}_{x_1, x_2} f$  and the result will follow by repeated application of this result, replacing  $x_1$  with the group of variables  $(x_1, \dots, x_{j-1})$  and  $x_2$  with  $x_j$ .

$$\begin{aligned} \mathbf{E}_{x_2} \mathbf{Var}_{x_1} f + \mathbf{E}_{x_1} \mathbf{Var}_{x_2} f &\geq \mathbf{Var}_{x_1} \left( \mathbf{E}_{x_2} f \right) + \mathbf{E}_{x_1} \mathbf{Var}_{x_2} f \\ &= \mathbf{E}_{x_1} \left( \mathbf{E}_{x_2} f \right)^2 - \left( \mathbf{E}_{x_1, x_2} f \right)^2 + \mathbf{E}_{x_1, x_2} f^2 - \mathbf{E}_{x_1} \left( \mathbf{E}_{x_2} f \right)^2 \\ &= \mathbf{E}_{x_1, x_2} f^2 - \left( \mathbf{E}_{x_1, x_2} f \right)^2 = \mathbf{Var}_{x_1, x_2} f , \end{aligned}$$

where the first inequality uses Jensen's inequality and the convexity of  $\mathbf{Var} f$ . ■

Now, we will show that if we consider an SDP relaxation for Sparsest Cut, the SDP value for  $G^{\square k}$  will be  $\frac{1}{k} \text{Opt}$ , where  $\text{Opt}$  is the SDP value for  $G$ . This theorem holds for the following families of SDP relaxations - standard SDP, SDP with triangle inequalities, SDP with  $k$ -gonal inequalities, standard SDP with  $t$  levels of Sherali-Adams constraints (for any fixed  $t$ ) and SDP at  $t^{\text{th}}$  levels of Lasserre hierarchy.

**Theorem B.3 (SDP Value for Sparsest Cut)** *Let  $\Psi$  be an SDP relaxation for Sparsest Cut that is one of the following: the standard relaxation with  $t$  levels of SA variables, the standard relaxation lifted to  $t$  level of Lasserre hierarchy ( $t$  is arbitrary). Then, denoting the relaxation  $\Psi$  applied to  $G$  as  $\Psi(G)$  and the optimum for  $\Psi(G)$  by  $\text{Opt}(G)$ ,*

$$\text{Opt}(G^{\square k}) \leq \frac{1}{k} \text{Opt}(G)$$

**Proof:** Let  $\{v_u\}_{u \in V(G)}$  be an optimum solution to the SDP  $\Psi(G)$ . We will construct a solution for  $\Psi(G^{\square k})$  with objective value  $\frac{1}{k} \text{Opt}(G)$ .

**Standard SDP** The Standard SDP for Sparsest Cut applied to graph  $G$  is the following:

$$\min \quad \mathbf{E}_{\{x,y\} \in E(G)} \|v_x - v_y\|_2^2 \quad (2)$$

$$\text{s.t.} \quad \mathbf{E}_{x,y} \|v_x - v_y\|_2^2 = 1 \quad (3)$$

Define the vectors  $v_x = \bigoplus_i v_{x_i}$  where  $x = (x_1, \dots, x_k) \in V(G^{\square k})$  and  $\bigoplus$  denotes the direct-sum operation. First, consider the objective function (2),

$$\begin{aligned} \mathbf{E}_{(x,y) \in E(G^{\square k})} \|v_x - v_y\|^2 &= \mathbf{E}_j \mathbf{E}_{(x,y) \in E_j(G^{\square k})} \|v_x - v_y\|^2 \\ &= \frac{1}{k} \mathbf{E}_j \mathbf{E}_{(x,y) \in E_j(G^{\square k})} \|v_{x_j} - v_{y_j}\|^2 \\ &= \frac{1}{k} \mathbf{E}_j \text{Opt}(G) = \frac{1}{k} \text{Opt}(G) \end{aligned}$$

Now, let us verify that the *spreadness* constraint (3) is satisfied.

$$\mathbf{E}_{x,y \in V(G^{\square k})} \|v_x - v_y\|^2 = \mathbf{E}_{x,y \in V(G^{\square k})} \mathbf{E}_j \|v_{x_j} - v_{y_j}\|^2 = \mathbf{E}_{x,y \in V(G)} \|v_x - v_y\|^2$$

**$k$ -gonal inequalities** The relaxation  $\Psi$  may contain a certain family of constraints that must be satisfied by the vectors (e.g. triangle inequalities,  $k$ -gonal inequalities). The best approximation algorithm by Arora, Rao and Vazirani[1] uses triangle inequalities in the SDP relaxation.

$$\forall x, y, z \in V(G), \quad \|v_x - v_y\|_2^2 + \|v_y - v_z\|_2^2 \geq \|v_x - v_z\|_2^2$$

We know that the vectors we have constructed satisfy  $\langle v_x, v_y \rangle = \mathbf{E}_j \langle v_{x_j}, v_{y_j} \rangle$ . It follows that as long as these set of constraints is invariant under a permutation of the vertices and is linear in the dot product of the vectors (which is true in case of the examples mentioned), the new vectors also satisfy the corresponding constraints by linearity.

**SDP with Sherali-Adams Constraints** Let us now consider the Sherali-Adams constraints. The Sherali-Adams SDP for level  $t$  requires that for every set  $T \subseteq V(G)$  of at most  $t$  vertices, there must be a distribution  $\mathcal{D}_T$  on the integral assignments to  $T$  ( $-1, 1$  valued). There are consistency constraints requiring that for two sets  $T_1, T_2$ , the marginals of the distributions  $\mathcal{D}_{T_1}, \mathcal{D}_{T_2}$  on  $T_1 \cap T_2$  are identical. There are also consistency constraints with the vector solution requiring that

$$\forall x, y \in V(G), \quad \langle v_x, v_y \rangle = \mathbf{E}_{(z_x, z_y) \sim \mathcal{D}_{\{x,y\}}} z_x z_y$$

Now, let us define the distributions that correspond to the solution for  $\Psi(G^{\square k})$ . In order to sample from  $\mathcal{D}_{\{x^1, \dots, x^t\}}$ , we pick a random  $j \in [k]$ , draw a sample  $z \sim \mathcal{D}_{\{x_j^1, \dots, x_j^t\}}$  and output  $z$  (Note that we are concerned only about sets, so if an element appears more than once, it's treated as if it appeared just once). Consistency of marginals follows easily because of linearity. Again, since  $\langle v_x, v_y \rangle = \mathbf{E}_j \langle v_{x_j}, v_{y_j} \rangle$ , by linearity, we get that the distributions constructed are consistent with the vectors.

**Lasserre SDP** The SDP at  $t^{\text{th}}$  round of the Lasserre Hierarchy has vectors for every subset  $S$  with at most  $t$  vertices in the graph and has the following family of constraints,

$$\langle v_{S_1}, v_{S_2} \rangle = \langle v_{T_1}, v_{T_2} \rangle \text{ whenever } S_1 \Delta S_2 = T_1 \Delta T_2 .$$

Given vectors for  $S \subseteq V(G)$ , we wish to construct vectors for  $T \subseteq V(G^{\square k})$  so that the consistency constraints are satisfied. Define sets  $T_j \subseteq V(G)$  as follows.

$$T_j = \{ y \mid \#\{x \in T \mid x_j = y\} \text{ is odd} \}$$

Now,  $v_T$  is the direct sum of these  $k$  vectors,  $v_T = \bigoplus_{j=1}^k v_{T_j}$ . It is easy to verify that these vectors satisfy the consistency constraints.  $\blacksquare$

## C Irregular graphs

All of the above results hold for irregular and weighted graphs too. Here, we detail the right definitions for irregular graphs.

Let  $\mu$  be a probability distribution on unordered pairs of vertices, that describes the edges. Let  $\pi : V(G) \rightarrow \mathbb{R}$  be the stationary distribution on the vertices of the graph for the random walk determined by  $\mu$ . The two distributions satisfy the following consistency requirement.

$$2\pi(v) = \sum_u \mu(\{u, v\})$$

The definition of inner product for the space of functions  $V(G) \rightarrow \mathbb{R}$  remains the same,

$$\langle f, g \rangle = \mathbf{E}_{x \sim \pi} [f(x)g(x)].$$

The variance operator is defined as follows,

$$\mathbf{Var}(f) = \mathbf{E}_{x \sim \pi} f(x)^2 - \left( \mathbf{E}_{x \sim \pi} f(x) \right)^2$$

The normalized Laplacian of the graph,  $\mathbf{L}_G$  is defined as,

$$(\mathbf{L}_G)_{ij} = \begin{cases} \pi(i) & , i = j \\ -\frac{\mu(\{i, j\})}{2} & , i \neq j \end{cases}$$

It is easy to verify that  $\mathbf{L}_G$  satisfies,

$$\langle f, \mathbf{L}_G f \rangle = \frac{1}{2} \mathbf{E}_{\{u, v\} \sim \mu} (f(x) - f(y))^2$$

Now, we need to define the distributions on the vertices and edges of  $G^{\square k}$ . The distribution on the vertices is just the product distribution,

$$\pi(v_1, \dots, v_k) = \prod_{i=1}^k \pi(v_i).$$

The distribution on the edges needs to be carefully defined. As before,  $\mu$  will be zero unless the pairs of vertices have  $k - 1$  coordinates that are identical. For the remaining pairs, we define  $\mu$  as follows,

$$\begin{aligned} & \mu(\{(v_1, \dots, v_{j-1}, u, v_{j+1}, \dots, v_k), (v_1, \dots, v_{j-1}, v, v_{j+1}, \dots, v_k)\}) \\ &= \frac{1}{2k} \pi(v_1) \cdot \dots \cdot \pi(v_{j-1}) \cdot \mu(\{u, v\}) \cdot \pi(v_{j+1}) \cdot \dots \cdot \pi(v_k) \end{aligned}$$

It is easy to see that the distributions defined above are consistent.

Under these definitions, it still holds that

$$\langle u_1 \otimes u_2, v_1 \otimes v_2 \rangle = \langle u_1, v_1 \rangle \langle u_2, v_2 \rangle$$

Denote by  $\Pi$  the diagonal matrix with  $(\Pi)_{ii} = \pi(v_i)$ . We define the operator  $\mathbf{L}_j$  to be

$$\mathbf{L}_j = \Pi \otimes \dots \otimes \Pi \otimes \mathbf{L}_G \otimes \Pi \dots \otimes \Pi$$

where the matrix  $\mathbf{L}_G$  is in the  $j^{\text{th}}$  position. It is easy to verify that,

$$\mathbf{L}_{G^{\square k}} = \mathbf{E}_j \mathbf{L}_j.$$

The volume of a set  $\text{Vol}(S)$  will now be the measure of  $S$  under  $\pi$ ,

$$\text{Vol}(S) = \sum_{v \in S} \pi(v).$$

The definition of the sparsity then becomes,

$$\Phi(G) = \min_{S \subseteq V(G)} \frac{\sum_{u \in S, v \in \bar{S}} \mu(\{u, v\})}{\text{Vol}(S) \text{Vol}(\bar{S})} = \min_{f: V(G) \rightarrow \{-1, 1\}} \frac{\langle f, \mathbf{L}_G f \rangle}{2 \mathbf{Var}(f)}$$

Using these definitions, all the previous proofs go through without modifications.