

# $L_1$ AND $L_2$ EMBEDDINGS OF THE SYMMETRIC GROUP

COSMAS KRAVARIS

**ABSTRACT.** We show that the Cayley graph of the symmetric group  $Sym_n$  generated by the cycle  $(123\dots n)$  and the transposition  $(12)$  embeds into  $L_1$  with bi-Lipschitz distortion  $O(1)$ . This answers a question of Ostrovskii, and along with Kassabov’s theorem gives the first example of a sequence of groups which embed bi-Lipschitzly into  $L_1$  for one choice of bounded size generating sets, but not for another choice of bounded size generating sets. In particular, the Cayley graphs generated by the cycle and the transposition cannot contain coarsely any unbounded sequence of expander graphs. Moreover, within the context of the Ribe program, they are a new example of bounded degree Cayley graphs which are test spaces for Rademacher type.

## 1. Introduction

The **rank** of a finitely generated group  $G$ , denoted by  $rank(G)$ , is the smallest size of a generating set of  $G$ . Given a group  $G$  with a finite generating set  $S$ , denote by  $c_1(G, d_S)$  the bi-Lipschitz distortion of the shortest path metric  $d_S$  of the Cayley graph  $\Gamma(G, S)$  into  $L_1$  (see the background below).

We start with the following observation and fundamental question, shown to us by Naor. A sequence of finite groups  $\{G_n\}_n$  of bounded rank,  $\sup_n rank(G_n) < \infty$ , tautologically exhibits one of the following three behaviors:

**Never- $L_1$  behavior:** For any sequence of generating sets  $\{S_n \subset G_n\}_n$  with  $\sup_n |S_n| < \infty$ ,

$$\sup_n c_1(G_n, d_{S_n}) = \infty.$$

**Always- $L_1$  behavior:** For any sequence of generating sets  $\{S_n \subset G_n\}_n$  with  $\sup_n |S_n| < \infty$ ,

$$\sup_n c_1(G_n, d_{S_n}) < \infty.$$

**Mixed- $L_1$  behavior:** There exist two sequences of generating sets  $\{S_n \subset G_n\}_n$  and  $\{S'_n \subset G_n\}_n$  with  $\sup_n |S_n| < \infty$  and  $\sup_n |S'_n| < \infty$  such that

$$\sup_n c_1(G_n, d_{S_n}) < \infty \quad \text{and} \quad \sup_n c_1(G_n, d_{S'_n}) = \infty.$$

The existence of a sequence of groups with Never- $L_1$  behavior follows from a theorem of Breuillard and Gamburd [8], and the well-known fact that expander graphs do not embed bi-Lipschitzly into  $L_1$  [26]. It was shown to us by Naor that bounded rank abelian groups exhibit Always- $L_1$  behavior (see Proposition 2 in Section 7). See Section 7 for the  $L_2$ -counterparts of these behaviors.

**The main result of this paper, see Theorem 1 below, provides the first example of a sequence of groups with Mixed- $L_1$  behavior.**

**Question 1** (Naor). *Classify (sequences of) groups  $\{G_n\}$  according to the above trichotomy. Namely, obtain a useful structural characterization of when they exhibit never, always or mixed  $L_1$  behavior.*

In particular, Naor highlighted the following concrete particular case of Question 1, which could be quite tractable and enlightening to resolve: into which part of the above trichotomy do the lamplighter groups  $\{\mathbb{Z}/2\mathbb{Z} \wr \mathbb{Z}/n\mathbb{Z}\}_n$  fall?

A landmark result of Kassabov [22] states that there exists a sequence of generating sets  $T_n \subset \text{Sym}_n$  of the symmetric groups of bounded size,  $\sup_n |T_n| < \infty$ , such that the Cayley graphs  $\{\Gamma(\text{Sym}_n, T_n)\}_n$  form a family of expanders. In light of Kassabov's theorem, Ostrovskii asked the following question [41] which was subsequently posted in [37].

**Question 2** (Ostrovskii). *Do there exist  $C \in (0, \infty)$  and  $r \in \mathbb{N}$  and generating sets  $T_n \subset \text{Sym}_n$  each of size  $|T_n| \leq r$  such that for each  $n$ , the symmetric group with word metric  $(\text{Sym}_n, d_{T_n})$  embeds into  $L_1$  with bi-Lipschitz distortion  $\leq C$ ?*

We give a positive answer to Question 2. Combined with Kassabov's theorem (and the well-known fact that bounded degree expander graphs do not embed bi-Lipschitzly into  $L_1$  [26]), we conclude that the symmetric groups  $\{\text{Sym}_n\}_{n=1}^\infty$  exhibit Mixed- $L_1$  behavior.

**Theorem 1** (Cycle and transposition live in  $L_1$ ). .

*For any  $n \in \mathbb{N}$ , the symmetric group  $\text{Sym}_n$  on  $\mathbb{Z}/n$  endowed with the word metric generated by the transposition  $t := (01)$  and the cycle  $c := (0123\dots n-1)$  embeds into  $L_1$  with bi-Lipschitz distortion  $O(1)$ . Specifically, we have*

$$\sup_n c_1(\text{Sym}_n, d_{\{t,c\}}) < 1000.$$

The constant 1000 is not optimized.

It is well-known that  $\Gamma(\text{Sym}_n, \{t, c\})$  is not an expander (e.g. see Remark 2 on page 510 in [4] or Section 11.4 in the survey [20]). Theorem 1 implies that  $\Gamma(\text{Sym}_n, \{t, c\})$  is **very far** from being an expander in the following strong sense.

**Corollary 1** (Coarse opposite to Kassabov's theorem). .

*The sequence of Cayley graphs  $\{\Gamma(\text{Sym}_n, \{t, c\})\}_n$  cannot contain coarsely an unbounded sequence of bounded degree expanders. That is, given two increasing functions  $\alpha, \beta : [0, \infty) \rightarrow [0, \infty)$  with  $\lim_{t \rightarrow \infty} \alpha(t) = \infty$ , and a finite graph  $\Gamma$  with (normalized) spectral gap  $1 - \lambda$  and maximum degree  $\Delta$ , if there exists  $f : \Gamma \rightarrow \text{Sym}_n$  with*

$$\alpha(d_\Gamma(u, v)) \leq d_{\{c,t\}}(f(u), f(v)) \leq \beta(d_\Gamma(u, v)) \quad \text{for all } u, v \in V(\Gamma)$$

*(where  $d_\Gamma$  denotes the shortest path metric on  $\Gamma$  and  $d_{\{c,t\}}$  the word metric on  $\text{Sym}_n$ ), then the number of vertices in  $\Gamma$  is bounded:*

$$|V(\Gamma)| \leq \Delta^{\alpha^{-1}(C\beta(1)/(1-\lambda))}$$

*where  $C > 0$  is a universal constant.*

The proof that Theorem 1 implies Corollary 1 follows from Gromov [18] and it is a standard argument nowadays (see Theorem 5.7 and the proof of Theorem 4.9 in [40]). Note that the dependence on the max-degree  $\Delta$  is necessary:  $\{\Gamma(Sym_n, \{t, c\})\}_n$  contain the Hamming cubes bi-Lipschitz which are unbounded degree spectral expanders (see Subsection 1.2 and Section 8).

**1.1. Background.** The **L<sub>1</sub> bi-Lipschitz distortion** of a metric space  $(\mathcal{M}, d)$ , denoted by  $c_1(\mathcal{M})$ , is the infimal  $D > 1$  for which there exists a map  $f : \mathcal{M} \rightarrow L_1$  with

$$d(x, y) \leq \|f(x) - f(y)\|_1 \leq D d(x, y) \quad \text{for all } x, y \in \mathcal{M}.$$

Understanding the smallest possible distortion  $D > 1$  for embedding a metric space is a central question in the theory of metric embeddings (see for instance [29, 40, 32, 12, 20]). The question is equivalent to bi-Lipschitz approximation by a measured walls structure which has been fruitfully investigated in the context of geometric group theory (see for example Prop. 2.6 in [10], [9], Chapter 6 in [15] and references within).

For a Banach space  $X$ , the bi-Lipschitz distortion  $c_X(\mathcal{M})$  is defined analogously. When  $X = L_2$  we write  $c_2(\mathcal{M}) := c_{L_2}(\mathcal{M})$ .  $L_2$ -distortion will be discussed in Section 7. (We remark that for finite metric spaces, whether the target space is  $L_1([0, 1])$  or  $l_1(\mathbb{N})$  makes no difference. For every  $\epsilon > 0$ , any finite dimensional subspace of one  $L_1$ -space is  $(1 + \epsilon)$ -isomorphic to some finite dimensional subspace of the other  $L_1$ -space. The same remark holds for  $L_2$ .)

**Notation:** For each  $n \in \mathbb{N}$  let  $Sym_n$  be the symmetric group on  $\mathbb{Z}/n = \{0, 1, 2, \dots, n - 1\}$ ,  $t = (01)$  be the transposition of the first two elements and  $c = (0123\dots n - 1)$  be the cyclic permutation. We consider the (left) Cayley graphs of the symmetric groups  $\Gamma(Sym_n, \{t, c\})$  with vertices  $Sym_n$  and edges  $\{(t\pi, \pi), (c\pi, \pi) : \pi \in Sym_n\}$  and view  $Sym_n$  as a metric space endowed with the shortest path metric (that is, the word metric). We denote the metric by  $d(\cdot, \cdot)$  and the distance to the identity (or word length) by  $|\cdot|$ . We will write:  $S := \{t, c\}$  for the generating set. For each  $k, l \in \mathbb{Z}/n$  we denote by  $d_{\mathbb{Z}/n}(k, l)$  the distance between  $k$  and  $l$  on the Cayley graph of  $\mathbb{Z}/n$  generated by  $\{+1\}$ . Finally, the product of two permutations is read from right to left, as in function composition. This means that, for instance,  $(12)(23) = (123)$  and  $(23)(12) = (132)$ .

**Remark:** It is well-known that if we do not insist that the generating sets are of bounded size, then it is easy to embed the symmetric groups into  $L_1$ . For each  $n$ , take the generating set of all transpositions. Map  $f : Sym_n \rightarrow l_1([n]^2)$  by mapping each permutation  $\pi$  to its associated permutation matrix  $A_\pi$ . Whenever two permutations differ by a transposition, their difference in the image is exactly 2. By the triangle inequality, this gives the upper bound  $\|A_\pi - A_\tau\|_1 \leq 2d(\pi, \tau)$  for any  $\pi, \tau \in Sym_n$ . For the lower bound, observe that  $\|A_\pi - A_\tau\|_1 = \|Id - A_{\tau\pi^{-1}}\|_1$  so it suffices to show  $\|Id - A_\pi\|_1 \geq d(1, \pi)$ . Observe that by our choice of generating set,  $d(1, \pi) = \#\{k \in [n] : \pi(k) \neq k\} - \#\{\text{cycles of } \pi\}$  and also  $\|Id - A_\pi\|_1 = 2\#\{k \in [n] : \pi(k) \neq k\}$ ; so the lower bound also follows.

## 1.2. New example of bounded degree Cayley test spaces for Rademacher type. .

In this subsection and in Section 8 we discuss an application of Theorem 1 to the geometry of Banach spaces. Those who are not interested in Banach spaces can skip these parts. .

A central aspect of the Ribe program [32, 5] aims to provide metric characterizations of local

properties of Banach spaces (i.e. properties that depend only on the linear structure of the finite dimensional subspaces). An important early result is a metric characterization of Rademacher type<sup>1</sup> through test spaces: a Banach space  $X$  has trivial Rademacher type if and only if  $X$  contains the Hamming cubes with uniformly bounded bi-Lipschitz distortion, i.e.  $\sup_n c_X((\{0, 1\}^n, \|\cdot\|_1)) < \infty$ . The "only if" direction follows from Pisier [42]. For the "if" direction note the folklore result that for every  $0 < \epsilon < 1$  and every finite subset  $F \subset L_1$ ,  $F$  embeds into  $\{0, 1\}^n$  with distortion  $1 + \epsilon$  for some  $n \in \mathbb{N}$  (this follows from a standard approximation argument; see for instance the proof of Lemma 5.4 in [31]). By Ribe's theorem [44] (see also [6] for quantitative bounds), taking an  $\epsilon$ -net of the unit ball of any finite dimensional subspace of  $L_1$  we see that  $L_1$  is crudely finitely representable in  $X$ , hence  $X$  has trivial Rademacher type. (Almost sharp quantitative bounds for this metric characterization of type follow from [7] and truly sharp bounds follow from [21].)

A sequence of metric spaces  $((M_n, d_n))_n$  is a **family of test spaces** for a Banach space  $Z$  whenever for all Banach spaces  $X$  we have that  $Z$  is finitely representable in  $X$  if and only if  $\sup_n c_X((M_n, d_n)) < \infty$  [40]. Ostrovskii [36, 35] showed that for any Banach space  $Z$  there exists a family of 3-regular graphs whose shortest-path metrics form a family of test spaces for  $Z$ . However, these test spaces are not vertex-transitive, even for  $L_1$ .

A theorem of Naor and Peres [33] states that the lamplighter groups  $\mathbb{Z}/n \wr \mathbb{Z}/n$  with the standard generators (either move the pointer by a unit or edit the current lamp by a unit) embed into  $L_1$  with uniformly bounded bi-Lipschitz distortion. On the other hand, Arzhantseva, Guba and Sapir [2] showed that every Hamming cube  $\{0, 1\}^n$  embeds into  $\mathbb{Z}/m \wr \mathbb{Z}/m$  for some  $m$  with uniformly bounded bi-Lipschitz distortion. It follows that the sequence of 4-regular Cayley graphs of  $(\mathbb{Z}/n \wr \mathbb{Z}/n)_n$  form test spaces for Rademacher type. As an application of Theorem 1, we have a new example of bounded degree Cayley test spaces of Rademacher type which are 3-regular and are Cayley graphs of the symmetric groups.

**Proposition 1** (Bounded degree Cayley test spaces for Rademacher type). .

*A Banach space  $X$  has trivial Rademacher type if and only if*

$$\sup_n c_X(Sym_n, d_{\{t, c\}}) < \infty.$$

### 1.3. $L_2$ -embeddings of Cayley graphs into Hilbert space. .

A conjecture of Cornulier-Tessera-Valette [11] states that a finitely generated group whose word metric embeds bi-Lipschitzly into Hilbert space must have an abelian subgroup of finite index. In Section 7, we state a version of this conjecture due to Naor about finite groups, and prove that a positive answer to the conjecture of Cornulier-Tessera-Valette implies a positive answer to the conjecture of Naor. One should view Section 7 as the Euclidean counterpart to Question 1. This section also contains several results and conjectures due to Naor which appear for the first time with his permission.

---

<sup>1</sup>Rademacher type is a fundamental invariant of the local geometry of a Banach spcae (see [30]). We do not need to recall it here in this paper, because we will not use it.

#### 1.4. Overview of the proof of Theorem 1 and paper organization.

In Section 2 we begin with an explicit formula for the word metric of  $(\text{Sym}_n, \{c, t\})$  up to constant factors. The formula is inspired by the word metric formula of Yadin for the lampshuffler groups [45]. The *lampshuffler group*  $\text{Sym}_{00}(\mathbb{Z}) \times \mathbb{Z}$  consists of all pairs  $(\pi, x)$  where  $\pi$  is a finite support permutation on  $\mathbb{Z}$  and  $x \in \mathbb{Z}$  is a pointer. Roughly speaking, this is a variation of the lamplighter on  $\mathbb{Z}$ , but the generator moves are: just moving the pointer and moving the pointer while transposing the two elements along the movement. In fact, it is not hard to show that the lampshuffler group  $\text{Sym}_{00}(\mathbb{Z}) \times \mathbb{Z}$  embeds bi-Lipschitzly into  $L_1$  (see [25]). In the symmetric group  $\text{Sym}_n$ , the cycle  $c$  plays the role of the "pointer" and  $t$  takes the role of the transposition generator (though no knowledge of lamplighter nor lampshufflers is required for the proof).

The **key difficulty** in analyzing the word metric on  $\text{Sym}_n$  as opposed to  $\text{Sym}_{00}(\mathbb{Z}) \times \mathbb{Z}$  (or even  $\text{Sym}_{00}(\mathbb{Z}/n) \times \mathbb{Z}/n$ ) is that **the pointer (or frame of reference) is ambiguous**. For example, whenever  $\pi = c^2tctc^{-4}$  then it is obvious that the "pointer" should be  $-1$ , but for a more complicated permutation this is not clear. The first step of the proof is to show that the distance between two permutations  $\pi$  and  $\tau$  is given by:

$$|\tau\pi^{-1}| \asymp \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right)$$

where  $d_{\mathbb{Z}/n}$  denotes the distance on the  $n$ -cycle  $\mathbb{Z}/n$ . Observe that we have to take the minimum over all potential "pointer positions"  $l \in \mathbb{N}$ .

**The presence of this minimum makes the metric difficult to analyze.**

In Section 3 we split the minimum of the sum of two terms into the sum of two minimums:

$$|\tau\pi^{-1}| \asymp \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) =: T_1 + T_2$$

This is achieved by high-low casework. Whenever  $T_1 < n/3$ , for some  $l$  there are at most  $n/3$  nonzero terms in the sum, so the "potential pointer" is obvious, and the same value  $l$  will be the optimal one in all three minimums. Whenever  $T_1 > n/3$ , the first term dominates. This is because we always have  $T_2 \leq n/2$  (since it is a diameter) and hence:  $T_1 \leq T_2 \leq (1 + 3/2)T_2$ . The rest of the proof deals with embedding each term separately.

The first term,  $\min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k))$ , is a metric on  $\text{Sym}_n$  and can be viewed as a subset of the abelian group  $[n-1]^{\mathbb{Z}/n}$  (where addition is pointwise addition of functions  $f : [n-1] \rightarrow \mathbb{Z}/n$ ) with word metric given by the generating set  $\{(1, 0, \dots, 0), (0, 1, 0, \dots, ), \dots, (0, \dots, 0, 1), (1, 1, \dots, 1)\}$ . We follow the embedding method of Austin-Naor-Valette [3] (see also section 4 in Naor-Peres [33]), and write the embedding into  $L_1$  using the representation theory of abelian groups (that is, exponential sums). We interpret

$$\min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) = \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - \tau(k), l)$$

as the minimum sum of distances of the cloud of points  $\{\pi(k) - \tau(k)\}$  to a "median" point  $l \in \mathbb{Z}/n$ . It is well-known (e.g. see [43]) that this quantity is, up to a factor of 2, the average distance of the cloud of points. Based on this observation, we construct the embedding via exponential sums.

(Sidenote: similar to how [3] and [33] proceed, the embedding was found by reverse-engineering: first searching through all possible  $L_2$ -representations of this abelian group and pointing out which representations will not work.)

The second term,  $\min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\})$  is a variation of the lamplighter metric, but in which the lamplighter has "forgotten his position". It was shown by Naor-Peres [33] that the lamplighter metric embeds into  $L_1$ . (Later Ostrovskii-Randrianantoanina [38] gave a different embedding which works more generally into any non-superreflexive Banach space.) Here we modify the second embedding given in Naor-Peres by identifying/collapsing certain specific coordinates of  $l_1(\mathbb{N})$ . Whenever the first term is small, e.g.  $T_1 < n/3$ , the optimal "pointer"  $l$  is obvious, and the coordinate identifications do not "influence" the lower bound of the embedding.

In Section 6 we combine the estimates for each of the two terms and prove Theorem 1.

In Section 7 we discuss  $L_2$ -embeddings of Cayley graphs into Hilbert space.

Finally, in Section 8 we prove the test space characterization for  $L_1$ . The only missing step is showing that the Hamming cube embeds into the Cayley graph of a cycle and a transposition. The embedding is similar to that of Arzhantseva, Guba and Sapir [2] for  $\mathbb{Z}/n \wr \mathbb{Z}/n$ , and the analysis uses the word metric formula in Section 2.

## 2. The word metric of cycle and transposition

We begin by slowly examining the word length of various types of permutations.

**Adjacent transpositions:** Via conjugating by cyclic permutations, we can obtain any transposition between adjacent numbers:

$$c^k tc^{-k} = (k(k+1)) \quad \text{for all } k \in \mathbb{Z}/n.$$

**General transpositions:** Any transposition between non-adjacent numbers can be obtained by applying transpositions between adjacent numbers. For example,  $(13) = (12)(23)(12)$ . More generally, for any  $k \in \mathbb{Z}/n$ ,  $(k(k+2)) = (k(k+1))((k+1)(k+2))(k(k+1))$ . Substituting the expression of  $(k(k+1))$  in terms of our generators, we observe *cancellation* between the conjugation exponents:

$$(k(k+2)) = c^k tc^{-k} c^{k+1} tc^{-k-1} c^k tc^{-k} = c^k t c t c^{-1} t c^{-k}.$$

We now want to express a general transposition as the product of generators. For every  $2 \leq l \leq \lfloor n/2 \rfloor + 1$  we have:

$$(0l) = (01)(12)\dots((l-2)(l-1))((l-1)l)((l-2)(l-1))\dots(12)(01) = (tc)^{l-1} t (tc)^{-(l-1)}$$

while for  $\lfloor n/2 \rfloor + 1 \leq l \leq n-1$  we have a shorter expression since  $(0(n-1)) = c^{-1} tc =: t'$

$$(0l) = (t' c^{-1})^{l-1} t' (t' c^{-1})^{-(l-1)}.$$

To obtain a general transposition  $(k(k+l))$  we simply conjugate the transposition  $(0l)$ . We have the word length estimates:

$$|(0l)|_S \leq 4d_{\mathbb{Z}/n}(0, l) \quad \text{and} \quad |(k(k+l))|_S \leq 4d_{\mathbb{Z}/n}(0, l) + 2d_{\mathbb{Z}/n}(0, k).$$

**Cyclic permutations:** Take a cyclic permutation  $(k_1 k_2 \dots k_m)$  where  $m \in [n]$  and  $k_1, \dots, k_m \in \mathbb{Z}/n$  are distinct. We write the cyclic permutation as the product of transpositions and observe that there is *cancellation* between the conjugation exponents:

$$(k_1 \dots k_m) = (k_1 k_2)(k_2 k_3) \dots (k_{m-2} k_{m-1})(k_{m-1} k_m) \\ = c^{k_1}(0(k_2 - k_1))c^{-k_1}c^{k_2}(0(k_3 - k_2))c^{-k_2} \dots c^{k_{m-1}}(0(k_m - k_{m-1}))c^{-k_{m-1}} = c^{k_1} \left( \prod_{i=1}^{m-1} (0(k_{i+1} - k_i))c^{k_{i+1} - k_i} \right) c^{-k_m}.$$

By the triangle inequality we arrive at the estimate:

$$|(k_1 \dots k_m)|_S \leq d_{\mathbb{Z}/n}(0, k_1) + d_{\mathbb{Z}/n}(0, k_m) + 5 \sum_{i=1}^{m-1} d_{\mathbb{Z}/n}(k_i, k_{i+1}) \leq 2d_{\mathbb{Z}/n}(0, k_1) + 6 \sum_{i=1}^{m-1} d_{\mathbb{Z}/n}(k_i, k_{i+1})$$

The following estimate on the word length of a permutation is a variation of a lemma due to Yadin about the lampshuffler groups (in the paper titled "Rate of Escape of the Mixer Chain").

**Lemma 1.** *For each permutation  $\pi \in Sym_n$ , the word length is given by*

$$|\pi|_S \asymp \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(k, \pi(k) + l) + diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi(p) \neq p - l\}) \right).$$

To be more precise, we have:

$$\frac{1}{3} \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(k, \pi(k) + l) + diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi(p) \neq p - l\}) \right) \\ \leq |\pi|_S \leq \min_{l \in \mathbb{Z}/n} \left( 6 \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(k, \pi(k) + l) + 2diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi(p) \neq p - l\}) \right).$$

*Proof.* **Upper bound (construction)** We will show the upper bound for  $l = 0$ . For any  $l \in \mathbb{Z}/n$ , the bound follows by applying the formula to  $c^l \pi$ . Write  $\pi$  as the product of disjoint cyclic permutations, say

$$\pi = \prod_{j=1}^s c^{l_j} \gamma_j c^{-l_j}$$

where for each  $j = 1, \dots, s$ ,  $l_j \in \mathbb{Z}/n$  and  $\gamma_j = (0 k_2^{(j)} k_3^{(j)} \dots k_{m_j}^{(j)})$  for some  $m_j \in \mathbb{N}$  and  $k_2^{(j)}, \dots, k_{m_j}^{(j)} \in \mathbb{Z}/n$  all distinct and different from 0. After relabeling the indices, we may assume that the path  $0 \mapsto l_1 \mapsto l_2 \mapsto \dots \mapsto l_s$  is a path of length at most twice the diameter:

$$d_{\mathbb{Z}/n}(0, l_1) + \sum_{j=2}^s d_{\mathbb{Z}/n}(l_{j-1}, l_j) \leq 2diam_{\mathbb{Z}/n}(\{0\} \cup \{k : \pi(k+l) \neq k\}).$$

(Why twice? It would be the case that 0 sits at the middle of the cloud of points  $\{l_1, \dots, l_s\}$ .)

We now write:

$$\pi = \prod_{j=1}^s c^{l_j} \gamma_j c^{-l_j} = c^{l_1} \gamma_1 \prod_{j=2}^s c^{-l_{j-1} + l_j} \gamma_j$$

and the upper bound follows from the triangle inequality and the estimate on the word length of a cycle.

**Lower bound** We show that the right hand side (*RHS* in short) does not change much when we apply a generator:

When we multiply  $\pi$  by the cyclic permutation, the *RHS* changes by at most 1 unit.

When we multiply  $\pi$  by a transposition, the *RHS* can change by at most 3 units.

For the identity permutation,  $RHS = 0$ . Moving across the shortest path from  $\pi$  to 1, the sum drops to zero, and hence the number of steps to reach the identity is at least

$$|\pi|_S \geq \frac{1}{3} \min_{l \in \mathbb{Z}/n} \left( d_{\mathbb{Z}/n}(0, l) + \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(k, \pi(k+l)) + \text{diam}_{\mathbb{Z}/n}(\{0\} \cup \{k : \pi(k+l) \neq k\}) \right),$$

so we get our estimate. ■

Applying the formula to the permutation  $\tau\pi^{-1}$  we get:

$$\begin{aligned} |\tau\pi^{-1}| &\asymp \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(k, \tau\pi^{-1}(k) + l) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \tau\pi^{-1}(p) \neq p - l\}) \right) \\ &= \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right) \end{aligned}$$

### 3. Splitting the minimum into two terms

**Lemma 2.** *For any  $\pi, \tau \in \text{Sym}_n$ , we have*

$$\begin{aligned} &\min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right) \\ &\leq 2 \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \end{aligned}$$

*Proof.* Note that the direction  $\geq$  is trivial (when we drop the constants 4 and 3). The point of this claim is to show the reverse non-trivial direction. We will write:

$$T_1 := \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)), \quad T_2 := \min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\})$$

**CASE A:**  $T_1 = \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) \geq n/2$

Observe that always  $T_2 \leq n$ , so the second term is of lower order compared to the first term, and we get the trivial estimate:

$$\begin{aligned} &\min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right) \\ &\leq T_1 + n/2 \leq T_1 + T_1 \leq 2T_1 + T_2. \end{aligned}$$

**CASE B:**  $T_1 = \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) < n/2$

Note that in this case  $|\{k : \pi(k) - l \neq \tau(k)\}| < n/2$  for any value of  $l$  which attains this minimum. Also, via a change of variables:

$$|\{k : \pi(k) - l \neq \tau(k)\}| = |\{k : k - l \neq \tau(\pi^{-1}(k))\}| = |\{k : \tau^{-1}(k - l) \neq \pi^{-1}(k)\}|,$$

so for any other value  $l' \neq l$  we get that

$$\begin{aligned} \text{diam}_{\mathbb{Z}/n}(\{0, l'\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l')\}) &\geq |\{k : \tau^{-1}(k - l') \neq \pi^{-1}(k)\}| \\ &\geq |\{k : \tau^{-1}(k - l) = \pi^{-1}(k)\}| > n/2. \end{aligned}$$

We have two further subcases:

**CASE B1:** The minimum of  $T_2$  is attained at the same value  $l$  as the minimum of  $T_1$ .

In this case the claim is trivial (since the left-hand side equals  $T_1 + T_2$ ).

**CASE B2:** The minimum of  $T_2$  is attained at a different value  $l' \neq l$  than the minimum of  $T_1$ .

In this case, as we saw above,  $T_2 > n/2$  and we get:

$$\begin{aligned} T_1 + T_2 &> \frac{n}{2} + T_1 = \frac{n}{2} + \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) \\ &\geq \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right). \end{aligned}$$

(Aside: since the resulting inequality is strict, CASE B2 can never happen.) ■

#### 4. Embedding the 1st term: distance to the median

**Lemma 3** (Embedding the 1st term). .

There exists a map  $\Phi_1 : \text{Sym}_n \rightarrow L_1$  such that for all  $\pi, \tau \in \text{Sym}_n$  we have

$$\|\Phi_1(\pi) - \Phi_1(\tau)\|_1 \asymp \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)).$$

In particular, the bi-Lipschitz distortion is  $< 2\pi \approx 6.28$ .

The following observation is well-known (e.g. see [43]).

**Observation 1.** For any metric space  $(X, d)$  and any  $x_1, \dots, x_n \in X$

$$\frac{1}{2} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j) \leq \min_{1 \leq r \leq n} \frac{1}{n} \sum_{j=1}^n d(x_r, x_j) \leq \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j).$$

*Proof.* The second inequality is trivial since the minimum of a list of numbers is always  $\leq$  the average. For the first inequality, for all  $r \in [n]$  we use the triangle inequality:

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d(x_i, x_j) \leq \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (d(x_r, x_j) + d(x_r, x_i)) = \frac{2}{n} \sum_{j=1}^n d(x_r, x_j)$$
■

**Observation 2.** (on the circle every cloud contains one of its median points) .

For any  $n, m \in \mathbb{N}$  and  $x_1, \dots, x_m \in \mathbb{Z}/n$  we have:

$$\min_{1 \leq r \leq m} \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, x_r) = \min_{l \in \mathbb{Z}/n} \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, l).$$

*Proof.* Let  $l$  be a minimizer for  $\min_{l \in \mathbb{Z}/n} \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, l)$ . If  $l \in \{x_1, \dots, x_m\}$  then we are done, so suppose not. Consider the antipodal point  $\bar{l} := \lfloor n/2 \rfloor + l$  and the two intervals  $[l, \bar{l}]$  and  $[\bar{l}, l]$  whose union is the whole circle  $\mathbb{Z}/n$ . Without loss of generality, assume that

$$|\{x_1, \dots, x_m\} \cap [\bar{l}, l]| \geq |\{x_1, \dots, x_m\} \cap [l, \bar{l}]|.$$

Let  $x_r \in \{x_1, \dots, x_m\} \cap [\bar{l}, l]$  be the point in  $\{x_1, \dots, x_m\} \cap [\bar{l}, l]$  which is closest to  $l$ . We claim that  $x_r$  is also a midpoint:

$$\begin{aligned} \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, x_r) &= \sum_{x_k \in [\bar{l}, l]} d_{\mathbb{Z}/n}(x_k, x_r) + \sum_{x_k \in (\bar{l}, l]} d_{\mathbb{Z}/n}(x_k, x_r) \\ &\leq \sum_{x_k \in [\bar{l}, l]} (d_{\mathbb{Z}/n}(x_k, l) - d_{\mathbb{Z}/n}(x_r, l)) + \sum_{x_k \in (\bar{l}, l]} (d_{\mathbb{Z}/n}(x_k, l) + d_{\mathbb{Z}/n}(x_r, l)) \\ &= \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, l) - d_{\mathbb{Z}/n}(x_r, l) (|\{x_1, \dots, x_m\} \cap [\bar{l}, l]| - |\{x_1, \dots, x_m\} \cap [l, \bar{l}]|) \leq \sum_{k=1}^m d_{\mathbb{Z}/n}(x_k, l). \end{aligned}$$

■

*Proof for embedding the first term in  $L_1$ .* We map  $\Phi_1 : Sym_n \rightarrow l_1^{n^2}$  by

$$\Phi_1(\pi) := \left( e^{2\pi i(\pi(k)-\pi(r))/n} \right)_{k,r \in [n]} \quad \text{for all } \pi \in Sym_n.$$

For all  $\pi, \tau \in Sym_n$  we have:

$$\begin{aligned} ||\Phi_1(\pi) - \Phi_1(\tau)||_1 &= \sum_k \sum_r |e^{2\pi i(\pi(k)-\pi(r))/n} - e^{2\pi i(\tau(k)-\tau(r))/n}| \\ &\asymp \sum_k \sum_r \frac{1}{n} d_{\mathbb{Z}/n}(\pi(k) - \pi(r), \tau(k) - \tau(r)) = \frac{1}{n} \sum_k \sum_r d_{\mathbb{Z}/n}(\pi(k) - \tau(k), \pi(r) - \tau(r)) \\ &\asymp \min_r \sum_k d_{\mathbb{Z}/n}(\pi(k) - \tau(k), \pi(r) - \tau(r)) = \min_{l \in \mathbb{Z}/n} \sum_k d_{\mathbb{Z}/n}(\pi(k) - \tau(k), l). \end{aligned}$$

Finally, observe that in each of the two  $\asymp$  steps we pay factors  $\pi$  and 2 respectively. ■

## 5. Embedding the 2nd term: a lamplighter who forgets his position

**Lemma 4** (Embedding the 2nd term). .

There exists a 4-Lipschitz function  $\Phi_2 : Sym_n \rightarrow L_1$  such that for all  $\pi, \tau \in Sym_n$

**if**  $\min_l \sum_k d(\pi(k), \tau(k) + l) < n/3$ , **then**

$$\|\Phi_2(\pi) - \Phi_2(\tau)\|_1 \geq \frac{1}{8} \min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}).$$

Denote by  $\mathcal{J}$  the set of all intervals  $J = [a, b] = \{a, a+1, \dots, b\} \subset \mathbb{Z}/n$  and by  $J^{oo} := [a+2, b-2]$  the double-interior of  $J$  (which is empty when  $|J| \leq 2$ ). Given any function  $f : J \rightarrow \mathbb{Z}/n$  and  $J = [a, b] \in \mathcal{J}$  we will write  $\text{list}(f, J)$  for the function

$$\text{list}(f, J) : \{0, 1, \dots, b-a\} \rightarrow \mathbb{Z}/n : k \mapsto f(k+a).$$

Let  $\{v_{k,f}\}_{k,f}$  be a standard basis of  $l_1^N$  where  $1 \leq k \leq n$  and  $f : \{0, 1, \dots, k\} \rightarrow \mathbb{Z}/n$  is a function, and  $N \in \mathbb{N}$  is the number of pairs  $(k, f)$ . Define:

$$\Phi_2(\pi) := \frac{1}{n} \sum_{J \in \mathcal{J}} 1_{\{0 \notin J^{oo}\}} v_{|J|, \text{list}(\pi^{-1}, J)}.$$

The role of the notation  $\text{list}(\pi^{-1}, J)$  is to record the permutation  $\pi^{-1}$  on the interval  $J$  but forget the starting point of the interval. The above embedding is similar to the one in [33] except for this identification of the coordinates.

**Upper Bound** We show that  $\Phi_2$  is Lipschitz with respect to the word metric. Fix  $\pi \in \text{Sym}_n$ .

**Across a transposition edge**,  $\pi^{-1}$  and  $((01)\pi)^{-1}$  are identical functions except at the points 0 and 1. This means that for all  $J \in \mathcal{J}$  with

$$\{0, 1\} \cap J = \emptyset \iff v_{|J|, \text{list}(\pi^{-1}, J)} = v_{|J|, \text{list}((01)\pi)^{-1}, J}.$$

Also, if  $0 \in J^{oo}$ , then the coefficient of  $v_{|J|, \text{list}(\pi^{-1}, J)}$  will vanish and likewise for  $v_{|J|, \text{list}((01)\pi)^{-1}, J}$ . There are at most  $4n$  intervals  $J$  with  $\{0, 1\} \cap J \neq \emptyset$  and  $0 \notin J^{oo}$ .

We conclude that  $\|\Phi_2(\pi) - \Phi_2((01)\pi)\|_1 \leq 4$ .

**Across a cyclic permutation edge** Observe that for all  $J \in \mathcal{J}$ ,

$$\text{list}(\pi^{-1}, J) = \text{list}((c\pi)^{-1}, J+1),$$

so the corresponding coordinates are identical. The only way to get a nonzero summand in the  $l_1^N$  norm is whenever  $0 \in J^{oo}$  and  $0 \notin (J+1)^{oo} = J^{oo} + 1$  or whenever  $0 \notin J^{oo}$  and  $0 \in (J+1)^{oo} = J^{oo} + 1$ . The number of such  $J$  is  $\leq 2n$ ,

so we conclude that  $\|\Phi_2(\pi) - \Phi_2(c\pi)\|_1 \leq 2$ .

### Lower Bound

To avoid any ambiguity about the interval notation, in what follows, for all  $x, y, z \in \mathbb{Z}/n$ , we write  $x < y < z$  to say that the counterclockwise path from  $x$  to  $z$  passes through  $y$ . Also, we use the notation  $J = [a, b] := \{x \in \mathbb{Z}/n : a \leq x \leq b\}$ . (That way, for instance,  $[0, n-3]$  is an interval of size  $n-2$  whereas  $[n-3, 0]$  is an interval of size 4.)

**Observation 3.** For any subset  $S \subset \mathbb{Z}/n$  with  $|S| \leq n/3$  we have:

$$|\{J \in \mathcal{J} | \emptyset \neq J \cap S \neq J \text{ and } 0 \notin J^{oo}\}| \geq \frac{n}{4} \text{ diam}_{\mathbb{Z}/n}(\{0\} \cup S).$$

*Proof.* For notational simplicity assume that  $n$  is even. We consider the points

$$x := \arg \max_{x \in S} d_{\mathbb{Z}/n}(x, 0), \text{ and } y := \arg \max_{y \in \mathbb{Z}/n - S} d_{\mathbb{Z}/n}(y, 0).$$

**CASE A:**  $d(0, x) \neq n/2 = \text{diam}(\mathbb{Z}/n)$ .

Without loss of generality suppose that  $0 < x < n/2$  (so  $x$  in the "right" side of the clock). By definition  $x + 1 \notin S$ . Any interval  $J = [a, b] \in \mathcal{J}$  with  $0 \leq a \leq x < x + 1 \leq b \leq 0$  will satisfy  $\emptyset \neq J \cap S \neq J$  and  $0 \notin J^{\text{oo}}$ . In total, the number of such intervals is

$$\geq (n - d_{\mathbb{Z}/n}(x, 0))d_{\mathbb{Z}/n}(x, 0) \geq (n/2)d_{\mathbb{Z}/n}(x, 0) \geq (n/4)\text{diam}_{\mathbb{Z}/n}(\{0\} \cup S).$$

**CASE B:**  $x = n/2$ .

Without loss of generality, suppose that  $0 < y < n/2$ . By definition  $y + 1 \in S$ . Also, since  $|S| \leq n/3$ , we have  $d_{\mathbb{Z}/n}(y, 0) \geq n/3$ . Any interval  $J = [a, b] \in \mathcal{J}$  with  $0 \leq a \leq y < y + 1 \leq b \leq 0$  will satisfy  $\emptyset \neq J \cap S \neq J$  and  $0 \notin J^{\text{oo}}$ . In total, the number of such intervals is

$$\geq (n - d_{\mathbb{Z}/n}(y, 0))d_{\mathbb{Z}/n}(y, 0) \geq (n/2)(n/3) = (n/3)\text{diam}_{\mathbb{Z}/n}(\{0\} \cup S)$$

■

Now fix  $\pi, \tau \in \text{Sym}_n$  with  $\min_l \sum_k d(\pi(k), \tau(k) + l) < n/3$ , so there exists an "obvious pointer"  $l \in \mathbb{Z}/n$  such that  $|\{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}| < n/3$ . We denote the set of "positions of different tiles" by

$$D := \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}.$$

It is obvious that if an interval  $J \in \mathcal{J}$  is disjoint from this set,  $J \cap D = \emptyset$ , then we get identical coordinates  $v_{|J|, \text{list}(\pi^{-1}, J)} = v_{|J|, \text{list}(\tau^{-1}, J+l)}$ . We also have a partial converse:

**Observation 4** (some important coordinates have not been collapsed). .

**If**  $\emptyset \neq J \cap D \neq J$  **then**  $v_{|J|, \text{list}(\pi^{-1}, J)} \neq v_{|J'|, \text{list}(\tau^{-1}, J')}$  **for all**  $J' \in \mathcal{J}$ .

*Proof.* Pick  $p \in J - D$  and observe that  $\pi^{-1}(k) = \tau^{-1}(k - l)$  so the only way we could have  $v_{|J|, \text{list}(\pi^{-1}, J)} = v_{|J'|, \text{list}(\tau^{-1}, J')}$  was if  $J' = J + l$ . (This is because in a permutation, each element appears only once.) Picking any other point  $p' \in J \cap D$ , we see that  $v_{|J|, \text{list}(\pi^{-1}, J)} \neq v_{|J|, \text{list}(\tau^{-1}, J+l)}$  showing the claim. ■

We use Observations 3 and 4 to each of the summands in the definition of  $\Phi_2(\pi)$  and  $\Phi_2(\tau)$ .

$$\begin{aligned} \|\Phi_2(\pi) - \Phi_2(\tau)\|_1 &\geq \frac{1}{n} |\{J | \emptyset \neq J \cap D \neq J \text{ and } 0 \notin J^{\text{oo}}\}| + \frac{1}{n} |\{J | \emptyset \neq (J+l) \cap D \neq (J+l) \text{ and } 0 \notin J^{\text{oo}} + l\}| \\ &\geq \frac{1}{4} (\text{diam}(\{0\} \cup D) + \text{diam}(\{-l\} \cup D)) \geq \frac{1}{8} \text{diam}(\{0, l\} \cup D) \\ &= \frac{1}{8} \min_{l \in \mathbb{Z}/n} \text{diam}_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}), \end{aligned}$$

and the proof of Lemma 4 is complete.

### 6. Putting everything together and finishing the proof of theorem 1:

Let  $\Phi_1$  and  $\Phi_2$  be the two embeddings in the above two sections and take their direct sum  $\Phi = \Phi_1 \oplus \Phi_2 : \pi \mapsto (\Phi_1(\pi), \Phi_2(\pi)) \in L_1 \oplus L_1$ . We may rescale the map  $\Phi_1$  so that  $\|\Phi_1\|_{Lip} \leq 1$  and  $\|\Phi_1^{-1}\|_{Lip} \leq 2\pi$  (where by  $\Phi_1^{-1}$  we mean the inverse on the image).

For all  $\pi, \tau \in Sym_n$  we have two cases:

**CASE 1:**  $\min_l \sum_k d(\pi(k), \tau(k) + l) \geq n/3$ . For the lower bound:

$$\begin{aligned} d(\pi, \tau) &\geq \frac{1}{3} \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right) \\ &\geq \frac{1}{3} \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) \geq \frac{1}{3} \|\Phi_1(\pi) - \Phi_1(\tau)\|_1 \\ &\geq \frac{1}{3} \|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + \frac{1}{3} \|\Phi_2(\pi) - \Phi_2(\tau)\|_1 - \frac{4}{3} d(\pi, \tau). \end{aligned}$$

where we used  $\|\Phi_2(\pi) - \Phi_2(\tau)\|_1 \leq 4d(\pi, \tau)$ . After rearranging terms, we obtain:

$$d(\pi, \tau) \geq 7(\|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + \|\Phi_2(\pi) - \Phi_2(\tau)\|_1).$$

while for the upper bound:

$$\begin{aligned} d(\pi, \tau) &\leq 6 \min_{l \in \mathbb{Z}/n} \left( \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \right) \\ &\leq 12 \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + 6 \min_{l \in \mathbb{Z}/n} diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \\ &\leq 12 \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + 3n \leq (12 + 9) \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) \\ &\leq 21 \times 2\pi \|\Phi_1(\pi) - \Phi_1(\tau)\|_1 \leq 21 \times 2\pi (\|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + \|\Phi_2(\pi) - \Phi_2(\tau)\|_1) \end{aligned}$$

**CASE 2:**  $\min_l \sum_k d(\pi(k), \tau(k) + l) < n/3$ . For the lower bound

$$\|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + \|\Phi_2(\pi) - \Phi_2(\tau)\|_1 \leq \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + 4d(\pi, \tau) \leq 7d(\pi, \tau),$$

while for the upper bound:

$$\begin{aligned} d(\pi, \tau) &\leq 12 \min_{l \in \mathbb{Z}/n} \sum_{k \in \mathbb{Z}/n} d_{\mathbb{Z}/n}(\pi(k) - l, \tau(k)) + 6 \min_{l \in \mathbb{Z}/n} diam_{\mathbb{Z}/n}(\{0, l\} \cup \{p : \pi^{-1}(p) \neq \tau^{-1}(p - l)\}) \\ &\leq 12 \times 2\pi \|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + 6 \times 8 \|\Phi_2(\pi) - \Phi_2(\tau)\|_1 \leq 12 \times 2\pi (\|\Phi_1(\pi) - \Phi_1(\tau)\|_1 + \|\Phi_2(\pi) - \Phi_2(\tau)\|_1) \end{aligned}$$

Combining all the constants from the upper and lower bounds:

$$c_1(\Gamma(Sym_n, \{t, c\})) \leq 7 \times 21 \times 2\pi \approx 923.63 < 1000.$$

## 7. Remarks on embedding groups into Hilbert space

This section contains results of Naor and are included with his permission. Given a metric space  $(X, d)$  denote by  $c_2(X, d)$  the infimal bi-Lipschitz distortion of  $X$  into Hilbert space  $L_2$ .

**Conjecture 1** (Cornulier-Tessera-Valette [11]). .

*Let  $G$  be a finitely generated group with a word metric  $d_S$ . We have:*

$$c_2(G, d_S) < \infty \iff G \text{ has an abelian subgroup of finite index.}$$

The following is a finite version of this conjecture. Given a set  $X$  and  $r \in \mathbb{N}$  we write  $\binom{X}{r}$  for the collection of all subsets of  $X$  of size  $r$ .

**Conjecture 2** (Naor). *Let  $\{G_n\}_n$  be a sequence of finite groups with  $\sup_n \text{rank}(G_n) < \infty$ .*

*The following five conditions are equivalent:*

1. *For all  $r > \sup_n \text{rank}(G_n)$ , all Cayley graphs of  $\{G_n\}_n$  with  $r$  generators embed into Hilbert space with uniformly bounded distortion:*

$$\sup_n \sup_{S \in \binom{G_n}{r} : \langle S \rangle = G_n} c_2(G_n, d_S) < \infty$$

2. *There exists  $r \in \mathbb{N}$  such that all Cayley graphs of  $\{G_n\}_n$  with  $r$  generators embed into Hilbert space with uniformly bounded distortion:*

$$\sup_n \sup_{S \in \binom{G_n}{r} : \langle S \rangle = G_n} c_2(G_n, d_S) < \infty$$

3. *For all  $r > \sup_n \text{rank}(G_n)$  there exists a generating set of size  $r$ ,  $S_n$ , for each  $G_n$ , such that the Cayley graphs embed into Hilbert space with uniformly bounded distortion:*

$$\sup_n \inf_{S \in \binom{G_n}{r} : \langle S \rangle = G_n} c_2(G_n, d_S) < \infty$$

4. *There exists  $r \in \mathbb{N}$  and a generating set of size  $r$ ,  $S_n$ , for each  $G_n$ , such that the Cayley graphs embed into Hilbert space with uniformly bounded distortion:*

$$\sup_n \inf_{S \in \binom{G_n}{r} : \langle S \rangle = G_n} c_2(G_n, d_S) < \infty$$

5. *There exists  $r \in \mathbb{N}$  and a rank  $\leq r$  abelian subgroup  $H_n \leq G_n$  for each  $G_n$  such that*

$$\sup_n [G_n : H_n] < \infty.$$

In this conjecture, of course (1)  $\implies$  (2)&(3), and (2)  $\implies$  (4) and (3)  $\implies$  (4). The direction (5)  $\implies$  (1) is not immediate, since one has to deal with arbitrary generating sets of abelian groups. Nonetheless, it is true:

**Proposition 2** (Naor). *In the above conjecture, (5)  $\implies$  (1).*

*That is, for each  $r \in \mathbb{N}$  and  $K \in \mathbb{N}$  there exists  $D > 1$  such that*

*If  $G$  is a finite group which has an abelian subgroup  $H \leq G$  of index  $[G : H] \leq K$ , and  $S = \{s_1, \dots, s_r\}$  is a generating set of  $G$  then  $c_2(G, d_S) \leq D$ .*

Since  $L_2$  embeds bi-Lipschitzly into  $L_1$ , Proposition 2 implies that abelian groups of bounded rank exhibit Always- $L_1$  behavior. Also, the proof below gives the dependence  $D \lesssim_K r^{3/2} \log r$ . It would be interesting to know the optimal dependence on  $r$  in Proposition 2 (see also Conjecture 3 below).

*Proof. CASE 1:  $G$  is abelian*

The word metric of  $y \in G$  is given by:

$$d(y, 0) = \min \left\{ \sum_{i=1}^r |k_i| : y = \sum_{i=1}^r k_i s_i, \quad k_1, \dots, k_r \in \mathbb{Z} \right\}.$$

Consider  $\Lambda := \text{kernel}\{k \mapsto \sum_{i=1}^r k_i s_i\}$  which is the kernel of a map  $\mathbb{Z}^r \rightarrow G$  and a lattice of rank  $\leq r$ . We may write our group as the direct product of cyclic ones  $G = \oplus_{j=1}^J \mathbb{Z}_{m_j}$  and pick canonical generators  $e_1, \dots, e_J$ , where  $e_j = (\delta_{jj'})_{j'=1}^J$  is a generator for  $\mathbb{Z}/m_j$ . We write each of the canonical generators in terms of the generators  $s_1, \dots, s_r$ : there exist  $(a_i^{(j)})_{i=1}^r \in \mathbb{Z}^r$  for each  $j \in [J]$  such that

$$e_j = \sum_{i=1}^r a_i^{(j)} s_i \quad \text{for all } j \in [J].$$

That way, given any  $g = \sum_{j=1}^J g_j e_j \in G$ , we have:

$$g = \sum_{i=1}^r \left( \sum_{j=1}^J g_j a_i^{(j)} \right) s_i.$$

Now, if we have any element  $(k_1, \dots, k_r) \in \Lambda$ , then

$$g = \sum_{i=1}^r \left( k_i + \sum_{j=1}^J g_j a_i^{(j)} \right) s_i.$$

Conversely, for any two representations

$$g = \sum_{i=1}^r k_i s_i = \sum_{i=1}^r k'_i s_i \implies (k_i - k'_i)_{i=1}^r \in \Lambda.$$

Writing  $(Ag)_i := \sum_{j=1}^J g_j a_i^{(j)}$  for each  $i \in [r]$ , we conclude that the word metric is given by:

$$d_{\text{word}}(g, 0) = \min_{k \in \Lambda} \sum_{i=1}^r |k_i - \sum_{j=1}^J g_j a_i^{(j)}| = \min_{k \in \Lambda} \|k - Ag\|_{l_1^r}.$$

Form the **flat torus**  $\mathbb{R}^r/\Lambda$  which is a Riemannian manifold (with universal cover  $\mathbb{R}^r$ ) and the Riemannian metric is inherited from the standard Riemannian metric on the cover  $\mathbb{R}^r$ . The metric on the flat torus is given by the formula

$$d_{\mathbb{R}^r/\Lambda}(x + \Lambda, y + \Lambda) := \min_{z \in \Lambda} \|x - y + z\|_{l_2^r},$$

so by the Cauchy-Schwarz  $\leq$  and the above observation, we get that for all  $g_1, g_2 \in G$ ,

$$d_{\mathbb{R}^r/\Lambda}(Ag_1 + \Lambda, Ag_2 + \Lambda) \leq d_{\text{word}}(g_1, g_2) \leq \sqrt{r} d_{\mathbb{R}^r/\Lambda}(Ag_1 + \Lambda, Ag_2 + \Lambda).$$

By Theorem 5.8 in Khot-Naor [24], any flat torus  $(\mathbb{R}^r/\Lambda, d_{\mathbb{R}^r/\Lambda})$  embeds into  $L_2$  with distortion  $O(r^{3r/2})$ . (Remark: the dependence on  $r$  was recently improved by Haviv-Regev to  $r\sqrt{\log r}$  [19]). Say the map  $f : \mathbb{R}^r/\Lambda \rightarrow L_2$  achieves this distortion bound. Then the map

$$G \rightarrow L_2 : y \mapsto f(Ay + \Lambda)$$

is an embedding of distortion  $O_r(1)$ , where the distortion does NOT depend on  $G$ , nor on the generators  $s_1, \dots, s_r$ .

### CASE 2: Passing to finite extensions

We will essentially replicate the proof of the Milnor–Schwarz lemma (see [15] for instance), and observe that all the involved constants can be bounded by some function of  $[G : H]$  (and hence are independent of  $G$  and of the generators  $s_1, \dots, s_r$ ).

The Schreier graph  $\Gamma(G/H, S)$  is connected and has  $[G : H]$ -many points, hence its diameter at most  $[G : H]$ . This means that we may pick representatives for each coset  $k_1, \dots, k_J$  (where  $J = [G : H]$ ) such that

$$|k_i|_T \leq [G : H] \quad \text{and} \quad \cup_{j=1}^J k_j H = G.$$

Consider the following subset of  $H$ :

$$T' := H \cap \{k_{j_1}^{-1} s_i k_{j_2} : j_1, j_2 \in [J] \text{ and } i \in [r]\}.$$

We claim that  $T'$  generates  $H$  and moreover, the word metric of  $T'$  is "undistorted" with respect to the word metric of  $T$ :

$$|h|_{T'} \leq |h|_T \leq (2[G : H] + 1)|h|_{T'} \quad \text{for all } h \in H.$$

Fix any geodesic of length  $M \in \mathbb{N}$  from 1 to  $h \in H$  with generators in  $T$ ,

$$1 \rightarrow g_1 \rightarrow g_2 \rightarrow \dots \rightarrow g_M = h$$

so  $g_{m+1}g_m^{-1} = s_{i_m}$  for some  $i_m \in [r]$ . We may express via coset representatives  $g_m = k_{j_m}h_m$  where  $h_m \in H$  and  $j_m \in [J]$ . Now we have for each  $m \in [M]$ :

$$h_{m+1} = k_{j_{m+1}}^{-1} s_{i_m} k_{j_m} h_m.$$

This means that  $k_{j_{m+1}}^{-1} s_{i_m} k_{j_m} \in H$  and moreover the path

$$1 \rightarrow h_1 \rightarrow h_2 \rightarrow \dots \rightarrow h_M = h$$

is a path from 1 to  $h$  in the Cayley graph of  $H$  generated by  $T'$ . This shows that  $\langle T' \rangle = H$  and moreover,  $|h|_{T'} \leq |h|_T$ . On the other hand, since  $|k_{j_1}^{-1} s_{i_m} k_{j_2}|_T \leq 2[G : H] + 1$ , we get that:

$$|h|_{T'} \leq |h|_T \leq (2[G : H] + 1)|h|_{T'} \quad \text{for all } h \in H.$$

Finally, we construct the embedding. From CASE 1, we can find  $f : H \rightarrow L_2$  with distortion  $O_r(1)$ . Let  $([J], d_{\text{trivial}})$  be the trivial metric on  $[J]$ , where any two distinct points have distance 1. (This metric embeds isometrically into Hilbert space as the vertices of the standard simplex.) Map  $F : G \rightarrow L_2 \times [J] : k_j h \rightarrow (f(h), j)$ . We check that for each  $k_j h, k_{j'} h' \in G$ .

$$\frac{1}{2[G : H] + 1} d_{H, T'}(h, h') + 1_{\{j=j'\}} \leq d_{G, T}(k_j h, k_{j'} h') \leq [G : H] 1_{\{j=j'\}} + d_{H, T'}(h, h')$$

and hence we get an embedding of  $G$  into  $L_2$  whose distortion depends only on  $r$  and  $[G : H]$ . ■

One of the main results of Khot-Naor [24], answering a question of W. B. Johnson, is that there exists a lattice  $\Lambda \leq \mathbb{R}^r$  such that  $c_1(l_2^r/\Lambda) \gtrsim \sqrt{r}$ . Case 1 in the proof of Proposition 2 yields the following strengthening of this result.

**Proposition 3** (Naor). *For every  $r \in \mathbb{N}$  there exists a lattice  $\Lambda \subset \mathbb{R}^r$  such that  $c_1(l_1^r/\Lambda) \gtrsim r$ , where the metric on  $l_1^r/\Lambda$  is given by:  $d(x + \Lambda, y + \Lambda) := \inf_{k \in \Lambda} \|x - y + k\|_1$  for all  $x, y \in \mathbb{R}^r$ . By Cauchy-Schwarz, this implies  $c_1(l_2^r/\Lambda) \gtrsim \sqrt{r}$*

*Proof.* The proof is probabilistic. Recall the following special case of the Alon-Roichman theorem [1]. There exists a constant  $c > 0$  for which we sample  $r$  iid elements  $S = \{s_1, \dots, s_r\}$  uniformly at random from  $G = (\mathbb{Z}/2)^{\lfloor cr \rfloor}$ . Then with probability  $> 1/2$ , the Cayley graph  $\Gamma(G, S)$  has Cheeger constant  $> 1/10$ .

Consider the map  $\mathbb{Z}^r \rightarrow G : (k_i)_{i=1}^r \mapsto \sum_{i=1}^r k_i s_i$  and its kernel  $\Lambda := \text{kernel}(k \mapsto \sum_{s \in S} k_s s)$  which is a random lattice  $\Lambda \leq \mathbb{R}^r$ . As we saw in Case 1 in the proof of Proposition 2, the shortest path metric on  $\Gamma(G, S)$  embeds isometrically into the  $l_1$ -torus  $l_1^r/\Lambda$ . We show  $c_1(\Gamma(G, S))$  is large.

By Propositions 3.4 and 3.5 in Newmann-Rabinovich [34] we have

$$\frac{1}{|G|^2} \sum_{g \in G} \sum_{h \in G} d_{word}(x, y) \geq \frac{\text{diam}(\Gamma(G, S))}{2} - 1 \gtrsim r.$$

(The first estimate is immediate: consider the smallest  $r$  such that  $|B|(1, r) | > |G|/2$ . Then  $LHS \geq r - 1 \geq \text{diam}(\Gamma(G, S))/2 - 1$ . The second estimate follows from counting all possible small geodesics in an abelian group; see [34] for details.)

Fix any  $D > 1$  and  $f : \Gamma(G, S) \rightarrow L_1$  with  $d(g, h) \leq \|f(g) - f(h)\|_1 \leq Dd(g, h)$  for all  $g, h \in G$ . By the  $L_1$ -Poincaré inequality for the Cheeger constant (i.e. Theorem 4.7 in [40]), we have with probability  $> 1/2$  that

$$r \lesssim \frac{1}{|G|^2} \sum_{g \in G} \sum_{h \in G} \|f(g) - f(h)\|_1 \leq \frac{20}{|E(\Gamma)|} \sum_{(g, h) \in E(\Gamma)} \|f(g) - f(h)\|_1 \leq 20D.$$

We conclude that  $c_1(l_1^r/\Lambda) \geq c_1(\Gamma(G, S)) \gtrsim r$  with probability  $> 1/2$ . ■

The following conjecture is an  $L_1$ -analogue of the result of Haviv-Regev [19]:

**Conjecture 3** (Naor). *There exists  $c > 0$  such that for any  $r \in \mathbb{N}$  and any lattice  $\Lambda \subset \mathbb{R}^n$ :*

$$c_1(l_1^r/\Lambda) \lesssim r(\log r)^c.$$

A positive answer would imply that Proposition 3 is sharp up to logarithmic factors. Perhaps the conjecture is true without the logarithmic factors.

We return to bi-Lipschitz embeddings of groups into Hilbert space.

**Proposition 4.** .

*Assuming conjecture 1, the direction (4)  $\implies$  (5) in conjecture 2 follows, that is:*

*If  $\{G_n\}_n$  are finite groups and  $\{S_n\}_n$  generating sets of size  $= r$  with  $\sup_n c_2(G_n, d_{S_n}) < \infty$ , then each  $G_n$  has an abelian subgroup  $H_n$  such that  $\sup_n [G_n : H_n] < \infty$ .*

*Proof.* **STEP 1: Compactness and convergence to a limit group:**

We recall the definition of **Grigorchuk's space of marked groups**  $\mathcal{M}_r$  (see Section 6 in [17]). The points of  $\mathcal{M}_r$  consist of pairs  $(G, S)$  where  $G$  is a group, and  $S$  is an ordered generating set of size  $= r$ . Every such pair  $(G, S)$  is in one-to-one correspondence with the kernel  $\ker(\pi)$  of the unique canonical epimorphism from the free group to  $G$ ,  $\pi : \mathbb{F}_r \twoheadrightarrow G$ , which preserves the order of the generators. Each subgroup can be viewed as a subset of  $\mathbb{F}_r$ , i.e.  $\ker(\pi) \in \{0, 1\}^{\mathbb{F}_r}$ .

The topology of  $\mathcal{M}_r$  is the subspace topology inherited from the Tychonoff topology on  $\{0, 1\}^{\mathbb{F}_r}$ . One checks that  $\mathcal{M}_r$  is a compact topological space (see Proposition 6.1 in [17]) and that this topology is metrizable (see end of page 286 in [17]).

For each  $n$ , order the generators  $S_n = \{s_1, \dots, s_r\}$  in any manner. Consider the sequence of points in the space of marked groups  $\{(G_n, S_n)\}_n \subset \mathcal{M}_r$  and pick a convergent subsequence  $\{(G_{n_k}, S_{n_k})\}_k$  which we will relabel and denote by  $\{(G_n, S_n)\}_n$  (a slight abuse of notation) for our convenience.

**STEP 2: The limit group must be virtually abelian**

Note that convergence in the space of marked groups implies local convergence in the sense of Benjamini and Schramm. For every  $R \in \mathbb{N}$  consider the ball of radius  $R$  in the free group  $\mathbb{F}_r$  with standard generators. There exists  $n_0$  such that for all  $n > n_0$ ,  $(G_n, S_n)$  and  $(G, S)$  satisfy the same relations of length  $\leq R$ . This implies that the ball of radius  $R/2$  in the Cayley graph  $\Gamma(G_n, S_n)$ , is graph-isomorphic to the ball of radius  $R/2$  in  $\Gamma(G, S)$  and consequently, the word metric on the ball of radius  $R/4$  in  $\Gamma(G_n, S_n)$  is isometric to that of  $\Gamma(G, S)$ . By the hypothesis, we conclude that every finite subset of  $\Gamma(G, S)$  embeds into Hilbert space with distortion  $O(1)$ .

A theorem of Ostrovskii [39] states that for every Banach space  $X$  and every locally finite metric space  $(M, d)$  (i.e. a space where each ball has finitely many points), if every finite subset of  $M$  embeds into  $X$  with bi-Lipschitz distortion  $\leq D$  for some  $D \in (0, \infty)$ , then the entire metric space  $(M, d)$  embeds with bi-Lipschitz distortion  $\leq CD$  where  $C \in (0, \infty)$  is a universal constant. It follows that the Cayley graph  $\Gamma(G, S)$  embeds bi-Lipschitzly into Hilbert space.

A positive answer to Conjecture 1 would imply that  $G$  has an abelian subgroup  $H \leq G$  of finite index which we denote by  $K := [G : H] < \infty$ .

**STEP 3: The finite groups must also be virtually abelian.**

This step of the proof was shown to us by Emmanuel Breuillard.

Consider the canonical projection homomorphism  $\pi : \mathbb{F}_r \twoheadrightarrow G$ . Recall that  $G \cong \mathbb{F}_r / \ker(\pi)$ . Consider the subgroup  $F := \pi^{-1}(H)$  which has index

$$[\mathbb{F}_r : F] = [\mathbb{F}_r / \ker \pi : F / \ker \pi] = [G : H] = K.$$

By the Milnor–Schwarz lemma (see [15]) every finite index subgroup of a finitely generated group is finitely generated. Pick a finite generating set  $F = \langle w_1, \dots, w_{r'} \rangle$  (where  $r' \in \mathbb{N}$  and  $\{w_1, \dots, w_{r'}\}$  are words in  $S \cup S^{-1}$ ).

Each  $G_n$ , has its canonical projection map  $\pi_n : \mathbb{F}_r \twoheadrightarrow G_n$ . Define the subgroup:  $H_n := \pi_n(F)$ .

We have  $H_n \cong F / (F \cap \ker \pi_n) \cong (F \ker \pi_n) / \ker \pi_n$  and the index bound:

$$[G_n : H_n] = \left[ \frac{\mathbb{F}_r}{\ker(\pi_n)} : \frac{F \ker(\pi_n)}{\ker(\pi_n)} \right] = [\mathbb{F}_r : F \ker(\pi_n)] \leq [\mathbb{F}_r : F] = K.$$

Since  $H$  is abelian,  $\pi([w_i, w_j]) = 1$  for all  $i, j \in [r']$ . There exists  $R \in \mathbb{N}$  such that each word  $[w_i, w_j]$  ( $i, j \in [r']$ ) has length  $< R$ . By the convergence  $(G_n, S_n) \rightarrow (G, S)$ , there exists  $n_0$  such that for all  $n > n_0$ ,  $(G_n, S_n)$  and  $(G, S)$  satisfy the same relation of length  $\leq R$ , and hence

$$[\pi_n(w_i), \pi_n(w_j)] = \pi_n([w_i, w_j]) = 1 \quad \text{for all } i, j \in [r'].$$

Therefore subgroup  $H_n \leq G_n$  is abelian for all  $n > n_0$ . ■

Combining Proposition 2 and Proposition 4, we get:

**Corollary 2.** *The conjecture of Cornulier-Tessera-Valette (Conjecture 1) implies the finite version of Naor (Conjecture 2).*

It is natural to ask the following  $L_2$ -variant of Ostrovkii's Question 2: Does there exist  $r \in \mathbb{N}$  and generating sets  $T_n \subset Sym_n$  each of size  $|T_n| \leq r$  such that  $\sup_n c_2(\Gamma(Sym_n, d_{T_n})) < \infty$ ? Assuming Conjecture 2 the answer is no. The same negative answer follows from the following conjecture:

**Conjecture 4** (Naor). *For any  $r \in \mathbb{N}$  and  $M > 0$  there exists  $n \in \mathbb{N}$  such that for any generating set  $T_n$  of  $Sym_n$  of size  $|T_n| \leq r$  we have the following super-diffusive drift estimate for the simple random walk  $\{W_t\}_{t=1}^\infty$  on  $\Gamma(Sym_n, T_n)$  starting at  $W_0 = 1$ :*

$$\sup_{t \in \mathbb{N}} \frac{\mathbb{E} d_{T_n}(1, W_t)^2}{t} > M.$$

Potentially, stronger super-diffusivity estimates hold. Using Markov type 2 as in Linial-Magenvaer [27] (alternatively see Chapter 8 in [40] or Chapter 13 in [28]) Conjecture 4 implies that for any  $r \in \mathbb{N}$  and any sequence of generating sets  $\{T_n \subset Sym_n\}$  of size  $|T_n| \leq r$  we have  $\sup_n c_2(Sym_n, d_{T_n}) = \infty$ .

Next, we verify Conjecture 4 for specific generating sets of the symmetric groups.

Let  $T_n = \{(0123\dots n-1), (01)\}$  is the cycle and the transposition. Then one checks that the ball of radius  $n/4$  of  $Sym_n$  is graph-isomorphic to the ball of radius  $n/4$  of the 1D lampshuffler group  $Sym_{oo}(\mathbb{Z}) \times \mathbb{Z}$  with standard generators [45]. (In other words, we have local convergence.) By a theorem of Yadin [45], the simple random walk on  $Sym_{oo}(\mathbb{Z}) \times \mathbb{Z}$  has drift exponent  $3/4$ . This means that for small times the same estimate holds on the symmetric group:

$$\mathbb{E} d(1, W_t) \gtrsim t^{3/4} \quad \text{for all } 1 \leq t \leq n/4.$$

In what follows, we need to recall the following **Basic Fact**: The random walk  $\{W_t\}_{t=1}^\infty$  on an  $n$ -vertex degree  $d$  graph  $\Gamma$  with spectral gap  $1 - \lambda$  satisfies:

$$\mathbb{E} d(W_0, W_t) \gtrsim_{d,\lambda} t \quad \text{for all } 0 \leq t \lesssim_{d,\lambda} \log n,$$

where  $d(\cdot, \cdot)$  is the shortest path metric on  $\Gamma$ , and we have suppressed the dependence on the degree  $d$  and the spectral gap  $1 - \lambda$ . (This estimate goes back to Theorem 5 in Kesten [23]; see also Proposition 6.9 in [28].)

Let  $T_n$  be the generating set from Kassabov's theorem. The Cayley graphs are expanders, so Conjecture 4 is verified by the above fact.

Finally, we show that uniformly random generators satisfy Conjecture 4. Recall the theorem of Dixon [13]: an iid sample of uniformly random permutations  $\pi_1, \dots, \pi_r$  ( $r \geq 2$ ) generate the subgroup  $Alt_n$  (the alternating group) or the entire symmetric group  $Sym_n$  with probability  $> 1 - O(1/\log \log n) = 1 - o(1)$ . In particular, the probability that the generated subgroup is  $Alt_n$  equals  $1/2^r \pm o(1)$ .

**Proposition 5** (Naor). *For each  $r \geq 2$  let  $\pi_1, \dots, \pi_r \in Sym_n$  be iid sampled permutations from the uniform distribution. Then with high probability*

$$\mathbb{E} d_{\{\pi_1, \dots, \pi_r\}}(1, W_t) \gtrsim t \quad \text{for all } t = 1, \dots, O(\log n),$$

where  $\{W_t\}_{t=1}^\infty$  is the simple random walk on  $\langle \pi_1, \dots, \pi_r \rangle$ .

*First proof.* Note that by a theorem of Dixon [13], with high probability  $\langle \pi_1, \dots, \pi_r \rangle = Sym_n$  or  $Alt_n$  (the alternating group). We will consider the action of the group  $\langle \pi_1, \dots, \pi_r \rangle \curvearrowright \binom{[n]}{4}$  on the collection of all 4-element subsets on  $n$ . We can build the Schreier graph  $\Gamma_n(4)$  with vertices  $\binom{[n]}{4}$ , where we connect  $(A, \pi_i[A])$  for each generator  $\pi_i$  and subset  $A \in \binom{[n]}{4}$ .

By a theorem of Friedman-Joux-Roichman-Stern-Tillich [16],  $\Gamma_n(4)$  is an expander with high probability. Denote by  $W_t$  the simple random walk on  $Sym_n$  (or  $Alt_n$ ) and by  $W_t(\{1, 2, 3, 4\})$  the image of the set  $\{1, 2, 3, 4\}$  under the permutation  $W_t$ , and by  $d_{\Gamma_n(4)}$  the shortest path metric on the expander graph. By the basic fact above we get:

$$\mathbb{E} d_{\Gamma_n(4)}(\{1, 2, 3, 4\}, W_t(\{1, 2, 3, 4\})) \gtrsim t \quad \text{for all } t = 1, \dots, O(\log n).$$

The shortest path metric on the Schreier graph is at most by the shortest path metric on the Cayley graph (as the latter covers the former). This means that

$$\mathbb{E} d_{\{\pi_1, \dots, \pi_r\}}(1, W_t) \gtrsim t \quad \text{for all } t = 1, \dots, O(\log n).$$

■

*Second proof.* By a theorem of Dixon-Pyber-Seress-Shalev [14] for any nonempty freely reduced word  $w$  in the generators and inverses  $\{\pi_1, \dots, \pi_r, \pi_1^{-1}, \dots, \pi_r^{-1}\}$ ,  $p_n := \mathbb{P}[w = 1] \rightarrow 0$  as  $n \rightarrow \infty$ . By the union bound on all words of length  $< \log_2 1/p_n - 1$ , we conclude that the

$$\mathbb{E} girth(\Gamma(Sym_n, \{\pi_1, \dots, \pi_r\})) \gtrsim \log \frac{1}{p_n} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

For small times, the simple random walk on a degree  $2r$  regular graph of high girth is identical to that of a  $2r$ -regular tree, therefore  $\mathbb{E} d_{\{\pi_1, \dots, \pi_r\}}(1, W_t) \gtrsim t$  for times  $1 \leq t \leq girth(\Gamma)/2$ . ■

## 8. Proof of test space characterization of type

If  $X$  has trivial Rademacher type, then  $L_1$  is finitely representable into  $X$  by Pisier [42]. By Theorem 1, for each  $n \in \mathbb{N}$ ,  $\Gamma(Sym_n, \{c, t\})$  embeds into  $l_1^m$  with distortion  $< 1000$  for sufficiently high  $m$ , hence  $\Gamma(Sym_n, \{c, t\})$  embeds into  $X$  with distortion  $< 1000$ .

For the other direction, it suffices to show the following claim. The embedding is similar to that of Arzhantseva, Guba and Sapir [2] for  $\mathbb{Z}/n \wr \mathbb{Z}/n$ .

**Claim 1.** For each  $n \in \mathbb{N}$ , the Hamming cube  $(\{0, 1\}^n, || \cdot ||_1)$  embeds into  $\Gamma(Sym_{4n^2}, \{c, t\})$  with bi-Lipschitz distortion  $< 100$ .

*Proof.* For each  $n \in \mathbb{N}$  consider the following  $n$  disjoint transpositions in  $Sym_{4n^2}$

$$t_1 := (0, n), t_2 := (1 n + 1), t_3 := (2 n + 2), \dots, t_n := (n - 1 2n - 1).$$

Map each  $(\epsilon_1, \dots, \epsilon_n) \in \{0, 1\}^n$  to the permutation  $f_\epsilon := t_1^{\epsilon_1} \dots t_n^{\epsilon_n}$ .

Fix distinct  $\epsilon = (\epsilon_1, \dots, \epsilon_n), \delta = (\delta_1, \dots, \delta_n) \in \{0, 1\}^n$ . By the word metric formula in Section 2, we have the upper bound:

$$\begin{aligned} d(f_\epsilon, f_\delta) &\leq \min_{l \in \mathbb{Z}/4n^2} \left( 5 \sum_{k \in \mathbb{Z}/4n^2} d_{\mathbb{Z}/4n^2}(f_\epsilon(k) - l, f_\delta(k)) + 3diam_{\mathbb{Z}/4n^2}(\{0, l\} \cup \{p : f_\epsilon^{-1}(p) \neq f_\delta^{-1}(p - l)\}) \right) \\ &\leq 5 \sum_{k \in \mathbb{Z}/4n^2} d_{\mathbb{Z}/4n^2}(f_\epsilon(k), f_\delta(k)) + 6n = 5n||\epsilon - \delta||_1 + 6n \leq 11n||\epsilon - \delta||_1, \end{aligned}$$

and the lower bound:

$$\begin{aligned} d(f_\epsilon, f_\delta) &\geq \frac{1}{3} \min_{l \in \mathbb{Z}/4n^2} \sum_{k \in \mathbb{Z}/4n^2} d_{\mathbb{Z}/4n^2}(f_\epsilon(k) - l, f_\delta(k)) \\ &= \frac{1}{3} \min_{l \in \mathbb{Z}/4n^2} \left( (4n^2 - 2n)d_{\mathbb{Z}/4n^2}(0, l) + \sum_{i: \epsilon_i = \delta_i} d_{\mathbb{Z}/4n^2}(0, l) + \sum_{\epsilon_i = 0, \delta_i = 1} d_{\mathbb{Z}/4n^2}(n, l) + \sum_{\epsilon_i = 1, \delta_i = 0} d_{\mathbb{Z}/4n^2}(-n, l) \right). \end{aligned}$$

By inspection, it is clear that the minimizer occurs at  $l = 0$  so we get

$$\frac{1}{3}n||\epsilon - \delta||_1 \leq d(f_\epsilon, f_\delta) \leq 11n||\epsilon - \delta||_1.$$

■

## Acknowledgments

I thank Assaf Naor for his results which have been included in Section 7 of this paper, for encouragement, insightful discussions, writing advice, tons of context around this work, and for helping me with the use of representation theory for the embedding in Section 4. I thank Mikhail Ostrovskii for telling me his Question 2, and for insightful discussions and context. I thank Emmanuel Breuillard for showing me step 3 in Proposition 4. I also thank Kunal Chawla, Seung-Yeon Ryoo, and Eduardo Silva for insightful discussions and context. Finally, I thank the Casa Matemática Oaxaca (CMO) and Banff International Research Station for Mathematical Innovation and Discovery (BIRS), as well as the American Institute of Mathematics (AIM) for giving me the opportunity to speak with Mikhail Ostrovskii and learn about his question.

## REFERENCES

- [1] Noga Alon and Yuval Roichman. Random cayley graphs and expanders. *Random Structures & Algorithms*, 5(2):271–284, 1994.
- [2] Goulnara N Arzhantseva, Víctor S Guba, and Mark V Sapir. Metrics on diagram groups and uniform embeddings in a hilbert space. *Commentarii Mathematici Helvetici*, 81(4):911–929, 2006.
- [3] Tim Austin, Assaf Naor, and Alain Valette. The euclidean distortion of the lamplighter group. *Discrete & Computational Geometry*, 44(1):55–74, 2010.
- [4] L. Babai, W.M. Kantor, and A. Lubotsky. Small-diameter cayley graphs for finite simple groups. *European Journal of Combinatorics*, 10(6):507–522, 1989.
- [5] Keith Ball. The ribe programme. *Séminaire Bourbaki*, 2011:1043–1058, 2012.
- [6] J Bourgain. Remarks on the extension of lipschitz maps defined on discrete sets and uniform homeomorphisms. In *Geometrical Aspects of Functional Analysis: Israel Seminar, 1985–86*, pages 157–167. Springer, 2006.
- [7] Jean Bourgain, Vitali Milman, and Haim Wolfson. On type of metric spaces. *Transactions of the American Mathematical Society*, 294(1):295–317, 1986.
- [8] Emmanuel Breuillard and Alex Gamburd. Strong uniform expansion in  $\mathrm{sl}(2, p)$ . *Geometric and Functional Analysis*, 20(5):1201–1209, 2010.
- [9] Indira Chatterji and Cornelia Druțu. Median geometry for spaces with measured walls and for groups. *Mathematische Annalen*, pages 1–28, 2025.
- [10] Yves Cornulier, Yves Stalder, and Alain Valette. Proper actions of wreath products and generalizations. *Transactions of the American Mathematical Society*, 364(6):3159–3184, 2012.
- [11] Yves De Cornulier, Romain Tessera, and Alain Valette. Isometric group actions on banach spaces and representations vanishing at infinity. *Transformation Groups*, 13(1):125–147, 2008.
- [12] Michel Marie Deza and Monique Laurent. *Geometry of cuts and metrics. Algorithms and Combinatorics*, 15. Springer-Verlag, Berlin, xii+587 pp. ISBN: 3-540-61611-X, 1997.
- [13] John D Dixon. The probability of generating the symmetric group. *Mathematische Zeitschrift*, 110(3):199–205, 1969.
- [14] John D Dixon, László Pyber, Ákos Seress, and Aner Shalev. Residual properties of free groups and probabilistic methods. *Journal für die reine und angewandte Mathematik*, 556:159–172, 2003.
- [15] Cornelia Druțu and Michael Kapovich. *Geometric group theory*, volume 63. American Mathematical Soc., 2018.
- [16] Joel Friedman, Antoine Joux, Yuval Roichman, Jacques Stern, and Jean-Pierre Tillich. The action of a few permutations on r-tuples is quickly transitive. *Random Structures & Algorithms*, 12(4):335–350, 1998.
- [17] Rostislav I Grigorchuk. Degrees of growth of finitely generated groups, and the theory of invariant means. *Mathematics of the USSR-Izvestiya*, 25(2):259, 1985.
- [18] Mikhail Gromov. Random walk in random groups. *Geometric and Functional Analysis*, 13(1):73–146, 2003.
- [19] Ishay Haviv and Oded Regev. The euclidean distortion of flat tori. *Journal of Topology and Analysis*, 5(02):205–223, 2013.
- [20] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.
- [21] Paata Ivanisvili, Ramon Van Handel, and Alexander Volberg. Rademacher type and enflo type coincide. *Annals of mathematics*, 192(2):665–678, 2020.
- [22] Martin Kassabov. Symmetric groups and expander graphs. *Inventiones mathematicae*, 170(2):327–354, 2007.
- [23] Harry Kesten. Symmetric random walks on groups. *Transactions of the American Mathematical Society*, 92(2):336–354, 1959.
- [24] Subhash Khot and Assaf Naor. Nonembeddability theorems via fourier analysis. *Mathematische Annalen*, 334(4):821–852, 2006.
- [25] Cosmas Kravaris and Eduardo Silva. Forthcoming manuscript, 2026. in preparation.
- [26] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

- [27] Nathan Linial, Avner Magen, and Assaf Naor. Girth and euclidean distortion. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 705–711, 2002.
- [28] Russell Lyons and Yuval Peres. *Probability on trees and networks*, volume 42. Cambridge University Press, 2017.
- [29] Jiri Matousek. *Lectures on discrete geometry*, volume 212. Springer Science & Business Media, 2013.
- [30] Bernard Maurey. Type, cotype and k-convexity. *Handbook of the geometry of Banach spaces*, 2:1299–1332, 2003.
- [31] Manor Mendel and Assaf Naor. Expanders with respect to hadamard spaces and random graphs. In *Proceedings of the 5th conference on Innovations in theoretical computer science*, pages 353–358, 2014.
- [32] Assaf Naor. An introduction to the ribe program. *Japanese Journal of Mathematics*, 7(2):167–233, 2012.
- [33] Assaf Naor and Yuval Peres. Embeddings of discrete groups and the speed of random walks. *imrn*, 2008:rnn076, 2008.
- [34] Ilan Newman and Yuri Rabinovich. Hard metrics from cayley graphs of abelian groups. *Theory OF Computing*, 5(1):125–134, 2009.
- [35] M. I. Ostrovskii. Different forms of metric characterizations of classes of Banach spaces. *Houston J. Math.*, 39(3):889–906, 2013.
- [36] Mikhail Ostrovskii. Metric characterizations of some classes of banach spaces. In *Harmonic Analysis, Partial Differential Equations, Complex Analysis, Banach Spaces, and Operator Theory (Volume 1) Celebrating Cora Sadosky’s life*, pages 307–347. Springer, 2016.
- [37] Mikhail Ostrovskii. Participant list: Metric embeddings workshop at american institute of mathematics. <http://admin.aimath.org/resources/metricembeddings/participantlist/>, June 2025.
- [38] Mikhail Ostrovskii and Beata Randrianantoanina. A characterization of superreflexivity through embeddings of lamplighter groups. *Proceedings of the American Mathematical Society*, 147(11):4745–4755, 2019.
- [39] Mikhail I Ostrovskii. Embeddability of locally finite metric spaces into banach spaces is finitely determined. *Proceedings of the American Mathematical Society*, pages 2721–2730, 2012.
- [40] Mikhail I Ostrovskii. *Metric embeddings: Bilipschitz and coarse embeddings into Banach spaces*, volume 49. Walter de Gruyter, 2013.
- [41] Mikhail I Ostrovskii. Personal communication, May 2025.
- [42] Gilles Pisier. Sur les espaces qui ne contiennent pas de l 1 n uniformément. *Séminaire Maurey-Schwartz*, pages 1–19, 1973.
- [43] Yuri Rabinovich. On average distortion of embedding metrics into the line. *Discrete & Computational Geometry*, 39(4):720–733, 2008.
- [44] Martin Ribe. On uniformly homeomorphic normed spaces. *Arkiv för Matematik*, 14(1):237–244, 1976.
- [45] Ariel Yadin. Rate of escape of the mixer chain. *Electronic Communications in Probability*, 14:347–357, 2009.