Sharpness of KKL on Schreier graphs

Ryan O'Donnell[*] and Karl Wimmer[**]

[1] Computer Science Department, Carnegie Mellon University
[2] Department of Mathematics and Computer Science, Duquesne University

Abstract. Recently, the Kahn–Kalai–Linial (KKL) Theorem on influences of functions on $\{0,1\}^n$ was extended to the setting of functions on Schreier graphs. Specifically, it was shown that for an undirected Schreier graph $\mathrm{Sch}(G, X, U)$ with log-Sobolev constant ρ and generating set U closed under conjugation, if $f\colon X \to \{0,1\}$ then $\mathcal{E}[f] \gtrsim \log(1/\mathcal{M}[f]) \cdot \rho \cdot \mathbf{Var}[f]$. Here $\mathcal{E}[f]$ denotes the average of f's influences, and $\mathcal{M}[f]$ denotes their maximum.

In this work we investigate the extent to which this result is sharp. We show:
  – The condition that U is closed under conjugation cannot in general be eliminated.
  – The log-Sobolev constant cannot be replaced by the modified log-Sobolev constant.
  – The result cannot be improved for the Cayley graph on S_n with transpositions.
  – The result *can* be improved for the Cayley graph on $\mathbb{Z}_m^n$ with standard generators.
  – Talagrand's strengthened version of KKL also holds in the Schreier graph setting:
$$\underset{u \in U}{\mathrm{avg}}\big\{\mathcal{I}_u[f]/\log(1/\mathcal{I}_u[f])\big\} \gtrsim \rho \cdot \mathbf{Var}[f].$$

1 Introduction

1.1 The KKL Theorem

The famed KKL (Kahn–Kalai–Linial) Theorem [KKL88] asserts that for any "roughly balanced" function $f\colon \{0,1\}^n \to \{0,1\}$, one of the coordinate $i \in [n]$ must have "influence" $\mathcal{I}_i[f] \gtrsim \frac{\log n}{n}$. This theorem, along with generalizations by Bourgain–Kahn–Kalai–Katznelson–Linial [BKK+92], Talagrand [Tal94], and Friedgut [Fri98], has important applications in numerous areas of computer science and mathematics, including distributed computing [BOL90], random k-SAT [Fri99] and random graphs [FK96], communication complexity [Raz95], hardness of approximation [DS05,CKK+06,KR08], metric embeddings [KR09,DKSV06], weak random sources [KZ06], learning theory [OS08,OW09], and extremal combinatorics [OW09]. To state the KKL Theorem more precisely, we first introduce some definitions.

Definition 1. *For $f\colon \{0,1\}^n \to \mathbb{R}$ we define:*

$$\mathbf{Var}[f] = \underset{\boldsymbol{x}}{\mathbf{E}}[f(\boldsymbol{x})^2] - \underset{\boldsymbol{x}}{\mathbf{E}}[f(\boldsymbol{x})]^2;$$

$$\mathcal{I}_i[f] = \tfrac{1}{2} \underset{\boldsymbol{x}}{\mathbf{E}}[(f(\boldsymbol{x}) - f(\boldsymbol{x}^{e_i}))^2], \quad i \in [n];$$

$$\mathcal{E}[f] = \underset{i \in [n]}{\mathrm{avg}}\{\mathcal{I}_i[f]\}; \qquad \mathcal{M}[f] = \underset{i \in [n]}{\max}\{\mathcal{I}_i[f]\}.$$

Here the random variable $\boldsymbol{x}$ is always uniformly distributed on $\{0,1\}^n$, and x^{e_i} denotes x with its ith coordinate flipped.[3]

---

[3] The factor of $\tfrac{1}{2}$ in our definition of $\mathcal{I}_i[f]$ is sometimes 1 or $\tfrac{1}{4}$ elsewhere in the literature; the difference is not important for this paper.

We may now state the KKL Theorem:

KKL Theorem. *Let $f\colon \{0,1\}^n \to \{0,1\}$. Then $\mathcal{M}[f] \gtrsim \frac{\log n}{n} \cdot \mathbf{Var}[f]$.*

(Throughout this paper, $A \gtrsim B$ means that $A \geq cB$ for some absolute constant $c > 0$, and log denotes the natural logarithm.) The KKL Theorem should be compared to the much easier "Poincaré Inequality" for $\{0,1\}^n$:

Poincaré Inequality. *Let $f\colon \{0,1\}^n \to \mathbb{R}$. Then $\mathcal{E}[f] \geq \frac{2}{n} \cdot \mathbf{Var}[f]$.*

Note that the Poincaré Inequality may be sharp even for functions $f\colon \{0,1\}^n \to \{0,1\}$; e.g., when $f(x) = x_i$ for some $i \in [n]$. The proof method used in [KKL88] can easily be extended to give the following sharper result, first stated by Talagrand [Tal94]:

KKL Theorem 2. *Let $f\colon \{0,1\}^n \to \{0,1\}$. Then $\mathcal{E}[f] \gtrsim \frac{\log(1/\mathcal{M}[f])}{n} \cdot \mathbf{Var}[f]$.*

Indeed, Talagrand [Tal94] gave an even sharper version:

Talagrand Theorem. *Let $f\colon \{0,1\}^n \to \{0,1\}$. Then $\mathrm{avg}_{i\in[n]}\left\{\mathcal{I}_i[f] \,/\, \log(1/\mathcal{I}_i[f])\right\} \gtrsim \mathbf{Var}[f]$.*

To see that this is strictly sharper, note that it rules out a function f having $\mathbf{Var}[f] \gtrsim 1$, $\mathcal{I}_1[f] = 1/\log n$, and $\mathcal{I}_i[f] = (\log\log n)/n$ for $i \geq 2$, something that isn't ruled out by KKL Theorem 2.

1.2 KKL on Schreier graphs

In a recent work [OW09], the authors generalized KKL Theorem 2 to the setting of functions on Schreier graphs. Let us recall this setting. Let G be a finite group acting transitively on a finite set X. Let $U \subseteq G$ be a generating set which is symmetric; i.e., closed under inverses. The associated *Schreier graph* $\mathrm{Sch}(G, X, U)$ is the undirected graph with vertex set X and edges (x, x^u) for each $x \in X$, $u \in U$, where x^u denotes the action of u on x. In the special case that G acts on $X = G$ by $x^u = xu$, the Schreier graph is simply the Cayley graph $\mathrm{Cay}(G, U)$. This is the case in the original KKL Theorem setting, where $G = X = \mathbb{Z}_2^n$ and U is the standard generating set $U = \{e_i : i \in [n]\}$.

There is a natural random walk on a Schreier graph $\mathrm{Sch}(G, X, U)$. Let K denote the Markov transition matrix for this walk and write $L = id - K$ for the normalized Laplacian. Since $\mathrm{Sch}(G, X, U)$ is undirected, regular, and connected, the random walk has a unique invariant probability measure: the uniform distribution on X. We denote this measure by π, and we write $\ell^2(\pi)$ for the inner product space of functions $f\colon X \to \mathbb{R}$ with inner product $\langle f, g \rangle = \mathbf{E}_{\boldsymbol{x}\sim\pi}[f(\boldsymbol{x})g(\boldsymbol{x})]$.

Thinking of K as an operator on $\ell^2(\pi)$, we have $K = \mathrm{avg}_{u\in U} K_u$, where K_u is the operator defined by $K_u f(x) = f(x^u)$. Similarly, $L = \mathrm{avg}_{u\in U} L_u$, where we define the operator $L_u = id - K_u$. We have the following simple facts:

Proposition 1. *The adjoint K_u^* of K_u is $K_{u^{-1}}$, and similarly $L_u^* = L_{u^{-1}}$.*

Proof. This holds because for each fixed $u \in U$, the pairs $(\boldsymbol{x}, \boldsymbol{x}^u)$ and $(\boldsymbol{x}^{u^{-1}}, \boldsymbol{x})$ have the same distribution when $\boldsymbol{x} \sim \pi$.

Proposition 2. *For all $f, g \in \ell^2(\pi)$ it holds that $\langle L_u f, L_u g \rangle = \langle f, L_u g \rangle + \langle f, L_u^* g \rangle$.*

Proof. We have $\langle L_u f, L_u g \rangle = \langle f - K_u f, g - K_u g \rangle = (\langle f, g \rangle - \langle f, K_u g \rangle) + (\langle K_u f, K_u g \rangle - \langle K_u f, g \rangle)$. The first quantity in parentheses is precisely $\langle f, L_u g \rangle$. As for the second quantity, we have $\langle K_u f, K_u g \rangle = \mathbf{E}_{\boldsymbol{x}\sim\pi}[f(\boldsymbol{x}^u)g(\boldsymbol{x}^u)] = \langle f, g \rangle$, because $\boldsymbol{x}^u$ is uniformly distributed when $\boldsymbol{x}$ is; furthermore, $\langle K_u f, g \rangle = \langle f, K_u^* g \rangle$. Hence the second quantity in parentheses is indeed $\langle f, L_u^* g \rangle$ as required.

We may now define "influences" in the Schreier graph setting:

Definition 2. *The influence of generator $u \in U$ on $f \in \ell^2(\pi)$ is defined to be*

$$\mathcal{I}_u[f] = \langle f, L_u f \rangle = \frac{1}{2}\|L_u f\|_2^2,$$

where the second equality is by Proposition 2. We also define $\mathcal{M}[f] = \max_{u \in U}\{\mathcal{I}_u[f]\}$ and

$$\mathcal{E}[f] = \underset{u \in U}{\operatorname{avg}}\left\{\mathcal{I}_u[f]\right\} = \langle f, Lf \rangle,$$

which is sometimes called the "energy" of f.

These definitions agree with those in the original KKL Theorem setting where $G = X = \mathbb{Z}_2^n$, $U = \{e_i : i \in [n]\}$.

To state the KKL Theorem in Schreier graphs we must also recall the "log-Sobolev inequality" for Markov chains. For a nonnegative function $f \in \ell^2(\pi)$, the *entropy* of f is defined to be

$$\mathbf{Ent}[f] = \underset{\boldsymbol{x} \sim \pi}{\mathbf{E}}[f(\boldsymbol{x})\log f(\boldsymbol{x})] - \underset{\boldsymbol{x} \sim \pi}{\mathbf{E}}[f(\boldsymbol{x})]\log\underset{\boldsymbol{x} \sim \pi}{\mathbf{E}}[f(\boldsymbol{x})],$$

with $0\log 0$ defined to be 0. The *log-Sobolev constant* for the Markov chain on $\operatorname{Sch}(G, X, U)$ is defined to be the largest constant ρ such that the following inequality holds:

Log-Sobolev Inequality. *For all nonconstant $f \in \ell^2(\pi)$, $\mathbf{Ent}[f^2] \leq 2\rho^{-1}\mathcal{E}[f]$.*

This notion was introduced by Gross [Gro75] who showed the following:

Gross's Theorem. *For $\operatorname{Cay}(\mathbb{Z}_2^n, U)$ with $U = \{e_i : i \in [n]\}$, the log-Sobolev constant is $\rho = \frac{2}{n}$.*

We may now state the authors' generalization [OW09] of KKL Theorem 2 to Schreier graphs:

Theorem 1. *Let $\operatorname{Sch}(G, X, U)$ be a Schreier graph with log-Sobolev constant ρ. Assume that U is closed under conjugation. Then for all $f \colon X \to \{0, 1\}$ it holds that $\mathcal{E}[f] \gtrsim \log(1/\mathcal{M}[f]) \cdot \rho \cdot \mathbf{Var}[f]$. In particular, $\mathcal{M}[f] \gtrsim \rho \log(1/\rho) \cdot \mathbf{Var}[f]$.*

The motivation for this theorem was the setting where X is the set of length-n binary strings of Hamming weight k, $G = S_n$ acts on X by permuting coordinates, and $U = \{(ij) : 1 \leq i < j \leq n\}$. The log-Sobolev constant for this Schreier graph is known [LY98] to be $\rho = \Theta(\frac{1}{n})$ assuming $\frac{k}{n}$ is bounded away from 0 and 1. Using the resulting KKL Theorem, the authors were able to give a "robust" version of the classical Kruskal–Katona theorem, as well as an optimal weak-learning algorithm for the class of monotone Boolean functions.

1.3 Our results on sharpness of the inequality

In this paper we address several natural questions one might ask regarding the sharpness of Theorem 1. The most obvious question is whether the condition that U be closed under conjugation is really necessary. Although originally inclined to believe it is not, we show here the following:

Theorem 2. *The assumption that U is closed under conjugation cannot in general be removed from Theorem 1, even for Cayley graphs.*

There are also natural cases where Theorem 1 does not give a strong result because the log-Sobolev constant is too small. One such example is the Cayley graph on S_n with generating set given by transpositions. In this case the log-Sobolev constant is known [DSC96,LY98] to be $\rho = \Theta(\frac{1}{n\log n})$. Hence for $f \colon S_n \to \{0, 1\}$ with $\mathbf{Var}[f] \gtrsim 1$, Theorem 1 only implies that $\mathcal{M}[f] \gtrsim \frac{1}{n}$. One might ask whether this inequality can nevertheless be improved. Unfortunately, the answer is no:

Theorem 3. *For the Cayley graph on S_n ($n > 1$) with generating set U equal to all transpositions, there is a function $f \colon S_n \to \{0, 1\}$ with $\mathbf{Var}[f] \geq 2/9$ and $\mathcal{I}_u[f] \leq \frac{2}{n}$ for all $u \in U$.*

The proof is short enough that we can give it here. Say that $f(\sigma) = 1$ if σ is a derangement (i.e., has no fixed point) and $f(\sigma) = 0$ otherwise. It is well known [**?**] that the fraction of permutations in S_n which are derangements is $\sum_{i=0}^{n} \frac{(-1)^i}{i!} \in [1/3, 1/2]$; thus $\mathbf{Var}[f] \geq 2/9$. By symmetry, all transpositions u have the same influence; the influence of (12), say, is

$$\Pr_{\boldsymbol{\sigma} \sim S_n} [\boldsymbol{\sigma} \text{ has no fixed point}, \boldsymbol{\sigma} \cdot (12) \text{ has a fixed point}].$$

For the event in question to occur, $\boldsymbol{\sigma}$ must have either $\boldsymbol{\sigma}(1) = 2$ or $\boldsymbol{\sigma}(2) = 1$. But the probability of this is at most $\frac{1}{n} + \frac{1}{n} = \frac{2}{n}$, completing the proof of Theorem 3. This theorem also immediately implies:

Corollary 1. *In general, one cannot replace the log-Sobolev constant ρ in Theorem 1 with the modified log-Sobolev constant ρ_0.*

The *modified log-Sobolev constant ρ_0*, which always satisfies $\rho_0 \geq \rho$, was introduced in several papers, dating back to [Wu00]; it is defined to be the largest the constant such that $\mathbf{Ent}[f^2] \leq \frac{1}{2} \rho_0^{-1} \langle f^2, L \log f^2 \rangle$. Corollary 1 follows from Theorem 3 because it is known [GQ03,Goe04,BT06] that the modified log-Sobolev constant for the Cayley graph of S_n with transpositions is $\rho_0 = \Theta(\frac{1}{n})$.

Another natural setting for which Theorem 1 does not give a strong result is the Cayley graph on $\mathbb{Z}_m^n$ with standard generating set $U = \{\pm e_i : i \in n\}$. It is known that the log-Sobolev constant for this Cayley graph satisfies $\rho = \Theta(\frac{1}{m^2 n})$; thus Theorem 1 shows that $\mathcal{M}[f] \gtrsim \frac{\log n + \log m}{m^2 n} \cdot \mathbf{Var}[f]$ for any $f \colon \mathbb{Z}_m^n \to \{0, 1\}$. In contrast to the case of S_n, we show that Theorem 1 *can* be improved for $\mathbb{Z}_m^n$.

Theorem 4. *For any $f \colon \mathbb{Z}_m^n \to \{0, 1\}$ it holds that $\mathcal{E}[f] \gtrsim \frac{\log(1/(m\mathcal{M}[f]))}{n} \cdot \mathbf{Var}[f]$. In particular, $\mathcal{M}[f] \gtrsim \frac{\log n}{mn} \cdot \mathbf{Var}[f]$. Further, these inequalities can be sharp up to a constant factor.*

Finally, one may ask whether Talagrand's strengthening of the KKL Theorem also holds in the Schreier graph setting. We establish this using a method of proof alluded to in Talagrand's paper [Tal94].

Theorem 5. *Let $\mathrm{Sch}(G, X, U)$ be a Schreier graph with log-Sobolev constant ρ. Assume that U is closed under conjugation. Then for all $f \colon X \to \{0, 1\}$ it holds that*

$$\underset{u \in U}{\mathrm{avg}} \left\{ \frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])} \right\} \gtrsim \rho \cdot \mathbf{Var}[f].$$

In particular, this proves the original KKL Theorem with a proof that has not previously appeared in the literature.

1.4 Organization of the remainder of the paper

In Section 2 we show Theorem 2, establishing that the condition that U be closed under conjugation cannot be removed, even for Cayley graphs. The example takes place on the semidirect product $\mathbb{Z}_2^n \rtimes \mathbb{Z}_n$. In Section 3 we show Theorem 4 regarding $\mathbb{Z}_m^n$; the proof follows from a simple combinatorial compression argument combined with the "BKKKL generalization" [BKK+92] of the KKL Theorem. Finally, in Section 4 we prove Theorem 5, the generalization of the Talagrand Theorem to the Schreier graph setting. This proof uses Orlicz norms and is probably the most technically interesting part of the paper.

2   When U is not closed under conjugation

In this section we prove Theorem 2, establishing that the condition that U be closed under conjugation cannot be removed from Theorem 1. Our counterexample will take place on a Cayley graph, $\mathrm{Cay}(G, U)$. The group G is the semidirect product $\mathbb{Z}_2^n \rtimes \mathbb{Z}_n$, where $\mathbb{Z}_n$ acts on $\mathbb{Z}_2^n$ by the natural cyclic shift of coordinates. I.e., for $(x, i), (y, j) \in \mathbb{Z}_2^n \times \mathbb{Z}_n$ the group multiplication is given by

$$(x, i) \cdot (y, j) = (x + y^{\pi_i}, i + j), \tag{1}$$

where $y^{\pi_i} \in \mathbb{Z}_2^n$ is the vector given by cyclically shifting y's coordinates i places to the right. We take U to be the following symmetric generating set of $2n$ elements:

$$U = \{(e_i, 0) : i \in [n]\} \cup \{(\mathbf{0}, j) : j \in \mathbb{Z}_n\}.$$

(If one prefers not to have the group identity in the generating set, it is not hard to alter our argument so that it works for $U \setminus \{(\mathbf{0}, 0)\}$.) We remark that U is not closed under conjugation; e.g.,

$$(e_1, 0) \cdot (\mathbf{0}, 1) \cdot (e_1, 0)^{-1} = (e_1, 0) \cdot (\mathbf{0}, 1) \cdot (e_1, 0) = (e_1, 0) \cdot (e_2, 1) = (e_1 + e_2, 1) \neq (\mathbf{0}, 1).$$

To show that the conclusion of Theorem 1 is not satisfied for $\mathrm{Cay}(G, U)$, it suffices to establish the following two lemmas:

Lemma 1. *There is a function* $f \colon G \to \{0, 1\}$ *with* $\mathbf{Var}[f] = 1/4$ *and* $\mathcal{I}_u[f] \leq \frac{1}{2n}$ *for all* $u \in U$.

Lemma 2. *Assuming* $n > 1$, *the log-Sobolev constant* ρ_G *for* $\mathrm{Cay}(G, U)$ *satisfies* $\rho_G = \frac{1}{n}$.

The first lemma is extremely easy; the function is $f((x, i)) = x_1$. Clearly $\mathbf{Pr}_{(\boldsymbol{x}, \boldsymbol{i})}[f((\boldsymbol{x}, \boldsymbol{i})) = 1] = 1/2$ and hence $\mathbf{Var}[f] = 1/4$. The generators $(\mathbf{0}, i) \in U$ have 0 influence on f, and it is easy to calculate that the generators $(e_i, 0)$ have influence $\frac{1}{2n}$ each. We now prove the second lemma.

Proof. (Lemma 2.) We determine ρ_G by comparing it with the log-Sobolev constant ρ_H of a related Cayley graph. Specifically, let X be the *set* $\mathbb{Z}_2^n \times \mathbb{Z}_n$; we write $G = (X, \cdot)$, where $\cdot$ is the group multiplication defined in (1). Let H be the *abelian direct product group* $\mathbb{Z}_2^n \times \mathbb{Z}_n$; we write $H = (X, +)$. We may interpret $U \subseteq X$ as both a subset of G and of H; it is a symmetric generating set of both. We may interpret any function $f \colon X \to \mathbb{R}$ as being both a function on G and on H; we distinguish the influence of $u \in U$ on f within G and within H as

$$\mathcal{I}_u^G[f] = \mathop{\mathbf{E}}_{\boldsymbol{x} \sim X}[f(\boldsymbol{x})(f(\boldsymbol{x}) - f(\boldsymbol{x} \cdot u))], \qquad \mathcal{I}_u^H[f] = \mathop{\mathbf{E}}_{\boldsymbol{x} \sim X}[f(\boldsymbol{x})(f(\boldsymbol{x}) - f(\boldsymbol{x} + u))].$$

We claim that

$$\mathcal{E}_G[f] = \mathop{\mathrm{avg}}_{u \in U}\{\mathcal{I}_u^G[f]\} = \mathop{\mathrm{avg}}_{u \in U}\{\mathcal{I}_u^H[f]\} = \mathcal{E}_H[f]$$

for all $f \colon X \to \mathbb{R}$. First, $\mathcal{I}_u^G[f] = \mathcal{I}_u^H[f]$ for any $u = (\mathbf{0}, j)$, since $\boldsymbol{x} \cdot u = \boldsymbol{x} + u$ for such u. Second, for the generators in U of the form $(e_i, 0)$ we have

$$\mathop{\mathrm{avg}}_{i \in [n]}\{\mathcal{I}_{(e_i, 0)}^G[f]\} = \mathop{\mathbf{E}}_{\boldsymbol{x} \sim X}\Big[f(\boldsymbol{x})(f(\boldsymbol{x}) - \mathop{\mathrm{avg}}_{i \in [n]}\{f(\boldsymbol{x} \cdot (e_i, 0))\})\Big]$$

$$= \mathop{\mathbf{E}}_{\boldsymbol{x} \sim X}\Big[f(\boldsymbol{x})(f(\boldsymbol{x}) - \mathop{\mathrm{avg}}_{i \in [n]}\{f(\boldsymbol{x} + (e_i, 0))\})\Big] = \mathop{\mathrm{avg}}_{i \in [n]}\{\mathcal{I}_{(e_i, 0)}^H[f]\},$$

because $\{x \cdot (e_i, 0) : i \in [n]\} = \{x + (e_i, 0) : i \in [n]\}$ (as multisets) for each $x \in X$.

Since we have $\mathcal{E}_G[f] = \mathcal{E}_H[f]$ for all $f \colon X \to \mathbb{R}$ and since $\mathbf{Ent}[f]$ does not depend on the group structure, it follows that $\rho_G = \rho_H$. It thus remains to show that $\rho_H = \frac{1}{n}$. The random walk on $\mathrm{Cay}(H, U)$ is the product two random walks, one the standard random walk on $\mathbb{Z}_2^n$ and one the random walk on the complete graph K_n with self-loops. It follows [DSC96, Lemma 3.2] that $\rho_H = \frac{1}{2}\min\{\rho_{\mathbb{Z}_2^n}, \rho_{K_n}\}$. Since $\rho_{\mathbb{Z}_2^n} = \frac{2}{n}$ ([Gro75]) and $\rho_{K_n} = \frac{2(1 - 2/n)}{\log(n-1)} \geq \frac{2}{n}$ ([DSC96, Theorem A.1]), it follows that $\rho_H = \frac{1}{n}$ as claimed.

3   Improved bounds for $\mathbb{Z}_m^n$

Here we prove Theorem 4.

Theorem 4 restated.   *For any $f\colon \mathbb{Z}_m^n \to \{0,1\}$ it holds that*

$$\mathcal{E}[f] \gtrsim \frac{\log(1/(m\mathcal{M}[f]))}{n} \cdot \mathbf{Var}[f]. \tag{2}$$

In particular, $\mathcal{M}[f] \gtrsim \frac{\log n}{mn}\cdot\mathbf{Var}[f]$. Further, these inequalities can be sharp up to a constant factor.

Proof. Let $g\colon \mathbb{Z}_m^n \to \{0,1\}$. For $x \in \mathbb{Z}_m^n$ and $i \in [n]$, we write $g_{i,x}\colon \mathbb{Z}_m \to \{0,1\}$ for the function $g_{i,x}(a) = g(x_1,\ldots,x_{i-1},a,x_{i+1},\ldots,x_n)$; we also identify $g_{i,x}$ with a subset of $\mathbb{Z}_m$. It follows from the definitions that $\mathcal{I}_{e_i}[g] = \mathcal{I}_{-e_i}[g] = \mathbf{E}_x[|\partial g_{i,x}|/2m]$, where we write $\partial g_{i,x} = \{x \in \mathbb{Z}_m : g(x) \neq g(x+1)\}$. We will write simply $\mathcal{I}_i[g]$ for this common value.

Consider the *jth compression operator* σ_j for $j \in [n]$; one may define $\sigma_j g\colon \mathbb{Z}_m^n \to \{0,1\}$ by stating that $(\sigma_j g)_{j,x}(a) = 1[a < |g_{j,x}|]$ for each $x \in \mathbb{Z}_m^n$. Note that $\mathbf{E}[\sigma_j g] = \mathbf{E}[g]$ always. It is a familiar fact in the study of influences (see, e.g., [BOL90]) that compressing a function does not increase any of its influences. In our particular context of $\mathbb{Z}_m^n$ the proof is straightforward and essentially appears in [BL91]. (That paper studies 'grids' rather than our 'discrete torus'; the only difference this makes is for the claim that $\mathcal{I}_i[\sigma_i g] \leq \mathcal{I}_i[g]$, but this follows immediately from the fact that $|\partial(\sigma_i g)_{i,x}| \leq |\partial g_{i,x}|$ for every $x \in \mathbb{Z}_m^n$.) If we now write $\tilde{g} = \sigma_1\sigma_2\cdots\sigma_n g$, we conclude that

$$\mathbf{E}[\tilde{g}] = \mathbf{E}[g], \qquad \mathcal{I}_i[\tilde{g}] \leq \mathcal{I}_i[g] \ \ \forall i, \qquad \forall i,x,\ \tilde{g}_{i,x} = \{0,1,\ldots,a-1\} \text{ for some } 0 \leq a < m.$$

Given any $f\colon \mathbb{Z}_m^n \to \{0,1\}$, the first two facts above show that to prove (2) for f, it suffices to prove it for $\tilde{f}$. Thus without loss of generality we may assume f satisfies the third condition above: for each $i \in [n]$ and $x \in \mathbb{Z}_m^n$ it holds that $f_{i,x} = \{0,1,\ldots,a-1\}$ for some $0 \leq a < m$. In this case, note that $|\partial f_{i,x}|$ is always either 0 or 2, depending on whether or not $f_{i,x}$ is a constant. It follows that

$$\mathcal{I}_i[f] = \mathcal{I}_i'[f]/m = \mathbf{Pr}_{\boldsymbol{x} \sim \mathbb{Z}_m^n}[f_{i,\boldsymbol{x}} \text{ not constant}]/m,$$

where $\mathcal{I}_i'[f]$ denotes the influence of the ith coordinate on f in the sense of Bourgain–Kahn–Kalai–Katznelson–Linial [BKK$^+$92]. Hence proving (2) for f is equivalent to proving

$$\sum_{i=1}^{n} \mathcal{I}_i'[f] \gtrsim \log\big(1/\max_i\{\mathcal{I}_i'[f]\}\big) \cdot \mathbf{Var}[f]. \tag{3}$$

But the inequality (3) was proved for by Friedgut and Kalai [FK96] (building on [BKK$^+$92]) for any $f : \Omega^n \to \{0,1\}$, where Ω^n is a product probability space.

Finally, to show that (2) may be sharp up to a constant, consider functions $f\colon \mathbb{Z}_m^n \to \{0,1\}$ of the form $f(x) = h(\lfloor 2x_1/m \rfloor, \ldots \lfloor 2x_m/m \rfloor)$, where $h\colon \{0,1\}^n \to \{0,1\}$. Then inequality (2) is sharp up to a constant for f if and only if inequality (3) is sharp up to a constant for h with respect to Ω, a p-biased probability space on $\{0,1\}^n$. Here $p = 1/2$ if m is even and $p = 1/2 - 1/2m \in [1/3,1/2)$ if m is odd. In either case, it is well known [FK96] that there are function families h (namely "Tribes") which are sharp for (3) on Ω up to a universal constant.

4   The Talagrand Theorem in Schreier graphs

Recall Talagrand's Theorem, which generalizes the KKL Theorem: for all $f : \{0,1\}^n \to \{0,1\}$,

$$\underset{i \in [n]}{\mathrm{avg}} \left\{ \mathcal{I}_i[f] \ / \ \log(1/\mathcal{I}_i[f]) \right\} \gtrsim \mathbf{Var}[f]. \tag{4}$$

In fact, in [Tal94] Talagrand also proved a version of this result for $\{0,1\}^n$ equipped with the p-biased measure, $p \neq 1/2$. Talagrand straightforwardly deduced (4) from the following Fourier-theoretic inequality:

Talagrand's Inequality.   *For $g : \{0,1\}^n \to \mathbb{R}$ with $\mathbf{E}[g] = 0$,*

$$\sum_{\emptyset \neq S \subseteq [n]} \frac{\widehat{g}(S)^2}{|S|} \lesssim \|g\|_M^2. \tag{5}$$

Here $\| \cdot \|_M$ denotes a certain Orlicz-type norm with $M(t) \sim t^2/\log t$.

Talagrand's proof of (a p-biased version of) (5) is slightly lengthy. It relies in part on a hypercontractive inequality for $\{0,1\}^n$ under the p-biased distribution, which Talagrand proves by reduction to the standard ([Bon70]) $p = 1/2$ case. After the proof, Talagrand remarks that one can obtain (5) in the $p = 1/2$ case "by duality from an inequality of L. Gross [the log-Sobolev inequality], that itself follows from [the hypercontractive inequality]". However it would be two years before the sharp log-Sobolev and hypercontractive constants for the p-biased distribution were obtained [DSC96]; as Talagrand wrote, this "creates complications in using this [log-Sobolev and duality] approach when $p \neq 1/2$". In this section we deduce the approach Talagrand presumably had in mind, and show that it can be extended to the setting of Schreier graphs.

4.1   Basics of Orlicz norms

We begin by recalling some basics of Orlicz norms; see, e.g., [KR61,RR91]. A function $\Phi : \mathbb{R} \to \mathbb{R}^{\geq 0}$ is called a *Young function* if it is convex, even, and satisfies $\Phi(0) = 0$, $\lim_{t \to \infty} \Phi(t) = \infty$. Each such function has a *complementary* Young function Ψ, defined by $\Psi(s) = \sup\{t|s| - \Phi(t) : t \geq 0\}$. Given a measure space (Ω, σ, μ) and a Young function Φ, one can define the following "gauge" (or "Luxemburg") norm on functions $f : \Omega \to \mathbb{R}$:

$$\|f\|_\Phi = \inf\{c > 0 : \int_\Omega \Phi(f/c) \, d\mu \leq 1\}.$$

This norm is closely equivalent to the "Orlicz" norm. Given two Young functions Φ_1 and Φ_2 one writes $\Phi_1 \prec \Phi_2$ if there exist constants $c, t_0 > 0$ such that $\Phi(t) \leq \Phi(ct)$ for all $t \geq t_0$. If $\Phi_1 \prec \Phi_2$ and also $\Phi_2 \prec \Phi_1$, then we call Φ_1 and Φ_2 *equivalent* and write $\Phi_1 \sim \Phi_2$. In this case the associated norms are also equivalent [KR61, (9.24) and (13.7)]:

$$\|f\|_{\Phi_1} \lesssim \|f\|_{\Phi_2} \lesssim \|f\|_{\Phi_1} \qquad \text{for all } f.$$

Finally, if Φ and Ψ are a complementary Young pair we have the following generalized *Hölder's inequality* [RR91, Prop. 1] for functions $f, g : \Omega \to \mathbb{R}$:

$$\int_\Omega |fg| \, d\mu \leq 2\|f\|_\Phi \|g\|_\Psi.$$

4.2 Talagrand's key inequality for Markov chains

We now show how to generalize Talagrand's key inequality (5) to the setting of Markov chains.

These Markov chains need not be random walks on Schreier graphs; for this subsection, we merely assume that we have an irreducible finite Markov chain X with a transition matrix K which is not necessarily reversible. We write π for the (unique) probability distribution on X which is invariant for K, and $\ell^2(\pi)$ for the inner product space of functions $f\colon X \to \mathbb{R}$ with inner product $\langle f, g \rangle = \mathbf{E}_{\boldsymbol{x} \sim \pi}[f(\boldsymbol{x})g(\boldsymbol{x})]$. We also let $L = id - K$ be the (normalized) Laplacian of the Markov chain. The entropy and energy of functions $f \in \ell^2(\pi)$ are defined as in the Schreier graph case, as is the log-Sobolev constant of the chain. We write $\ell_0^2(\pi)$ for the subspace of functions f with $\mathbf{E}_{\boldsymbol{x} \sim \pi}[f(\boldsymbol{x})] = 0$. By the assumption that the chain is irreducible, L is invertible when restricted to $\ell_0^2(\pi)$; we write L^{-1} for its inverse on this subspace. For example, in the setting of the natural random walk on $\{0, 1\}^n$ we have

$$Lf = \sum_{S \subseteq [n]} |S|\widehat{f}(S)\chi_S, \qquad L^{-1}f = \sum_{\emptyset \neq S \subseteq [n]} \frac{\widehat{f}(S)}{|S|}\chi_S$$

for all $f \in \ell^2(\pi), \ell_0^2(\pi)$ respectively.

The generalization of Talagrand's inequality (5) involves certain gauge norms on $\ell^2(\pi)$. Let

$$M(t) = \frac{t^2}{\log(e + t^2)}, \qquad N'(t) = \frac{t^2\log(e + t^2)}{4},$$

which are easily verified to be Young functions. M and N' are not complementary but by [KR61, Lem. 7.2] we have that $M \sim M'$, where M' is the Young complement of N'. Let us also introduce the Young function

$$N(t) = t^2\log(1 + t^2),$$

for which it is easy to verify $N(t/2) \leq N'(t) \leq N(t)$ when $t \geq 1$, and hence $N \sim N'$. Bobkov and Götze [BG99] have shown that the norm $\|\cdot\|_N$ is closely related to entropy:

Proposition 3. *For any $f \in \ell_0^2(\pi)$ it holds that $\|f\|_N^2 \lesssim \sup_{c \in \mathbb{R}}\{\mathbf{Ent}[(f + c)^2]\} \lesssim \|f\|_N^2$.*

Note that

$$\sup_{c \in \mathbb{R}}\{\mathbf{Ent}[(f + c)^2]\} \leq 2\rho^{-1}\mathcal{E}[f] \tag{6}$$

follows immediately from the log-Sobolev inequality because $\mathcal{E}[f + c] = \mathcal{E}[f]$ for all constants c.

We now state and prove the generalization of Talagrand's inequality (5) to our setting of Markov chains:

Theorem 6. *For all $f \in \ell_0^2(\pi)$ it holds that $\langle f, L^{-1}f \rangle \lesssim \rho^{-1}\|f\|_M^2$.*

Proof. The result is trivial if $f = 0$. Otherwise,

$$\langle f, L^{-1}f \rangle^2 \lesssim \|f\|_{M'}^2 \|L^{-1}f\|_{N'}^2 \lesssim \|f\|_M^2 \|L^{-1}f\|_N^2 \lesssim \|f\|_M^2 \cdot \sup_{c \in \mathbb{R}}\{\mathbf{Ent}[(L^{-1}f + c)^2]\}$$

$$\lesssim \rho^{-1}\|f\|_M^2 \cdot \mathcal{E}[L^{-1}f] = \rho^{-1}\|f\|_M^2 \cdot \langle L^{-1}f, f \rangle,$$

as required, where the first inequality uses generalized Hölder, the second inequality uses $M' \sim M$ and $N' \sim M$, the third inequality uses Proposition 3, and the fourth inequality is the log-Sobolev inequality (or rather, (6)).

4.3 The Talagrand Theorem for Schreier graphs

We now return to our setting of Schreier graphs and prove Theorem 5, the generalization of the Talagrand Theorem. We begin with a simple calculation (cf. [KR61, (9.23)]):

Fact 7 *Let $f\colon X \to \{0,1\}$ be the indicator of a subset of measure τ. Then*

$$\|f\|_M \leq \sqrt{\frac{\tau}{\log(e + 1/\tau)}}.$$

Proof. Write $c = c(\tau) = \sqrt{\frac{\tau}{\log(e+1/\tau)}}$. If $\tau = 0$ then f is 0 almost everywhere and hence $\|f\|_M = 0 = c(0)$. Otherwise, we verify that $\int_\Omega M(f/c)\,d\mu \leq 1$:

$$\int_\Omega M(f/c)\,d\mu = \tau M(1/c) = \frac{\tau}{c^2 \log(e + 1/c^2)} = \frac{\log(e + 1/\tau)}{\log(e + \frac{\log(e+1/\tau)}{\tau})} \leq 1,$$

as $\log(e + 1/\tau) \geq 1$.

With this calculation in hand, we are able to deduce Theorem 5 from Theorem 6. The deduction is not quite as straightforward as in Talagrand's case, since our operators L_u are not self-adjoint.

Theorem 5 restated. *Let $\mathrm{Sch}(G, X, U)$ be a Schreier graph with log-Sobolev constant ρ. Assume that U is closed under conjugation. Then for all $f\colon X \to \{0,1\}$ it holds that*

$$\operatorname*{avg}_{u \in U} \left\{ \frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])} \right\} \gtrsim \rho \cdot \mathbf{Var}[f].$$

Proof. Given $u \in U$, let $g = L_u f$. Since $\boldsymbol{x}^u$ is uniformly distributed when $\boldsymbol{x}$ is, it follows that $\mathbf{E}_{\boldsymbol{x} \sim \pi}[g(\boldsymbol{x})] = 0$. Hence we may apply Theorem 6, obtaining

$$\langle g, L^{-1}g \rangle \lesssim \rho^{-1}\|g\|_M^2. \tag{7}$$

Since $|g|$ is the 0-1 indicator of a set of measure $2\mathcal{I}_u[f]$, we conclude from Fact 7 that

$$\|g\|_M^2 = \|\,|g|\,\|_M^2 \leq \frac{2\mathcal{I}_u[f]}{\log(e + 1/2\mathcal{I}_u[f])} \lesssim \frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])}$$

(using $0 \leq \mathcal{I}_u[f] \leq 1/2$). Thus from (7) we deduce

$$\frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])} \gtrsim \rho \cdot \langle L_u f, L^{-1} L_u f \rangle.$$

Since U is closed under conjugation, it follows that L_u commutes with L (see [OW09]). Hence L_u commutes with L^{-1} on $\ell_0^2(\pi)$ and we obtain

$$\langle L_u f, L^{-1} L_u f \rangle = \langle L_u f, L_u L^{-1} f \rangle = \langle f, L_u L^{-1} f \rangle + \langle f, L_{u^{-1}} L^{-1} f \rangle,$$

using Propositions 2 and 1. Thus

$$\frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])} \gtrsim \rho \cdot \left(\langle f, L_u L^{-1} f \rangle + \langle f, L_{u^{-1}} L^{-1} f \rangle \right).$$

Averaging over $u \in U$ and noting that $\operatorname{avg}_{u \in U}\{L_u\} = L = \operatorname{avg}_{u \in U}\{L_{u^{-1}}\}$ (because U is closed under inverses), we get

$$\operatorname*{avg}_{u \in U} \left\{ \frac{\mathcal{I}_u[f]}{\log(1/\mathcal{I}_u[f])} \right\} \gtrsim 2\rho \cdot \langle f, LL^{-1} f \rangle = 2\rho \cdot \langle f, f - \mathbf{E}[f] \rangle = 2\rho \cdot \mathbf{Var}[f],$$

completing the proof.

References

BG99.   Sergey Bobkov and Friedrich Götze. Exponential integrability and transportation cost related to logarith-
           mic Sobolev inequalities. *Journal of Functional Analysis*, 163(1):1–28, 1999.
BKK+92.   Jean Bourgain, Jeff Kahn, Gil Kalai, Yitzhak Katznelson, and Nathan Linial. The influence of variables
           in product spaces. *Israel Journal of Mathematics*, 77(1):55–64, 1992.
BL91.   Béla Bollobás and Imre Leader. Compressions and isoperimetric inequalities. *Journal of Combinatorial
           Theory, Series A*, 56(1):47–62, 1991.
BOL90.   Michael Ben-Or and Nathan Linial. Collective coin flipping. In Silvio Micali, editor, *Randomness and
           Computation*. Academic Press, New York, 1990.
Bon70.   Aline Bonami. Étude des coefficients Fourier des fonctions de $L^p(G)$. *Annals de l'institute Fourier*,
           20(2):335–402, 1970.
BT06.   Sergey Bobkov and Prasad Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *Journal
           of Theoretical Probability*, 19(2):289–336, 2006.
CKK+06.   Shuchi Chawla, Robert Krauthgamer, Ravi Kumar, Yuval Rabani, and D. Sivakumar. On the hardness of
           approximating Multicut and Sparsest-Cut. *Computational Complexity*, 15(2):94–114, 2006.
DKSV06.   Nikhil Devanur, Subhash Khot, Rishi Saket, and Nisheeth Vishnoi. Integrality gaps for sparsest cut and
           minimum linear arrangement problems. In *Proc. 38th ACM Symposium on Theory of Computing*, pages
           537–546, 2006.
DS05.   Irit Dinur and Samuel Safra. On the hardness of approximating minimum vertex cover. *Annals of Math-
           ematics*, 162(1):439–485, 2005.
DSC96.   Persi Diaconis and Laurent Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Annals
           of Applied Probability*, 6(3):695–750, 1996.
FK96.   Ehud Friedgut and Gil Kalai. Every monotone graph property has a sharp threshold. *Proceedings of the
           American Mathematical Society*, 124(10):2993–3002, 1996.
Fri98.   Ehud Friedgut. Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*,
           18(1):27–36, 1998.
Fri99.   Ehud Friedgut. Sharp thresholds of graph properties, and the k-SAT problem. *Journal of the American
           Mathematical Society*, 12(4):1017–1054, 1999.
Goe04.   Sharad Goel. Modified logarithmic Sobolev inequalities for some models of random walk. *Stochastic
           Processes and Their Applications*, 114(1):51–79, 2004.
GQ03.   Fuqing Gao and Jeremy Quastel. Exponential decay of entropy in the random transposition and Bernoulli–
           Laplace models. *Annals of Applied Probability*, 13(4):1591–1600, 2003.
Gro75.   Leonard Gross. Logarithmic Sobolev inequalities. *American Journal of Mathematics*, 97(4):1061–1083,
           1975.
KKL88.   Jeff Kahn, Gil Kalai, and Nathan Linial. The influence of variables on Boolean functions. In *Proceedings
           of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 68–80, 1988.
KR61.   Mark Krasnosel'skii and Yakov Rutickii. *Convex functions and Orlicz spaces*. P. Noordhoff Ltd., Groningen,
           1961.
KR08.   Subhash Khot and Oded Regev. Vertex Cover might be hard to approximate to within $2 - \epsilon$. *Journal of
           Computer and System Sciences*, 74(3):335–349, 2008.
KR09.   Robert Krauthgamer and Yuval Rabani. Improved lower bounds for embeddings into L_1. *SIAM Journal
           on Computing*, 38(6):2487–2498, 2009.
KZ06.   Jesse Kamp and David Zuckerman. Deterministic extractors for bit-fixing sources and exposure-resilient
           cryptography. *SIAM Journal on Computing*, 36(5):1231–1247, 2006.
LY98.   Tzong-Yau Lee and Horng-Tzer Yau. Logarithmic Sobolev inequality for some models of random walks.
           Annals of Probability, 26(4):1855–1873, 1998.
OS08.   Ryan O'Donnell and Rocco Servedio. The Chow Parameters problem. In *Proceedings of the 40th Annual
           ACM Symposium on Theory of Computing*, pages 517–526, 2008.
OW09.   Ryan O'Donnell and Yi Wu. 3-Bit dictator testing: 1 vs. 5/8. In *Proceedings of the 20th Annual ACM-SIAM
           Symposium on Discrete Algorithms*, pages 365–373, 2009.
Raz95.   Ran Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*,
           5(3):205–221, 1995.
RR91.   Malempati Rao and Zhong Ren. *Theory of Orlicz spaces*. Marcel Dekker, New York, 1991.
Tal94.   Michel Talagrand. On Russo's approximate zero-one law. *Annals of Probability*, 22(3):1576–1587, 1994.
Wu00.   Liming Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several
           applications. *Probability Theory and Related Fields*, 118(3):427–438, 2000.