Expander Random Walks: A Fourier-Analytic Approach

Gil Cohen (Tel Aviv University)

Based on joint works with

Noam Peri (TAU), **Amnon Ta-Shma** (TAU) (STOC'21)

Dor Minzer (MIT), **Shir Peleg** (TAU),
Aaron Potechin (U of Chicago), **Amnon Ta-Shma** (TAU)

September 20, 2021

Outline

Spectral Expanders

Let $G = (V, E)$ be a d-regular undirected graph on n vertices.

$$(\mathbf{W}_G)_{u,v} = \begin{cases} \frac{1}{d}, & uv \in E, \\ 0, & \text{otherwise.} \end{cases}$$

The eigenvalues of $\mathbf{W}_G$ are real, satisfying

$$-1 \leq \lambda_n \leq \cdots \leq \lambda_2 \leq \lambda_1 = 1.$$

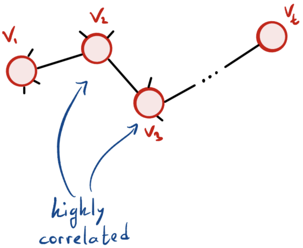
The spectral expansion of G is given by

$$\lambda(G) = \max(\lambda_2, -\lambda_n).$$

The smaller $\lambda(G)$ is the better. The "best" spectral expanders, dubbed Ramanujan graphs, satisfy

$$\lambda = \frac{2\sqrt{d-1}}{d}.$$

Expander random walks



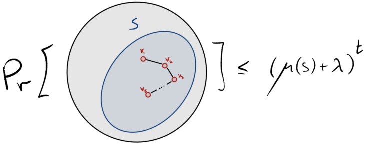
$$v_1 \sim V$$
$$v_2 \sim N(v_1)$$
$$\vdots$$
$$v_t \sim N(v_{t-1})$$

Meta question

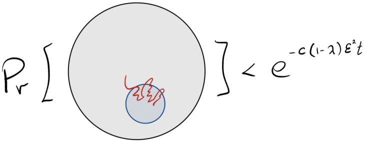
How random are random walks on expanders?

Some pseudorandom properties

The Expander hitting property (Ajtai-Komlós-Szemerédi'87)



The Expander Chernoff bound (AKS'87,Gillman'98,Healy'08)



$$\Pr_{(v_1,\ldots,v_t)\sim \mathrm{RW}}\left[\left|\frac{|\{v_1,\ldots,v_t\}\cap S|}{t}-\mu\right|\geq\varepsilon\right]\leq e^{-c(1-\lambda)\varepsilon^2 t}$$

Outline

Question

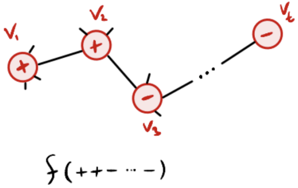
Question

What test functions are fooled by a random walk on expanders?

For $G = (V, E)$ and $\mathsf{val} : V \to \{\pm 1\}$ let $\mathsf{RW}_{G,\mathsf{val}} \in \{\pm 1\}^t$ be the distribution $\mathsf{val}(v_1), \ldots, \mathsf{val}(v_t)$ where $v_1, \ldots, v_t$ is a random walk in G.

Given $f : \{\pm 1\}^t \to \{\pm 1\}$ define

$$\mathcal{E}_{G,\mathsf{val}}(f) = \left| \mathbb{E}\left[f(\mathsf{RW}_{G,\mathsf{val}}) \right] - \mathbb{E}\left[f(\mathsf{val}(V)^t) \right] \right| .$$



Question

Recall

$$\mathcal{E}_{G,\mathrm{val}}(f) = \left|\mathbb{E}\left[f(\mathrm{RW}_{G,\mathrm{val}})\right] - \mathbb{E}\left[f(\mathrm{val}(V)^t)\right]\right|.$$

Definition

For λ, μ and $f : \{\pm 1\}^t \to \{\pm 1\}$, define

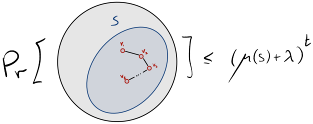
$$\mathcal{E}_{\lambda,\mu}(f) = \sup_{G,\mathrm{val}} \mathcal{E}_{G,\mathrm{val}}(f),$$

where

- $G = (V, E)$ ranges over all λ-spectral expanders; and
- val ranges over all valuations with $\mathbb{E}\left[\mathrm{val}(V)\right] = \mu$.

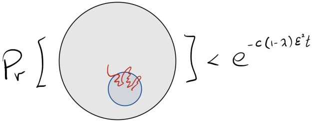
Pseudorandom properties revisited

The Expander hitting property (Ajtai-Komlós-Szemerédi'87)



$$\mathcal{E}_{\lambda,\mu}(\mathrm{AND}_t) \leq (\mu + \lambda)^t.$$

The Expander Chernoff bound (AKS'87,Gillman'98,Healy'08)



$$\mathcal{E}_{\lambda,\mu}\left(\mathbf{1}_{[(\mu-\varepsilon)t,(\mu+\varepsilon)t]}\right) \leq e^{-c(1-\lambda)\varepsilon^2 t}.$$

Other known results

CLT for expander random walks wrt CDF.

Theorem (Kipnis-Varadhan'86, Lezaud'01, Kloeckner'17)

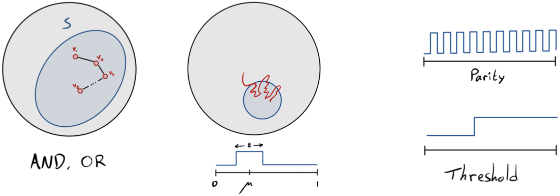
$$\mathcal{E}_{\lambda,\mu}\left(\sum_{i=1}^{t} x_i \geq k\right) = O\left(\frac{1}{(1-\lambda)^2\sqrt{t}}\right).$$

Perhaps most surprising is

Theorem (Ta-Shma'17 (see also Alon'93, Wigderson-Rozenman'04))

$$\mathcal{E}_{\lambda,\mu}\left(\text{Parity}\right) \leq (\mu + 2\lambda)^{t/2}.$$

What about other test functions?



What about other symmetric functions? Namely, CLT wrt TVD (raised independently by Guruswami-Kumar'20).

Beyond symmetric functions.

1. $\mathbf{AC^0}$ circuits.
2. Bounded-width (any order) read-once branching programs.
3. Low query complexity.

Question

Meta question

How random are random walks on expanders?

Question

What test functions are fooled by a random walk on expanders?

Question (formalized)

Given $f : \{\pm 1\}^t \to \{\pm 1\}$, bound $\mathcal{E}_{\lambda,\mu}(f)$.

Remarks.

- For simplicity, we only consider $\mu = 0$ and denote $\mathcal{E}_{\lambda,\mu}$ by $\mathcal{E}_{\lambda}$.
- To bound $\mathcal{E}_{\lambda}(f)$ it suffices to bound $\mathcal{E}_{G,\mathsf{val}}(f)$ for all relevant G, val. We fix G, val and write $\mathcal{E}$ for $\mathcal{E}_{G,\mathsf{val}}$.

Outline

Symmetric functions

Theorem

For every symmetric function $f : \{\pm 1\}^t \to \{\pm 1\}$,

$$\mathcal{E}_\lambda(f) = O(\lambda).$$

Further, the bound is tight for this class.

Theorem

$$\mathcal{E}_\lambda(\mathbf{1}_w) = O\left(\frac{\lambda}{\sqrt{t}}\right),$$

$$\mathcal{E}_\lambda(\text{Majority}) = O\left(\frac{\lambda^2}{\sqrt{t}}\right).$$

Beyond symmetric functions

Theorem (AC^0 circuits)

For every f that is computable by a size-s depth-d circuit,

$$\mathcal{E}_\lambda(f) = O\left(\sqrt{\lambda} \cdot (\log s)^{2(d-1)}\right).$$

This is essentially tight.

Theorem (Any order ROBP)

For every $f : \{\pm 1\}^t \to \{\pm 1\}$ that is computable by any order width-w ROBP,

$$\mathcal{E}_\lambda(f) = O\left(\sqrt{\lambda} \cdot (\log t)^{2w}\right).$$

For permutation ROBP, $\mathcal{E}_\lambda(f) = O\left(\sqrt{\lambda} \cdot w^4\right)$.

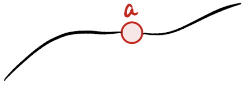
Outline

Our approach

Our Approach

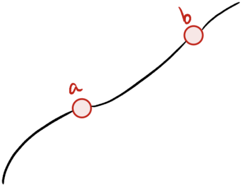
1. Bound $\mathcal{E}(\chi_S) = |\mathbb{E}[\chi_S(\mathrm{RW})]|$ for a general character
$$\chi_S(x_1, \ldots, x_t) = \prod_{i \in S} x_i$$
with $\emptyset \neq S \subseteq [t]$.
2. Expand f in the Fourier basis
$$f(x) = \sum_{S \subseteq [t]} \widehat{f}(S) \chi_S(x).$$
3. Conclude that
$$\mathcal{E}(f) \leq \sum_{\emptyset \neq S \subseteq [t]} \left|\widehat{f}(S)\right| \mathcal{E}(\chi_S).$$

Degree 1 characters



For every $a \in [t]$, v_a is uniformly distributed over V and so $\mathcal{E}(x_a) = 0$.

Degree 2 characters



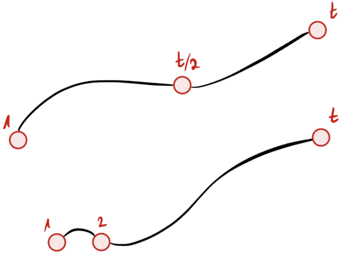
By the expander mixing lemma,

$$\mathcal{E}(x_a x_{a+1}) \leq \lambda.$$

The general case can be reduced to the above by considering G^{b-a}, and so

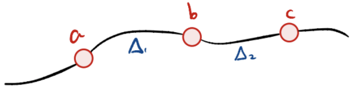
$$\mathcal{E}(x_a x_b) \leq \lambda^{b-a}.$$

Degree 3 characters



Test your intuition.

Degree 3 characters



Informally. A random step on a λ-spectral expander G can be thought of as sampling a vertex uniformly at random with probability $1 - \lambda$ and letting an adversary select a vertex with probability λ.

Formally.

$$\mathbf{W}_G = \mathbf{J} + \mathbf{E},$$

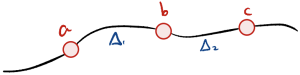
where $\mathbf{J} = (\frac{1}{n})_{i,j}$ and $\|\mathbf{E}\| = \max\{\|\mathbf{E}\mathbf{x}\|_2 : \|\mathbf{x}\|_2 = 1\} \leq \lambda$.

Informally.

- Except with probability λ^{Δ_1}, $b \mid a \sim V$.
- Independent of that, e.w.p λ^{Δ_2}, $c \mid b \sim V$.

Thus, e.w.p. $\lambda^{\Delta_1+\Delta_2}$, the bias is 0 and so $\mathcal{E}(x_a x_b x_c) \leq \lambda^{\Delta_1+\Delta_2}$.

Degree 3 characters



Formally. Let $\mathbf{1} = \left(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}\right)^{\mathsf{T}}$, $\mathbf{P} = \mathrm{diag}(\mathrm{val}(v))_{v \in V}$. Observe that

$$\mathbb{E}_{\mathsf{RW}}(\chi_{a,b,c}) = \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{W}^{\Delta_2} \mathbf{P} \mathbf{W}^{\Delta_1} \mathbf{P} \mathbf{1}.$$

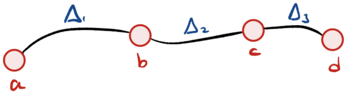
Now, $\mathbf{W}^{\Delta_i} = \mathbf{J} + \mathbf{E}_i$, with $\|\mathbf{E}_i\| \le \lambda^{\Delta_i}$, and so

$$\begin{aligned}
\mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{W}^{\Delta_2} \mathbf{P} \mathbf{W}^{\Delta_1} \mathbf{P} \mathbf{1} = \; & \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{J} \mathbf{P} \mathbf{J} \mathbf{P} \mathbf{1} + \\
& \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{J} \mathbf{P} \mathbf{E}_1 \mathbf{P} \mathbf{1} + \\
& \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{E}_2 \mathbf{P} \mathbf{J} \mathbf{P} \mathbf{1} + \\
& \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{E}_2 \mathbf{P} \mathbf{E}_1 \mathbf{P} \mathbf{1}.
\end{aligned}$$

As $\mathbf{J} \mathbf{P} \mathbf{1} = (\mathbf{1} \mathbf{1}^{\mathsf{T}}) \mathbf{P} \mathbf{1} = \mathbf{1}(\mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{1}) = \mathbf{0}$, we have

$$\mathbb{E}_{\mathsf{RW}}(\chi_{a,b,c}) = \mathbf{1}^{\mathsf{T}} \mathbf{P} \mathbf{E}_2 \mathbf{P} \mathbf{E}_1 \mathbf{P} \mathbf{1} \quad \Longrightarrow \quad \mathcal{E}(\chi_{a,b,c}) \le \|\mathbf{P} \mathbf{E}_2 \mathbf{P} \mathbf{E}_1 \mathbf{P}\| \le \lambda^{\Delta_1 + \Delta_2}.$$

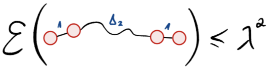
Degree 4 characters

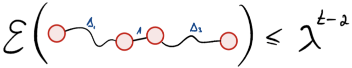


$$
\begin{aligned}
\mathbb{E}_{\mathrm{RW}}(\chi_{a,b,c,d}) &= \mathbf{1}^{\mathsf{T}}\mathbf{P}\mathbf{W}^{\Delta_3}\mathbf{P}\mathbf{W}^{\Delta_2}\mathbf{P}\mathbf{W}^{\Delta_1}\mathbf{P}\mathbf{1} \\
&= \mathbf{1}^{\mathsf{T}}\mathbf{P}\mathbf{E}_3\mathbf{P}\mathbf{W}^{\Delta_2}\mathbf{P}\mathbf{E}_1\mathbf{P}\mathbf{1}
\end{aligned}
$$

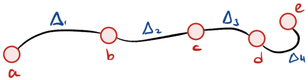
and so

$$
\mathcal{E}(\chi_{a,b,c,d}) \le \|\mathbf{P}\mathbf{E}_3\mathbf{P}\mathbf{W}^{\Delta_2}\mathbf{P}\mathbf{E}_1\mathbf{P}\| \le \lambda^{\Delta_1+\Delta_3}.
$$





Degree 5 characters



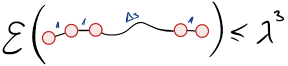
$$\mathbf{1}^{\mathsf{T}} \; P \; W^{\Delta_4} \; P \; W^{\Delta_3} \; P \; W^{\Delta_2} \; P \; W^{\Delta_1} \; P \; \mathbf{1}$$

E	J	J	E
E	J	E	E
E	E	J	E
E	E	E	E

$$\mathbf{JPJ} = (\mathbf{1}\mathbf{1}^{\mathsf{T}})\mathbf{P}(\mathbf{1}\mathbf{1}^{\mathsf{T}}) = \mathbf{1}(\mathbf{1}^{\mathsf{T}}\mathbf{P}\mathbf{1})\mathbf{1}^{\mathsf{T}} = \mathbf{0}$$

and so

$$\mathcal{E}(\chi_{a,b,c,d,e}) \leq \lambda^{\Delta_1+\Delta_2+\Delta_4} + \lambda^{\Delta_1+\Delta_3+\Delta_4} + \lambda^{\Delta_1+\Delta_2+\Delta_3+\Delta_4} \approx \lambda^{\Delta_1+\min(\Delta_2,\Delta_3)+\Delta_4}.$$



Degree 6 characters

$$\mathcal{E}(\chi_{a,b,c,d,e,f}) \lessapprox \lambda^{\Delta_1+\Delta_3+\Delta_5} + \lambda^{\Delta_1+\Delta_2+\Delta_4+\Delta_5}.$$

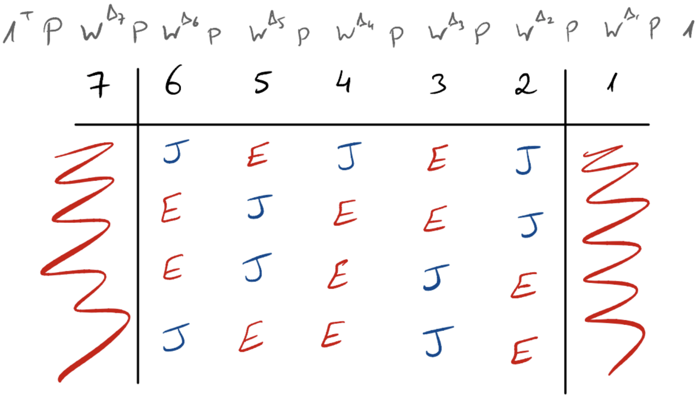
Degree 7 characters

$$\mathbb{1}^T P W^{\Delta_6} P W^{\Delta_5} P W^{\Delta_4} P W^{\Delta_3} P W^{\Delta_2} P W^{\Delta_1} P \mathbb{1}$$

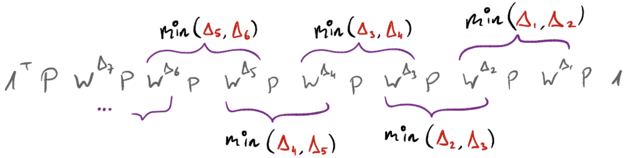
6	5	4	3	2	1
	E	J	E	J	
	J	E	E	J	
	J	E	J	E	

$$\mathcal{E} \lesssim \lambda^{\Delta_1+\Delta_6}\left(\lambda^{\Delta_3+\Delta_5} + \lambda^{\Delta_3+\Delta_4} + \lambda^{\Delta_2+\Delta_4}\right)$$

Degree 8 characters (this is addicting...)



The λ-tail



Definition

Let $S = \{s_1 < \cdots < s_k\} \subseteq [t]$, $k \geq 4$, and let $\Delta_i = s_{i+1} - s_i$. Define

$$\Delta(S) = \sum_{i=1}^{k-2} \min(\Delta_i, \Delta_{i+1}).$$

For $k = 2$, $\Delta(S) = \Delta_1$. For $k = 3$, $\Delta(S) = \Delta_1 + \Delta_2$.

Lemma

$$\mathcal{E}(\chi_S) \leq 2^{|S|} \cdot \lambda^{\frac{\Delta(S)}{2}}.$$

Beyond symmetric functions

Definition

A function $f : \{\pm 1\}^t \to \{\pm 1\}$ has b-Fourier decay if

$$\forall k \in [n] \quad \sum_{\substack{S \subseteq [t] \\ |S| = k}} |\widehat{f}(S)| \leq b^k.$$

Known Fourier decay results.

- For size-s depth-d circuits, $b = O((\log s)^{d-1})$ (Linial-Mansour-Nisan'93, Tal'17).
- For permutation width-w ROBP, $b \leq 2w^2$ (Reingold-Steinke-Vadhan'13).
- For general width-w ROBP, $b = O((\log t)^w)$ (Chattopadhyay-Hatami- Reingold-Tal'18).
- For every f, $b \leq \mathrm{DT}(f)$.

Beyond symmetric functions

Definition

A function $f : \{\pm 1\}^t \to \{\pm 1\}$ has b-Fourier decay if

$$\forall k \in [t] \quad \sum_{\substack{S \subseteq [t] \\ |S| = k}} |\widehat{f}(S)| \leq b^k.$$

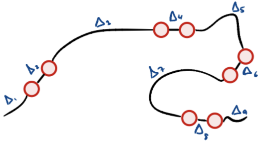
Lemma

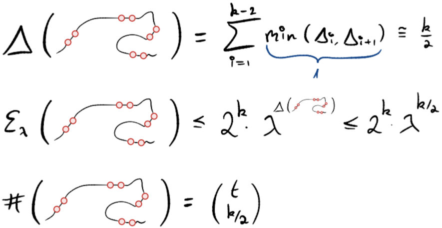
$$\mathcal{E}(f) = O(\sqrt{\lambda} \cdot b(f)^2).$$

Symmetric functions

Lemma (Main technical lemma)

$$\beta_k = \sum_{\substack{S\subseteq[t]\\|S|=k}} \mathcal{E}(\chi_S) \lessapprox 2^k \binom{t}{\lfloor \frac{k}{2} \rfloor} \lambda^{\lceil \frac{k}{2} \rceil}.$$





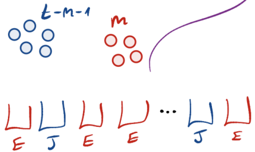
Symmetric functions

$$\beta_k = \sum_{\substack{S \subseteq [t] \\ |S| = k}} \mathcal{E}(\chi_S) = \sum_{\substack{S \subseteq [t] \\ |S| = k}} \underbrace{\left| 1^T P W^{\Delta_{k-1}} P \cdots W^{\Delta_1} P 1 \right|}_{\leq \; \sum_{\tau} \lambda^{\tau(S)}}$$

$$\tau \in \left\{ \begin{matrix} EJEJ \cdots EJE \\ EEJE \cdots EJE \\ \vdots \\ EJEE \cdots JEE \end{matrix} \right\}$$

$$\Rightarrow \quad \beta_k \leq \sum_{\tau} \sum_{\substack{S \subseteq [t] \\ |S| = k}} \lambda^{\tau(S)} = \sum_{\tau} \sum_{m} c_m(\tau) \lambda^m$$

τ :



Summary and open problems

Summary.

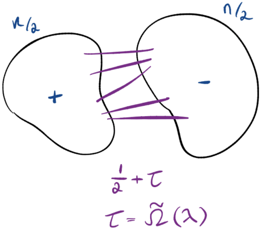
- Expander random walks fool all symmetric functions.
- Stronger Fourier decay $\implies$ better fooled by RW

Open problems.

- Tightness
- "Tailor-made" expanders
- Applications

Thanks!

The cross cut function



$$f(x_1, \ldots, x_t) = \text{Threshold}_{\frac{\tau}{2}t}\left(\sum_{i=1}^{t-1} x_i x_{i+1}\right).$$

We have that

$$|\mathbb{E}[f(U_t)]| \le e^{-\Omega(\tau^2 t)} \to 0,$$
$$|\mathbb{E}[f(\mathrm{RW})]| \ge 1 - e^{-\Omega(\tau^2 t)} \to 1.$$

Tightness

Lemma

Let $G = (V, E)$ be a Cayley graph on the Boolean hypercube and consider a labelling of V by an eigenvector corresponding to λ_2.

For $f : \{\pm 1\}^t \to \{\pm 1\}$ define $g : \{\pm 1\}^{2t} \to \{\pm 1\}$ by

$$g(x_1, \ldots, x_{2t}) = f(x_1 x_2, x_3 x_4, \ldots, x_{2t-1} x_{2t}).$$

Then,

$$\mathbb{E}[g(\mathrm{RW})] = (T_\lambda f)(\mathbf{1}).$$