Pseudorandomness of Expander Random Walks for Symmetric Functions and Permutation Branching Programs

Louis Golowich
Harvard University, Cambridge, MA, USA

Salil Vadhan
Harvard University, Cambridge, MA, USA

Abstract

We study the pseudorandomness of random walks on expander graphs against tests computed by symmetric functions and permutation branching programs. These questions are motivated by applications of expander walks in the coding theory and derandomization literatures. A line of prior work has shown that random walks on expanders with second largest eigenvalue λ fool symmetric functions up to a $O(\lambda)$ error in total variation distance, but only for the case where the vertices are labeled with symbols from a binary alphabet, and with a suboptimal dependence on the bias of the labeling. We generalize these results to labelings with an arbitrary alphabet, and for the case of binary labelings we achieve an optimal dependence on the labeling bias. We extend our analysis to unify it with and strengthen the expander-walk Chernoff bound. We then show that expander walks fool permutation branching programs up to a $O(\lambda)$ error in ℓ_2-distance, and we prove that much stronger bounds hold for programs with a certain structure. We also prove lower bounds to show that our results are tight. To prove our results for symmetric functions, we analyze the Fourier coefficients of the relevant distributions using linear-algebraic techniques. Our analysis for permutation branching programs is likewise linear-algebraic in nature, but also makes use of the recently introduced singular-value approximation notion for matrices (Ahmadinejad et al. 2021).

2012 ACM Subject Classification Theory of computation → Pseudorandomness and derandomization

Keywords and phrases Expander graph, Random walk, Pseudorandomness

Digital Object Identifier 10.4230/LIPIcs.CCC.2022.27

Related Version *Full Version*: `https://eccc.weizmann.ac.il/report/2022/024/` [9]

Funding *Louis Golowich*: Supported by Harvard College Herchel Smith Fellowship.
Salil Vadhan: Supported by NSF grant CCF-1763299 and a Simons Investigator Award.

Acknowledgements S.V. thanks Dean Doron, Raghu Meka, Omer Reingold, and Avishay Tal, conversations with whom inspired some of this research. We also thank the anonymous CCC reviewers for helpful comments that have improved the presentation.

1 Introduction

Random walks on expander graphs have numerous applications in computer science due to their pseudorandom properties (see e.g. [13] for a survey). Typically, an expander random walk is used to provide a randomness-efficient means for generating a sequence of vertices $v_0, \ldots, v_{t-1}$. In a given application, this expander walk will be used to "fool" certain desired test functions f, in the sense that the distribution of $f(v_0, \ldots, v_{t-1})$ is approximately the same whether the vertices $v_0, \ldots, v_{t-1}$ are sampled from a random walk on an expander, or independently and uniformly at random (which is equivalent to using a random walk on a complete graph with self loops). In this paper, we prove tight bounds on the extent to which expander graph random walks fool certain functions f of interest, namely, symmetric


© Louis Golowich and Salil Vadhan;
licensed under Creative Commons License CC-BY 4.0

functions as well as functions computable by permutation branching programs. These results improve on a recent line of work [11, 6, 5]. Our results also yield further implications, including a strengthening of the expander-walk Chernoff bound [7, 12].

An expander graph is a graph that is sparse but well connected. In this paper we consider regular λ-spectral expanders, which are constant-degree graphs for which all nontrivial eigenvalues of the random walk matrix have absolute value at most λ. Intuitively, the spectrum of an expander graph approximates that of the complete graph, so an expander provides a sparsification of the complete graph. Random walks on expander graphs therefore provide a derandomized approximation for random walks on complete graphs. A major aim of this paper is to obtain tight bounds on the error in this approximation.

Many explicit constructions of λ-spectral expanders are known for arbitrarily small $\lambda > 0$ (e.g. [18, 17, 19, 4]). Random walks on such expanders have many applications, such as in randomness-efficient error reduction, error-correcting codes, and small-space derandomization (see the surveys [13, 21, 10]). Randomness-efficient error reduction uses the ability of expander random walks to fool threshold functions, while Ta-Shma's recent breakthrough construction of ϵ-balanced codes [20] uses their ability to fool the parity function. Meanwhile, work on small-space derandomization starting from [15] uses the ability of expander walks to fool branching programs. In this paper, we prove new bounds on the extent to which expander walks fool symmetric functions (which include the threshold and parity functions), as well as (permutation) branching programs.

Specifically, we strengthen and generalize a result of Cohen et al. [5], which shows that a random walk on a sequence of λ-spectral expanders fools symmetric functions up to a $O(\lambda)$ error in total variation distance. Our result extends the result of Cohen et al. [5] to labelings of the vertices by symbols from an arbitrary alphabet and, in the binary case, achieves the optimal dependence on the bias of the labeling; the Cohen et al. [5] result only applies to binary labelings, and has a suboptimal dependence on the labeling bias. We also unite this total variation bound with a tail bound, which yields a strengthening of the expander-walk Chernoff bound. We furthermore show that expander random walks fool width-w permutation branching programs up to a $O(\lambda)$ error in ℓ_2-distance and a $O(\sqrt{w} \cdot \lambda)$ error in total variation distance, which extends a result of [2, 14] to walks of length > 2, and also strengthens the $O(w^4 \cdot \sqrt{\lambda})$ total variation bound of Cohen et al. [6]. For programs possessing a certain structure, we prove much stronger bounds. We also present several lower bounds that show our upper bounds to be tight.

The organization of the remainder of this extended abstract is as follows. Section 2 describes the main problem we consider, and introduces notation. Section 3 describes our contributions. We present proof outlines of our results for symmetric functions and for permutation branching programs in Section 4 and Section 5 respectively. For complete proofs of all results, the reader is referred to the full version of this paper [9].

2 Problem overview

For a sequence $\mathcal{G} = (G_1, \ldots, G_{t-1})$ of graphs on a shared vertex set V, let $\mathrm{RW}^t_{\mathcal{G}}$ denote the random variable taking values in V^t that is given by taking a length-t random walk on V, where the ith step is taken in the graph G_i. If all $G_i = G$ then we write $\mathrm{RW}^t_{\mathcal{G}} = \mathrm{RW}^t_G$.

For some fixed integer $d \geq 2$, we are given a labeling $\mathrm{val} : V \to [d] = \{0, \ldots, d-1\}$, which we extend to act on sequences componentwise, that is, $\mathrm{val}(v_0, \ldots, v_{t-1}) = (\mathrm{val}(v_0), \ldots, \mathrm{val}(v_{t-1}))$. We let the tuple $p = (p_0, \ldots, p_{d-1}) \in [0,1]^d$ specify the weights of the labels, so that p_b equals the fraction of vertices with label $b \in [d]$.

In this paper, we study the distribution of $\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}})$ for a sequence $\mathcal{G}$ of λ-spectral expanders. In particular, letting J denote the complete graph with self-loops, we will compare the distributions of $f(\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}))$ and $f(\mathrm{val}(\mathrm{RW}^t_J))$ for certain test functions f on $[d]^t$. Specifically, we study functions f that are either symmetric or computable by a permutation branching program.

Let $\Sigma : [d]^{[t]} \to [t+1]^{[d]}$ be the histogram function, so that $(\Sigma a)_b = |\{i \in [t] : a_i = b\}|$ denotes the number of copies of b in the sequence a. All symmetric functions factor through Σ, so to study symmetric functions we restrict attention to Σ.

3 Contributions

This section describes our main results. The reader is referred to the full version [9] for theorem statements containing explicit constants.

3.1 Symmetric functions

A major objective of this paper is to study the extent to which expander walks fool symmetric functions. In our notation, for a sequence $\mathcal{G}$ of λ-spectral expanders, we would like to bound the distance between the distributions of $\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}})$ and $\Sigma\,\mathrm{val}(\mathrm{RW}^t_J)$ as a function of λ, regardless of the choice of $\mathcal{G}$. Rather than directly comparing these distributions, in the following theorem we bound the change in $\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}})$ when one of the graphs G_u in the sequence $\mathcal{G}$ is changed. We then apply a hybrid argument by changing the graphs in $\mathcal{G}$ to J one at a time.

Thus the consideration of arbitrary expander sequences $\mathcal{G}$ is inherent in our proof. Yet as a side benefit, we are able to show fine-grained bounds on the distance between $\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}})$ and $\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}'})$ when $\mathcal{G}$ and $\mathcal{G}'$ only differ at a few steps. Such bounds are used in a follow-up work [8] to prove a new Berry-Esseen theorem for expander walks.

The following theorem considers the case of $d = 2$ possible labels; we will subsequently show a similar result for $d > 2$. In a slight abuse of notation below, we let G both denote a graph and its random walk matrix. We use $\|\cdot\|$ to denote the spectral norm of a matrix.

▶ **Theorem 1.** *Fix positive integers $u < t$. Let $\mathcal{G} = (G_i)_{1\le i\le t-1}$ and $\mathcal{G}' = (G'_i)_{1\le i\le t-1}$ be sequences of regular 1/100-spectral expanders on a shared vertex set V such that $G_i = G'_i$ for all $i \neq u$. Fix a labeling $\mathrm{val} : V \to [2]$ that assigns each label $b \in [2]$ to p_b-fraction of the vertices. Then for every $c \ge 0$,*

$$\sum_{j\in[t+1]:|j-p_1 t|\ge c} \left|\Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}'}) = (t-j,j)] - \Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}) = (t-j,j)]\right| = O\left(\frac{\|G'_u - G_u\| \cdot e^{-c^2/8t}}{t}\right).$$

Theorem 1 bounds the change in the distribution of $\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}})$ when the graph at a single step in $\mathcal{G}$ is changed. A key point is that the bound decays linearly in t. That is, the longer the walk, the less effect changing one of the graphs has. By changing all the graphs to the complete graph with self loops J one step at a time, we obtain the following corollary.

▶ **Corollary 2.** *For all positive integers t and all $0 \le \lambda \le 1/100$, let $\mathcal{G} = (G_i)_{1\le i\le t-1}$ be a sequence of regular λ-spectral expanders on a shared vertex set V with labeling $\mathrm{val} : V \to [2]$. Then for every $c \ge 0$,*

$$\sum_{j\in[t+1]:|j-p_1 t|\ge c} \left|\Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}) = (t-j,j)] - \Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_J) = (t-j,j)]\right| = O(\lambda \cdot e^{-c^2/8t}).$$

The bounds in Theorem 1 and Corollary 2 both provide unified bounds for two different notions of distance, namely total variation distance and tail bounds. Specifically, when $c = 0$ then the results above bound total variation distance, while as c grows large they provide tail bounds, as $p_1 t$ is the expected value of $(\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t))_1$.

Both the total variation and tail bounds above are novel, to the best of our knowledge. Our tails bounds can be viewed as strengthening the expander-walk Chernoff bound [7, 12], and indeed our proof of Theorem 1 draws on similar techniques as used in Healy's [12] proof of the expander-walk Chernoff bound. Recall that for a sequence $\mathcal{G}$ of λ-spectral expanders with λ bounded away from 1, the expander-walk Chernoff bound states that

$$\sum_{j \in [t+1] : |j - p_1 t| \geq c} \Pr[\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t) = (t-j, j)] = O(e^{-\Omega(c^2/t)}),$$

that is, the tails of $\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)$ decay approximately as quickly as the tails of the binomial distribution as $c^2/t \to \infty$. Corollary 2 shows the stronger statement that as $\lambda \to 0$, the tails of $\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)$ converge to the tails of the binomial distribution $\Sigma \operatorname{val}(\mathrm{RW}_J^t)$, even when $c^2/t = O(1)$.

The $c = 0$ case of Corollary 2 shows a $O(\lambda)$ bound on the total variation distance between $\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)$ and $\Sigma \operatorname{val}(\mathrm{RW}_J^t)$. Equivalently, this result shows that every symmetric function $f : \{0,1\}^t \to \{0,1\}$ satisfies $|\mathbb{E}[f(\operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t))] - \mathbb{E}[f(\operatorname{val}(\mathrm{RW}_J^t))]| = O(\lambda)$, that is, random walks on λ-spectral expanders $O(\lambda)$-fool symmetric functions. This bound improves upon a line of prior work [11, 6, 5]. Guruswami and Kumar [11] initiated this line of work by showing a $O(\lambda)$ bound on the total variation distance between $\Sigma \operatorname{val}(\mathrm{RW}_G^t)$ and $\Sigma \operatorname{val}(\mathrm{RW}_J^t)$ for the special case where G is the 2-vertex sticky random walk. Cohen et al. [6] then showed a $O(\lambda(\log(1/\lambda))^{3/2})$ bound on this total variation distance for arbitrary expanders G with a balanced labeling, that is, when $p_0 = p_1 = 1/2$. A follow-up paper of Cohen et al. [5] generalized to arbitrary p, and improved the total variation distance bound to $O(\lambda/\sqrt{\min(p)})$, where $\min(p) = \min\{p_0, p_1\}$. In contrast, the $c = 0$ case of Corollary 2 strengthens this bound to $O(\lambda)$ regardless of p. Our results also allow for sequences $\mathcal{G}$ of λ-spectral expanders with different graphs at different steps, whereas the prior work [6, 5] assumed that the graph was the same at each step.

Theorem 1, Corollary 2, and all of the prior work [11, 6, 5] assumes a binary labeling $\operatorname{val} : V \to \{0,1\}$ on the expander graph's vertices. Jalan and Moshkovitz [16] asked whether these results generalize to labelings $\operatorname{val} : V \to [d]$ for $d > 2$. We provide an affirmative answer to this question in the following results, which generalizing the total variation distance bounds in Theorem 1 and Corollary 2 to arbitrary $d \geq 2$. Below, we let $\min(p) = \min_{b \in [d]} p_b$.

▶ **Theorem 3.** *For every integer $d \geq 2$ and every distribution $p \in [0,1]^d$ over the labels $[d]$, there exists a constant $\lambda_0 = \lambda_0(d, p) > 0$ such that the following holds. For all positive integers $u < t$, let $\mathcal{G} = (G_i)_{1 \leq i \leq t-1}$ and $\mathcal{G}' = (G_i')_{1 \leq i \leq t-1}$ be sequences of λ_0-spectral expanders on a shared vertex set V, such that for all $i \neq u$ we have $G_i = G_i'$. Let $\operatorname{val} : V \to [d]$ be any labeling that assigns each label $b \in [d]$ to p_b-fraction of the vertices. Then*

$$d_{TV}\left(\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}'}^t), \Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)\right) = O\left(\left(\frac{d}{\min(p)}\right)^{O(d)} \cdot \frac{\|G_u' - G_u\|}{t}\right).$$

▶ **Corollary 4.** *For all integers $t \geq 1$ and $d \geq 2$, let $\mathcal{G} = (G_i)_{1 \leq i \leq t-1}$ be a sequence of λ-spectral expanders on a shared vertex set V with labeling $\mathrm{val} : V \to [d]$ that assigns each label $b \in [d]$ to p_b-fraction of the vertices. Then*

$$d_{TV}\left(\Sigma \,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}), \Sigma \,\mathrm{val}(\mathrm{RW}^t_J)\right) = O\left(\left(\frac{d}{\min(p)}\right)^{O(d)} \cdot \lambda\right).$$

In the results above, it is helpful to think of d and p as fixed, so that Corollary 4 gives a $O(\lambda)$ bound on total variation distance. When $d = 2$, Theorem 1 and Corollary 2 with $c = 0$ show that the factor of $(d/\min(p))^{O(d)}$ in the bounds above can be removed. We suspect that this $(d/\min(p))^{O(d)}$ dependence for $d > 2$ is not tight, and we leave the determination of the optimal dependence on d and p as an open question.

To show that the $O(\lambda)$ upper bounds on total variation distance described above are tight, we present the following lower bound.

▶ **Theorem 5.** *For every $0 < \lambda < 1$ and $p = (p_0, p_1)$, there exists a sufficiently large $t_0 = t_0(p, \lambda) \in \mathbb{N}$ and a λ-spectral expander $G = G_{\lambda,p}$ with vertex labeling $\mathrm{val} : V \to [2]$ that has label weights given by p, such that for every $t \geq t_0$,*

$$d_{TV}(\Sigma \,\mathrm{val}(\mathrm{RW}^t_G), \Sigma \,\mathrm{val}(\mathrm{RW}^t_J)) = \Omega(\lambda).$$

Theorem 5 generalizes a similar result of Guruswami and Kumar [11] for the special case of $p_0 = p_1 = 1/2$, and indeed our proof method is similar to theirs. Cohen et al. [5] showed a similar $\Omega(\lambda)$ lower bound for all t but only when $p_0 = p_1 = 1/2$. Their result is incomparable to ours, as Theorem 5 considers all p but only sufficiently large t.

The graph $G_{\lambda,p}$ in Proposition 8 is the λ-sticky, p-biased random walk, which generalizes the sticky walk studied by Guruswami and Kumar [11] for the case where $p = (1/2, 1/2)$. From a given vertex $v \in V$, with probability $1 - \lambda$ the sticky walk chooses the next vertex $v' \in V$ uniformly at random, and with probability λ it instead chooses a random v' that has the same label $\mathrm{val}(v') = \mathrm{val}(v)$. This sticky walk is in some sense a canonical λ-spectral expander, and arises in all of our lower bounds in this paper.

The main idea to prove Theorem 5 is that by the Markov chain CLT, as $t \to \infty$ then $((\Sigma \,\mathrm{val}(\mathrm{RW}^t_{G_{\lambda,p}}))_1 - p_1 t)/\sqrt{p_0 p_1 t}$ converges in distribution (that is, in Kolmogorov distance) to a normal distribution with variance $(1 + \lambda)/(1 - \lambda)$. In contrast, the CLT implies that the normalized binomial distribution $((\Sigma \,\mathrm{val}(\mathrm{RW}^t_J))_1 - p_1 t)/\sqrt{p_0 p_1 t}$ converges to a normal distribution with variance 1. Theorem 5 then follows because the distance between these two normals is $\Omega(\lambda)$. All the details are provided the full version [9].

3.2 Permutation branching programs

This section describes our main results on the extent to which expander walks fool permutation branching programs.

To begin, we recall the formal definition of a permutation branching program $\mathcal{B}$, which sequentially reads in inputs a_i and updates its internal state according to a permutation $B_i(a_i)$.

▶ **Definition 6.** *A **permutation branching program** $\mathcal{B}$ of length t, width w, and degree d is a collection of functions $B_i : [d] \times [w] \to [w]$ for $i \in [t]$ such that for $b \in [d]$, each restriction $B_i(b) = B_i|_{\{b\} \times [w]} : [w] \to [w]$ is a permutation. The program is said to **compute** the function $B : [d]^t \to [w]$ defined by*[1]

$$B(a) = (B_{t-1}(a_{t-1}) \circ \cdots \circ B_0(a_0))(0).$$

[1] Without loss of generality the initial state is assumed to be $0 \in [w]$.

We first present a bound that makes no assumptions on the structure of the program.

▶ **Theorem 7.** *For integers $t \geq 1$, $w \geq 2$, and $d \geq 2$, let G be a λ-spectral expander with $\lambda < .1$, and assign some vertex labeling* $\mathrm{val} : V \to [d]$*. Let $B : [d]^t \to [w]$ be computed by a permutation branching program $\mathcal{B}$ of length t, width w, and degree d. Then*

$$d_{\ell_2}(B(\mathrm{val}(\mathrm{RW}_G^t)), B(\mathrm{val}(\mathrm{RW}_J^t))) = O(\lambda).$$

Note that the bound in Theorem 7 has no dependence on the width w of the branching program, but only bounds ℓ_2 rather than total variation distance. Applying the Cauchy-Schwartz inequality to this ℓ_2-bound gives the total variation bound

$$d_{\mathrm{TV}}(B(\mathrm{val}(\mathrm{RW}_G^t)), B(\mathrm{val}(\mathrm{RW}_J^t))) = O(\sqrt{w} \cdot \lambda). \tag{1}$$

This bound improves upon the work of Cohen et al. [6], who showed a $O(w^4 \cdot \sqrt{\lambda})$ bound on $d_{\mathrm{TV}}(B(\mathrm{val}(\mathrm{RW}_G^t)), B(\mathrm{val}(\mathrm{RW}_J^t)))$ for the special case where $d = 2$ and $p_0 = p_1 = 1/2$.

Theorem 7 is closely related to the analysis of the Impagliazzo-Nisan-Wigderson [15] pseudorandom generator studied by Hoza et al. [14], which also uses expander walks to fool permutation branching programs. Both Theorem 7 and the results of Hoza et al. [14] are also proven using similar matrix approximation notions. However, Hoza et al. [14] consider many length-2 expander walks, whereas we consider a single longer walk.

The following lower bound shows that Theorem 7 is tight.

▶ **Proposition 8.** *For every $0 \leq \lambda \leq 1$ and every $p = (p_0, p_1)$, there exists a λ-spectral expander $G = G_{\lambda,p}$ with vertex labeling* $\mathrm{val} : V \to [2]$ *that assigns each label $b \in [2]$ to p_b-fraction of the vertices, such that the following hold:*

1. *There exists a permutation branching program $\mathcal{B}$ of length $t = 2$, width $w = 2$, and degree $d = 2$ such that*

$$\left|\Pr[B(\mathrm{val}(\mathrm{RW}_G^t)) = 0] - \Pr[B(\mathrm{val}(\mathrm{RW}_J^t)) = 0]\right| = 2p_0p_1\lambda.$$

2. *There exists a permutation branching program $\mathcal{B}$ of length $t = \lfloor 1/\min\{p_0, p_1\}\rfloor + 1$, width $w = t + 1$, and degree $d = 2$ such that*

$$\left|\Pr[B(\mathrm{val}(\mathrm{RW}_G^t)) = 0] - \Pr[B(\mathrm{val}(\mathrm{RW}_J^t)) = 0]\right| \geq \frac{\lambda}{2e^2}.$$

For these lower bounds, a smaller program length and width corresponds to a stronger result, as the length and width can be increased arbitrarily with padding. Proposition 8 implies a $\Omega(\lambda)$ lower bound for both the ℓ_2 and total variation distance between $B(\mathrm{val}(\mathrm{RW}_G^t))$ and $B(\mathrm{val}(\mathrm{RW}_J^t))$. This ℓ_2 lower bound meets the upper bound in Theorem 7. However, whereas the $\Omega(\lambda)$ total variation lower bound has no dependence on the program width w, the $O(\sqrt{w} \cdot \lambda)$ upper bound in (1) decays with w. It is an open question to resolve this gap.

The graph $G_{\lambda,p}$ in Proposition 8 is same the λ-sticky, p-biased random walk used to show Theorem 5, as described in Section 3.1. More details can be found in the full version [9].

Although Proposition 8 shows that Theorem 7 is tight in general, much stronger bounds hold for certain permutation branching programs.

▶ **Theorem 9.** *For integers $t \geq 1$, $w \geq 2$, and $d \geq 2$, let $\mathcal{G}$ be a sequence of λ-spectral expanders on a shared vertex set V with labeling* $\mathrm{val} : V \to [d]$*. Let $B^t : [d]^t \to [w]$ denote the sum modulo w, that is $B^t(a) = \sum_{i\in[t]} a_i \pmod w$. Then there exists a constant $c = c(d, w, p, \lambda) < 1$ such that*

$$d_{TV}(B^t(\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t)), B^t(\mathrm{val}(\mathrm{RW}_J^t))) \leq \sqrt{w} \cdot c^t.$$

That is, expander walks fool the small modular functions B^t, which are naturally computed by permutation branching programs, up to an exponentially small error. This result can be viewed as a generalization of the previously known fact that expander walks fool the parity function up to an exponentially small error, as can be recovered by letting $w = 2$ and $d = 2$ in Theorem 9. This fact that expander walks are good parity samplers played a pivotal role in Ta-Shma's breakthrough construction of almost optimal ϵ-balanced codes [20].

For arbitrary $w \geq 2$, Guruswami and Kumar [11] showed that the total variation distance between $B^t(\mathrm{val}(\mathrm{RW}^t_G))$ and $B^t(\mathrm{val}(\mathrm{RW}^t_J))$ is exponentially small in t when G is the 2-vertex sticky random walk. Theorem 9 generalizes this exponential decay bound to arbitrary expander walks.

Theorem 9 presents a particular class of permutation branching programs $\mathcal{B}$ for which $B(\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}))$ approaches a uniform distribution exponentially quickly. In the full version [9], we provide a more general class of such permutation branching programs $\mathcal{B}$, and deduce Theorem 9 as a special case. For illustrative purposes to avoid more cumbersome notation, we have omitted the more general case here.

4 Proof overview for symmetric functions

In this section, we outline the proof of Theorem 1, which contains many of the key technical insights in our paper. In particular, the proof of Theorem 3 follows the same general argument, so for the exposition in this section we focus on Theorem 1. All of the proof details can be found in the full version [9].

As in Theorem 1, for some $u < t$ let $\mathcal{G} = (G_i)_{1 \leq i \leq t-1}$ and $\mathcal{G}' = (G'_i)_{1 \leq i \leq t-1}$ be sequences of 1/100-spectral expanders that agree at all positions $i \neq u$, and again fix a vertex labeling $\mathrm{val} : V \to [2]$. Define $g \in [-1,1]^{[t+1]} \subseteq [-1,1]^{\mathbb{Z}}$ to be the difference between the probability mass functions of $(\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}'}))_1$ and $(\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}))_1$, that is,

$$g_j = \Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}'}) = (t-j, j)] - \Pr[\Sigma\,\mathrm{val}(\mathrm{RW}^t_{\mathcal{G}}) = (t-j, j)].$$

In this notation, the $c = 0$ case of Theorem 1 states that g has ℓ_1-norm $\|g\|_1 = O(\|G'_u - G_u\|/t)$, which is bounded by $O(\lambda/t)$ if G'_u and G_u are λ-spectral expanders.

We first show that the ℓ_2-norm of g satisfies

$$\|g\| = O\left(\frac{\|G'_u - G_u\|}{t} \cdot \frac{1}{(p_0 p_1 t)^{1/4}}\right). \tag{2}$$

The proof of this bound is sketched below in Section 4.1. We will then explain in Section 4.2 how to go from this ℓ_2-bound to the desired ℓ_1-bound. We compare our techniques to those of prior work in Section 4.3, and in particular we draw connections with Healy's proof of the expander-walk Chernoff bound [12].

4.1 Bounding the ℓ_2-distance $\|g\|$

In this section, we sketch the proof of the ℓ_2-bound (2). Because the Fourier transform preserves ℓ_2-norms, we will bound the ℓ_2-norm $\|\hat{g}\| = \|g\|$ of the Fourier transform $\hat{g}$ of g. Recall that here the Fourier transform is given by $\hat{g}(\theta) = \sum_{j \in \mathbb{Z}} e^{-i\theta j} g_j$, and has ℓ_2-norm $\|\hat{g}\| = \sqrt{\int_{\theta=-\pi}^{\pi} |\hat{g}(\theta)|^2 d\theta/2\pi}$.

To motivate this shift to the Fourier basis, recall that the Fourier transform interchanges convolution and multiplication, so that addition of independent random variables translates to multiplication of the Fourier transforms of their probability density functions (i.e. multiplication of their *characteristic functions*). Such products can be easier to analyze than convolutions, so the Fourier transform is a natural tool for analyzing sums of independent random variables, as is exemplified in proofs of the central limit theorem. Theorem 1 and Corollary 2 intuitively show that the expander walk distribution $(\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t))_1$ is close to the sum of independent variables, so it is also natural to analyze this distribution with the Fourier transform.

Whereas we apply the Fourier transform over the group $\mathbb{Z}$ to the random variable $\Sigma \operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)$ (which is distributed over $\mathbb{Z}$), the prior work of Cohen et al. [6] and Cohen et al. [5] applied the Fourier transform over the group $(\mathbb{Z}/2)^t$ to the random variable $\operatorname{val}(\mathrm{RW}_{\mathcal{G}}^t)$ (which is distributed over $\{0,1\}^t \cong (\mathbb{Z}/2)^t$). As described above, our approach seems well suited for symmetric functions, and it generalizes naturally to give Theorem 3 and Corollary 4 for alphabet sizes $d > 2$. In contrast, Cohen et al. [6] only consider $d = 2$, but they are able to apply their techniques to other classes of functions such as bounded-depth circuits, which we do not consider. More comparisons to prior techniques are provided in Section 4.3.

To begin, we express $\hat{g}(\theta)$ linear-algebraically. Specifically, let $\vec{1} = (1/\sqrt{|V|}, \dots, 1/\sqrt{|V|})$ denote the uniform unit vector, and define the diagonal matrix $P_\theta = \operatorname{diag}(x_\theta) \in \mathbb{C}^{V \times V}$, where $x_\theta \in \mathbb{C}^V$ is the vector with $(x_\theta)_v = e^{-i\theta(\operatorname{val}(v) - p_1)}$. Then it can be verified that

$$e^{i\theta p_1 t} \cdot \hat{g}(\theta) = \vec{1}^\top \left(\prod_{i=u+1}^{t} G_i P_\theta \right) (G'_u - G_u) \left(\prod_{i=0}^{u-1} P_\theta G_i \right) \vec{1},$$

where the products above multiply from right-to-left, and we take $G_0 = G_t = J$. This equality can be seen by expanding the right hand side above as a sum over all length-t walks $v_0, \dots, v_{t-1}$ on V. Therefore because $G'_u - G_u$ annihilates $\vec{1}$ from both sides, we have

$$|\hat{g}(\theta)| \le \left\| \left(\vec{1}^\top \left(\prod_{i=u+1}^{t} G_i P_\theta \right) \right)^\perp \right\| \cdot \|G'_u - G_u\| \cdot \left\| \left(\left(\prod_{i=0}^{u-1} P_\theta G_i \right) \vec{1} \right)^\perp \right\|, \tag{3}$$

where the notation $x^\perp$ denotes the projection of a vector x onto the orthogonal complement of $\vec{1}$. We will also use $x^\parallel$ to denote the projection of x onto $\vec{1}$.

We bound the rightmost factor above by induction on u. Splitting off a factor of $P_\theta G_{u-1}$ gives

$$\begin{aligned}
&\left\| \left(\left(\prod_{i=0}^{u-1} P_\theta G_i \right) \vec{1} \right)^\perp \right\| \\
&\quad \le \|(P_\theta \vec{1})^\perp\| \cdot \left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^\parallel \right\| + \|P_\theta\| \cdot \lambda(G_{u-1}) \cdot \left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^\perp \right\| \\
&\quad \le \|(P_\theta \vec{1})^\perp\| \cdot \left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^\parallel \right\| + \frac{1}{100} \cdot \left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^\perp \right\|,
\end{aligned} \tag{4}$$

where the last inequality follows because $\|P_\theta\| = 1$ and G_{u-1} is a 1/100-spectral expander. Thus if we can bound the first term on the right hand side of (4) by some $B(u)$ that decays less rapidly than 100^{-u} (i.e. $B(u) = \Theta(\beta^{-u})$ for $\beta < 100$), we can inductively bound the left

hand side by $B(u) + B(u-1)/100 + B(u-2)/100^2 + \cdots = O(B(u))$. Specifically, we will show this bound for $B(u) = \Theta(\sqrt{p_0 p_1} \cdot \theta \cdot e^{-\Omega(p_0 p_1 (u-1)\theta^2)})$. Intuitively, it suffices to bound what happens to the component parallel to $\vec{1}$, because the component orthogonal to $\vec{1}$ is shrunk by a factor of 100 with each application of $P_\theta G_i$.

Letting $F = J + (1/10)(I - J)$ be the matrix that preserves $\vec{1}$ and scales its orthogonal complement by $1/10$, then because by assumption all $i \neq u$ have $\lambda(G_i) \leq 1/100$, it follows that $\|F^{-1} G_i F^{-1}\| \leq 1$. Thus

$$\left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^{\|} \right\| = \left\| \vec{1}^\top \left(\prod_{i=0}^{u-2} F P_\theta F \cdot F^{-1} G_i F^{-1} \right) \vec{1} \right\| \leq \|F P_\theta F\|^{u-1}.$$

Next via some technical calculations, we show that for all $-\pi < \theta \leq \pi$,

$$\|(P_\theta \vec{1})^\perp\| = \frac{\|x_\theta^\perp\|}{\sqrt{|V|}} = \Theta(\sqrt{p_0 p_1} \cdot \theta). \tag{5}$$

For intuition, observe that if $p_0 p_1$ or θ equals 0, then all entries of x_θ are the same, so $x_\theta^\perp = 0$. Using (5), we also deduce that

$$\|F P_\theta F\| \leq 1 - \Omega((\|(P_\theta \vec{1})^\perp\|^2) = e^{-\Omega(p_0 p_1 \theta^2)}.$$

Here for intuition, as F is a 1/10-spectral expander, we should expect $\|F P_\theta F\|$ to be close to $\|J P_\theta J\| = \|(P_\theta \vec{1})^{\|}\| = \sqrt{1 - \|(P_\theta \vec{1})^\perp\|^2} = 1 - \Omega(\|(P_\theta \vec{1})^\perp\|^2)$. Thus (4) becomes

$$\left\| \left(\left(\prod_{i=0}^{u-1} P_\theta G_i \right) \vec{1} \right)^{\perp} \right\| \leq O\left(\sqrt{p_0 p_1} \cdot \theta \cdot e^{-\Omega(p_0 p_1 (u-1)\theta^2)} \right) + \frac{1}{100} \cdot \left\| \left(\left(\prod_{i=0}^{u-2} P_\theta G_i \right) \vec{1} \right)^{\perp} \right\|.$$

Recursively applying this inequality to bound the last term on its right hand side then gives

$$\left\| \left(\left(\prod_{i=0}^{u-1} P_\theta G_i \right) \vec{1} \right)^{\perp} \right\| = O\left(\sqrt{p_0 p_1} \cdot \theta \cdot e^{-\Omega(p_0 p_1 u \cdot \theta^2)} \right).$$

We now apply the above bound on $\|((\prod_{i=0}^{u-1} P_\theta G_i)\vec{1})^\perp\|$, along with an analogous bound on $\|(\vec{1}^\top (\prod_{i=u+1}^{t} G_i P_\theta))^\perp\|$, in (3) to give

$$|\hat{g}(\theta)| = O\left(p_0 p_1 \cdot \theta^2 \cdot e^{-\Omega(p_0 p_1 t \cdot \theta^2)} \cdot \|G_u' - G_u\| \right).$$

We then obtain the desired ℓ_2-bound (2) by squaring and integrating this bound with the substitution $q = c\sqrt{p_0 p_1 t} \cdot \theta$ for a sufficiently small constant $c > 0$:

$$\begin{aligned} \|g\| = \|\hat{g}\| &= O\left(p_0 p_1 \cdot \|G_u' - G_u\| \cdot \sqrt{\int_{-\pi}^{\pi} \theta^4 e^{-\Omega(p_0 p_1 t \cdot \theta^2)} \frac{d\theta}{2\pi}} \right) \\ &= O\left(\frac{\|G_u' - G_u\|}{t \cdot (p_0 p_1 t)^{1/4}} \cdot \sqrt{\int_{-\infty}^{\infty} q^4 e^{-q^2} dq} \right) \\ &= O\left(\frac{\|G_u' - G_u\|}{t \cdot (p_0 p_1 t)^{1/4}} \right). \end{aligned}$$

4.2 Going from an ℓ_2 to ℓ_1 bound

In this section, we show how to extend the techniques for bounding $\|g\|$ described above to bound $\|g\|_1$, and more generally to prove Theorem 1.

First observe that by the expander-walk Chernoff bound, $\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t)$ and $\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}'}^t)$ are mostly supported in an interval of length $\ell \approx O(\sqrt{t})$ about their mean. Applying the Cauchy-Schwartz inequality to (2) on this interval (which costs a factor of $\sqrt{\ell} \approx O(t^{1/4})$ to convert from ℓ_2 to ℓ_1), and the expander-walk Chernoff bound on the tails lying outside of the interval, yields a total variation bound of

$$\|g\|_1 = O\left(\frac{\|G'_u - G_u\|}{t} \cdot \left(\frac{\log(\|G'_u - G_u\|/t)}{p_0 p_1}\right)^{1/4}\right).$$

However, the above ℓ_1-bound does not help us prove Theorem 1 when c^2/t is large. Furthermore, even to prove the $c = 0$ case Theorem 1, we need to remove the factor $(\log(\|G'_u - G_u\|/t)/p_0p_1)^{1/4}$ from the bound above.

To obtain these improvements, we first generalize (2) to bound the ℓ_2-norm of the vector $g^{(sr)} = (e^{sr(j-p_1t)}g_j)_{j\in[t+1]}$ for $s = \pm 1$ and various values of $r \geq 0$. The proof of this bound on $\|g^{(sr)}\|$ for general r simply generalizes the argument presented in Section 4.1. The special case $r = 0$ recovers $g = g^{(0)}$, while when $r > 0$ then the sum of the elements of $g^{(sr)}$ equals the difference between the moment generating functions $\mathbb{E}[e^{sr((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t))_1 - p_1t)}]$ and $\mathbb{E}[e^{sr((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}'}^t))_1 - p_1t)}]$ that are used in the proofs of Chernoff bounds.

We then partition $[t+1] \subseteq \mathbb{Z}$ into intervals of length approximately $\sqrt{p_0p_1t}$, and we bound the ℓ_1-norm of g restricted to each interval by applying the Cauchy-Schwartz inequality with our ℓ_2-bound on $g^{(sr)}$ for appropriately chosen s, r. Summing these bounds over all intervals lying at least some distance c from p_1t yields Theorem 1.

Intuitively, as $\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t)$ and $\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}'}^t)$ have standard deviation $\Theta(\sqrt{p_0p_1t})$, we would expect these distributions to be somewhat evenly distributed across an interval of length $\sqrt{p_0p_1t}$. This is the regime where Cauchy-Schwartz is tight. Appropriately choosing s, r allows us to "isolate" a given length-$\sqrt{p_0p_1t}$ interval, by ensuring that the components of $g^{(sr)}$ in that interval dominate components outside that interval.

4.3 Comparison with techniques in prior work

Our techniques described above to prove Theorem 1 are closely related to Healy's [12] proof of the expander-walk Chernoff bound. In some sense, Healy's proof [12] makes up "half" of our proof: Healy's proof bounds the moment-generating function $\mathbb{E}[e^{sr((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t))_1 - p_1t)}]$, but does not bound the characteristic function $\mathbb{E}[e^{-i\theta((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t))_1 - p_1t)}]$ as described in Section 4.1 (as the Fourier coefficient $\hat{g}(\theta)$ by definition equals the difference between the characteristic functions of $(\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t))_1$ and $(\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}'}^t))_1$). Intuitively, our proof combines the moment generating and characteristic function bounds, as in order to bound $\|g^{(sr)}\|$, we bound the difference $e^{i\theta p_1 t} \cdot \hat{g}^{(sr)}(\theta)$ between the generating functions $\mathbb{E}[e^{(sr-i\theta)((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t))_1 - p_1t)}]$ and $\mathbb{E}[e^{(sr-i\theta)((\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}'}^t))_1 - p_1t)}]$.

Although Cohen et al. [6] and Cohen et al. [5] also studied the extent to which expander walks fool symmetric functions, their proofs are less similar to ours. Most notably, both of these papers use Fourier analysis over the group $(\mathbb{Z}/2\mathbb{Z})^t$ by viewing $\mathrm{val}(\mathrm{RW}_G^t)$ as a distribution on $(\mathbb{Z}/2\mathbb{Z})^t$. In contrast, we use Fourier analysis over $\mathbb{Z}^{d-1}$ by viewing $\Sigma\,\mathrm{val}(\mathrm{RW}_{\mathcal{G}}^t)$ as a distribution on $\mathbb{Z}^d$ (or $\mathbb{Z}^{d-1}$, if we drop the first component). This explains why our results generalize more naturally to the case $d > 2$, which is not considered in [6, 5]. We

could also do our analysis using discrete Fourier analysis over $(\mathbb{Z}/m)^{d-1}$ instead, for any $m \geq t$, but then the modulus m is superfluous (as it is cleaner to avoid modular reduction) and only makes the notation more cumbersome.

5 Proof overview for permutation branching programs

In this section, we outline the proof of Theorem 7, which uses singular-value approximations as described below. We do not outline the proof of our other results for permutation branching programs, specifically Theorem 9, and instead refer the reader to the full version [9], as this latter result uses techniques somwhat similar to those described above in Section 4.

5.1 Proof outline

We now describe the proof of Theorem 7. As in the theorem statement, for arbitrary integers $t \geq 1$, $w \geq 2$, and $d \geq 2$, let $\mathcal{B}$ be a permutation branching program of length t, width w, and degree d that computes some function $B : [d]^t \to [w]$. Let G be a λ-spectral expander with $\lambda < .1$, and assign some vertex labeling $\mathrm{val} : G \to [d]$. We again let $g \in [-1,1]^{[w]}$ denote the difference between the distributions $B(\mathrm{val}(\mathrm{RW}_G^t))$ and $B(\mathrm{val}(\mathrm{RW}_J^t))$ of interest, that is,

$$g_j = \Pr[B(\mathrm{val}(\mathrm{RW}_G^t)) = j] - \Pr[B(\mathrm{val}(\mathrm{RW}_J^t)) = j].$$

In this notation, Theorem 7 states that $\|g\| = O(\lambda)$.

As in the proof of Theorem 1, we begin by expressing g linear-algebraically. Let $\tilde{P}$ be the operator on the vector space $\mathbb{R}^V \otimes \mathbb{R}^t \otimes \mathbb{R}^w$ given by

$$\tilde{P} = \sum_{v \in V, i \in [t]} \delta_v \delta_v^\top \otimes \delta_{i+1} \delta_i^\top \otimes B_i(\mathrm{val}(v)),$$

where $i+1$ is taken $(\mathrm{mod}\ t)$ above, and by abuse of notation $B_i(\mathrm{val}(v)) \in \mathbb{R}^{w \times w}$ refers to the permutation matrix associated to the permutation $B_i(\mathrm{val}(v)) : [w] \to [w]$. Also for $W = G$ or J, let $\tilde{W} = W \otimes I \otimes I$. Then for every $j \in [w]$,

$$g = (\vec{1} \otimes \delta_0 \otimes I)^\top ((\tilde{G}\tilde{P})^t - (\tilde{J}\tilde{P})^t)(\vec{1} \otimes \delta_0 \otimes \delta_0). \tag{6}$$

This equality can again be seen by expanding the right hand side above as a sum over all length-t walks on V.

We will bound the right hand side using singular-value approximations [1, 3]. A matrix $W' \in \mathbb{C}^{N \times N}$ is a singular-value ϵ-approximation of another matrix $W \in \mathbb{C}^{N \times N}$, written $W' \overset{\mathrm{sv}}{\approx}_\epsilon W$, if for all $x, y \in \mathbb{C}^N$,

$$|x^*(W' - W)y| \leq \frac{\epsilon}{2}(\|x\|^2 + \|y\|^2 - \|x^*W\|^2 - \|Wy\|^2),$$

where x^* denotes the conjugate transpose of x. The following properties were shown by Ahmadinejad et al. [1, 3]:

1. $\tilde{G} \overset{\mathrm{sv}}{\approx}_\lambda \tilde{J}$.

Assume that $W' \overset{\mathrm{sv}}{\approx}_\epsilon W$. Then:

2. For every matrix X with spectral norm $\|X\| \leq 1$, then $W'X \overset{\mathrm{sv}}{\approx}_\epsilon WX$.
3. If $\epsilon < .1$, then $(W')^t \overset{\mathrm{sv}}{\approx}_{\epsilon+5\epsilon^2} W^t$. (Importantly, the bound $\epsilon + O(\epsilon^2)$ does not grow with t.)
4. $\|W' - W\| \leq \epsilon$.

We now bound the right hand side of (6) using singular value approximations. Because by definition $\|\tilde{P}\| = 1$, property 1 and property 2 above imply that $\tilde{G}\tilde{P} \overset{\text{sv}}{\approx}_{\lambda} \tilde{J}\tilde{P}$. Then property 3 implies that $(\tilde{G}\tilde{P})^t \overset{\text{sv}}{\approx}_{\lambda+5\lambda^2} (\tilde{J}\tilde{P})^t$, and property 4 then gives that $\|(\tilde{G}\tilde{P})^t - (\tilde{J}\tilde{P})^t\| \leq \lambda + 5\lambda^2$, so $\|g\| \leq \lambda + 5\lambda^2 = O(\lambda)$.

5.2 Comparison with techniques in prior work

The proof of Theorem 7 described above is closely related to the analysis of the Impagliazzo-Nisan-Wigderson (INW) [15] pseudorandom generator in Hoza et al. [14]. Hoza et al. [14] use unit-circle approximations [2] to show that length-2 walks on λ-spectral expanders fool permutation branching programs up to a $O(\lambda)$ ℓ_2-error; the INW generator they study recursively applies many such length-2 walks. We generalize this $O(\lambda)$ bound to walks of arbitrary length, and simplify the analysis by replacing the unit-circle approximations with singular-value approximations. We obtain these improvements because the unit-circle approximations, though similar in nature to singular-value approximations, do not satisfy property 2 described above. Although our results do not directly translate to an improved pseudorandom generator, it is an interesting question whether longer walks could somehow be used to improve the seed length.

As described in Section 3.2, Theorem 7 implies a $O(\sqrt{w} \cdot \lambda)$ total variation distance bound, which improves upon the $O(w^4 \cdot \sqrt{\lambda})$ total variation bound of Cohen et al. [6]. However, Cohen et al. [6] prove their result using bounds on the Fourier tails over $(\mathbb{Z}/2\mathbb{Z})^t$ of permutation branching programs with alphabet size $d = 2$, differing significantly from our proof using singular-value approximations, which generalizes readily to $d > 2$.

References

1. AmirMahdi Ahmadinejad. *Computing stationary distributions: perron vectors, random walks, and ride-sharing competition.* PhD thesis, Stanford University, Stanford, California, 2020.
2. AmirMahdi Ahmadinejad, Jonathan Kelner, Jack Murtagh, John Peebles, Aaron Sidford, and Salil Vadhan. High-precision Estimation of Random Walks in Small Space. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1295–1306, November 2020. ISSN: 2575-8454. doi:10.1109/FOCS46700.2020.00123.
3. AmirMahdi Ahmadinejad, John Peebles, Aaron Sidford, and Salil Vadhan. Personal Communication, 2021.
4. Avraham Ben-Aroya and Amnon Ta-Shma. A Combinatorial Construction of Almost-Ramanujan Graphs Using the Zig-Zag Product. *SIAM Journal on Computing*, 40(2):267–290, January 2011. doi:10.1137/080732651.
5. Gil Cohen, Dor Minzer, Shir Peleg, Aaron Potechin, and Amnon Ta-Shma. Expander Random Walks: The General Case and Limitations. *Electronic Colloquium on Computational Complexity*, 2021. URL: https://eccc.weizmann.ac.il/report/2021/091/.
6. Gil Cohen, Noam Peri, and Amnon Ta-Shma. Expander random walks: a Fourier-analytic approach. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, pages 1643–1655, New York, NY, USA, June 2021. Association for Computing Machinery. doi:10.1145/3406325.3451049.
7. David Gillman. A Chernoff Bound for Random Walks on Expander Graphs. *SIAM Journal on Computing*, 27(4):1203–1220, August 1998. doi:10.1137/S0097539794268765.
8. Louis Golowich. A Berry-Esseen Theorem for Expander Walks. *Forthcoming*, 2022.
9. Louis Golowich and Salil Vadhan. Pseudorandomness of Expander Random Walks for Symmetric Functions and Permutation Branching Programs. *Electronic Colloquium on Computational Complexity*, 2022. URL: https://eccc.weizmann.ac.il/report/2022/024/.

10 Venkatesan Guruswami. Guest column: error-correcting codes and expander graphs. *ACM SIGACT News*, 35(3):25–41, September 2004. doi:10.1145/1027914.1027924.

11 Venkatesan Guruswami and Vinayak M. Kumar. Pseudobinomiality of the Sticky Random Walk. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 48:1–48:19, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2021.48.

12 Alexander D. Healy. Randomness-Efficient Sampling within NC1. *computational complexity*, 17(1):3–37, April 2008. doi:10.1007/s00037-007-0238-5.

13 Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. doi:10.1090/S0273-0979-06-01126-8.

14 William M. Hoza, Edward Pyne, and Salil Vadhan. Pseudorandom Generators for Unbounded-Width Permutation Branching Programs. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 7:1–7:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2021.7.

15 Russell Impagliazzo, Noam Nisan, and Avi Wigderson. Pseudorandomness for network algorithms. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of Computing*, STOC '94, pages 356–364, New York, NY, USA, May 1994. Association for Computing Machinery. doi:10.1145/195058.195190.

16 Akhil Jalan and Dana Moshkovitz. Near-Optimal Cayley Expanders for Abelian Groups. *arXiv:2105.01149 [cs]*, May 2021. arXiv:2105.01149v1.

17 A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, September 1988. doi:10.1007/BF02126799.

18 Grigorii Aleksandrovich Margulis. Explicit constructions of expanders. *Problemy Peredachi Informatsii*, 9(4):71–80, 1973. Publisher: Russian Academy of Sciences, Branch of Informatics, Computer Equipment and Automatization.

19 Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders. *The Annals of Mathematics*, 155(1):157, January 2002. doi:10.2307/3062153.

20 Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 238–251, Montreal Canada, June 2017. ACM. doi:10.1145/3055399.3055408.

21 Salil P. Vadhan. Pseudorandomness. *Foundations and Trends® in Theoretical Computer Science*, 7(1–3):1–336, December 2012. doi:10.1561/0400000010.