

# TOWARDS A CALCULUS FOR NON-LINEAR SPECTRAL GAPS [EXTENDED ABSTRACT]

MANOR MENDEL AND ASSAF NAOR

ABSTRACT. Given a finite regular graph  $G = (V, E)$  and a metric space  $(X, d_X)$ , let  $\gamma_+(G, X)$  denote the smallest constant  $\gamma_+ > 0$  such that for all  $f, g : V \rightarrow X$  we have:

$$\frac{1}{|V|^2} \sum_{x, y \in V} d_X(f(x), g(y))^2 \leq \frac{\gamma_+}{|E|} \sum_{xy \in E} d_X(f(x), g(y))^2.$$

In the special case  $X = \mathbb{R}$  this quantity coincides with the reciprocal of the absolute spectral gap of  $G$ , but for other geometries the parameter  $\gamma_+(G, X)$ , which we still think of as measuring the non-linear spectral gap of  $G$  with respect to  $X$  (even though there is no actual spectrum present here), can behave very differently.

Non-linear spectral gaps arise often in the theory of metric embeddings, and in the present paper we systematically study the theory of non-linear spectral gaps, partially in order to obtain a combinatorial construction of a super-expander — a sequence of bounded-degree graphs  $\{G_i = (V_i, E_i)\}_{i=1}^\infty$  which does not admit a coarse embedding into any uniformly convex normed space. In addition, the bi-Lipschitz distortion of  $G_i$  in any uniformly convex normed space is  $\Omega(\log |V_i|)$ , which is the worst possible behavior due to Bourgain’s embedding theorem [4]. Such remarkable graph families were previously known to exist due to a tour de force algebraic construction of Lafforgue [17]. Our construction is different and combinatorial, relying on the zigzag product of Reingold-Vadhan-Wigderson [38].

We show that non-linear spectral gaps behave sub-multiplicatively under zigzag products — a fact that amounts to a simple iteration of the inequality above. This yields as a special case a very simple (linear algebra free) proof of the Reingold-Vadhan-Wigderson theorem which states that zigzag products preserve the property of having an absolute spectral gap (with quantitative control on the size of the gap). The zigzag iteration of Reingold-Vadhan-Wigderson also involves taking graph powers, which is trivial to analyze in the classical “linear” setting. In our work, the behavior of non-linear spectral gaps under graph powers becomes a major geometric obstacle, and we show that for uniformly convex normed spaces there exists a satisfactory substitute for spectral calculus which makes sense in the non-linear setting. These facts, in conjunction with a variant of Ball’s notion of Markov cotype and a Fourier analytic proof of the existence of appropriate “base graphs”, are shown to imply that Reingold-Vadhan-Wigderson type constructions can be carried out in the non-linear setting.

---

2010 *Mathematics Subject Classification.* 51F99,05C12,05C50,46B85.

To be published (in a different layout) in the proceedings of the ACM-SIAM Symposium on Discrete Algorithms 2010 (SODA ’10).

## 1. INTRODUCTION

Let  $A = (a_{ij})$  be an  $n \times n$  symmetric stochastic matrix and let

$$1 = \lambda_1(A) \geq \lambda_2(A) \geq \dots \geq \lambda_n(A) \geq -1$$

be its eigenvalues. The reciprocal of the spectral gap of  $A$ , i.e., the quantity  $\frac{1}{1-\lambda_2(A)}$ , is the smallest constant  $\gamma > 0$  such that for every  $x_1, \dots, x_n \in \mathbb{R}$  we have

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2 \leq \frac{\gamma}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij} (x_i - x_j)^2. \quad (1)$$

By summing over the coordinates with respect to some orthonormal basis, (1) can be restated as follows: the value  $\frac{1}{1-\lambda_2(A)}$  is the smallest constant  $\gamma > 0$  such that for all  $x_1, \dots, x_n \in L_2$  we have

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \|x_i - x_j\|_2^2 \leq \frac{\gamma}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|x_i - x_j\|_2^2. \quad (2)$$

Once one realizes the validity of inequality (2) it is natural to generalize it in several ways. For example, we can replace the exponent 2 by some other exponent  $p > 0$  and, more crucially, we can replace the Euclidean geometry by some other metric space  $(X, d)$ . Such generalizations are standard practice in metric geometry, as we shall discuss below. For the sake of presentation, we can take this generalization to even greater extremes: let  $X$  be an arbitrary set and let  $K : X \times X \rightarrow [0, \infty)$  be a symmetric function. Such functions are often called *kernels* in the literature, and we shall adopt this terminology here. Define the reciprocal spectral gap of  $A$  with respect to  $K$ , denoted  $\gamma(A, K)$ , as the infimum over all  $\gamma \geq 0$  such that for all  $x_1, \dots, x_n \in X$  we have

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(x_i, x_j) \leq \frac{\gamma}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij} K(x_i, x_j). \quad (3)$$

In what follows we will also call  $\gamma(A, K)$  the Poincaré constant of  $A$  with respect to  $K$ . Readers are encouraged to focus on the case when  $K$  is a power of some metric on  $X$ , though as will become clear presently, a surprising amount of ground can be covered without any assumption on the kernel  $K$ . For concreteness we restate the above discussion: the standard gap in the linear spectrum of  $A$  corresponds to considering Poincaré constants with respect to Euclidean spaces (i.e., kernels which are squares of Euclidean metrics), but there is scope for a theory of non-linear spectral gaps when one considers inequalities such as (3) with respect to other geometries. The purpose of this paper is to make some steps towards such a theory, with emphasis on possible extensions of spectral calculus to non-linear (non-Euclidean) settings. We apply our new calculus for non-linear spectral gaps to construct new strong types of expanders, and to resolve a question of Lafforgue [17]. We obtain a new combinatorial construction of a remarkable type of bounded degree graphs whose shortest path metric is incompatible with the geometry of any uniformly convex normed space in a very strong sense (i.e., coarse non-embeddability). The existence of such graph families was first discovered by Lafforgue [17] via an algebraic construction. Our work indicates that there is hope for a useful theory of non-linear spectral gaps, beyond the sporadic examples that have been previously studied in the literature.

1.1. **Coarse non-embeddability.** A sequence of metric spaces  $\{(X_n, d_n)\}_{n=1}^\infty$  is said to embed coarsely (with uniform moduli) into a metric space  $(Y, d_Y)$  if there exist two non-decreasing functions  $\alpha, \beta : [0, \infty) \rightarrow [0, \infty)$  such that  $\lim_{t \rightarrow \infty} \alpha(t) = \infty$ , and mappings  $f_n : X_n \rightarrow Y$ , such that for all  $n \in \mathbb{N}$  and  $x, y \in X_n$  we have:

$$\alpha(d_{X_n}(x, y)) \leq d_Y(f_n(x), f_n(y)) \leq \beta(d_{X_n}(x, y)). \quad (4)$$

Equation (4) should be viewed as a weak form of “metric faithfulness” of the mappings  $f_n$ : this seemingly humble requirement can be restated informally as “large distances should map uniformly to large distances”. Nevertheless, this weak notion of embedding (which is much weaker than, say, bi-Lipschitz embeddability), has a wide range of applications in geometry and group theory: see for example the book [39] and the references therein for (a small part of) such applications.

Since coarse embeddability is a weak requirement, it is quite difficult to prove coarse non-embeddability: very few methods to establish such a result are known, among which is the use of non-linear spectral gaps, as pioneered by Gromov [10] (other such methods are coarse notions of metric dimension [9], or the use of metric cotype [23]. These methods do not seem to be applicable to the question that we study here). Gromov’s argument is simple: assume that  $X_n = (V_n, E_n)$  are regular graphs of bounded degree and that  $d_n(\cdot, \cdot)$  is the shortest-path metric on  $X_n$ . Assume also that for some  $p, \gamma \in (0, \infty)$  we have for all  $n \in \mathbb{N}$  and  $f : V_n \rightarrow Y$ :

$$\frac{1}{|V_n|^2} \sum_{u, v \in V_n} d_Y(f(u), f(v))^p \leq \frac{\gamma}{|E_n|} \sum_{x, y \in E_n} d_Y(f(x), f(y))^p. \quad (5)$$

A combination of (4) and (5) yields the bound  $\frac{1}{|V_n|^2} \sum_{u, v \in V_n} \alpha(d_n(u, v))^p \leq \gamma\beta(1)^p$ . But, since  $X_n$  is a bounded degree graph, at least half of the pairs of vertices  $u, v \in V_n$  satisfy  $d_n(u, v) \geq c \log |V_n|$ , where  $c$  is a constant which depends on the implied degree bound (but not on  $n$ ). Thus  $\alpha(c \log |V_n|)^p \leq 2\gamma\beta(1)^p$ , and in particular if  $\lim_{n \rightarrow \infty} |V_n| = \infty$  then we get a contradiction to the assumption  $\lim_{t \rightarrow \infty} \alpha(t) = \infty$ . Observe in passing that this argument also shows that  $X_n$  has bi-Lipschitz distortion  $\Omega(\log |V_n|)$  in  $Y$  — such an argument was first used by Linial, London and Rabinovich [19] (see also [20]) to show that Bourgain’s embedding theorem [4] is asymptotically sharp.

Assumption (5) can be restated as saying that  $\gamma(A_n, d_Y^p) \leq \gamma$ , where  $A_n$  is the normalized adjacency matrix of  $X_n$ . This condition can be viewed to mean that the graphs  $\{X_n\}_{n=1}^\infty$  are “expanders” with respect to  $(Y, d_Y)$ . Note that if  $Y$  contains at least two points then (5) implies that  $\{X_n\}_{n=1}^\infty$  are necessarily also expanders in the classical sense.

The key point in the coarse non-embeddability question is therefore to construct such  $\{X_n\}_{n=1}^\infty$  for which we can prove the inequality (5) for non-Hilbertian targets  $Y$ . This question has been previously investigated by several authors. Matoušek [20] devised an extrapolation method for Poincaré inequalities (see also the description of Matoušek’s argument in [3]) which establishes the validity of (5) for every expander when  $Y = L_p$ . The work of Ozawa [30] and Pisier [32, 35] proves (5) for every expander when  $Y$  is Banach space which satisfies certain geometric conditions (e.g.,  $Y$  can be taken to be a Banach lattice with finite cotype). In [28, 27] additional results of this type are obtained.

A normed space is called *super-reflexive* if it admits an equivalent norm which is uniformly convex. Recall that a normed space  $(X, \|\cdot\|_X)$  is called uniformly convex if for every  $\varepsilon \in (0, 1)$  there exists  $\delta = \delta_X(\varepsilon) > 0$  such that for any two vectors  $x, y \in X$  with  $\|x\|_X = \|y\|_X = 1$

and  $\|x - y\|_X \geq \varepsilon$  we have  $\left\|\frac{x+y}{2}\right\|_X \leq 1 - \delta$  (thus uniform convexity is a uniform version of *strict convexity*). The question whether there exists a sequence of arbitrarily large graphs of bounded degree which do not admit a coarse embedding into any super-reflexive normed space was posed by Kasparov and Yu in [15], and was solved in the remarkable work of Lafforgue [17] on the strengthened version of property (T) for  $SL_3(\mathbb{F})$  when  $\mathbb{F}$  is a non-Archimedean local field (thus, for concreteness, Lafforgue’s graphs can be obtained as Cayley graphs of finite quotients of co-compact lattices in  $SL_3(\mathbb{Q}_p)$ , where  $p$  is a prime and  $\mathbb{Q}_p$  is the  $p$ -adic rationals).

In this paper we obtain an alternative solution of the Kasparov-Yu problem using a combinatorial construction based on the “zigzag expanders” of Reingold, Vadhan, and Wigderson [38]. More specifically, we construct a family of 9-regular graphs which satisfies (5) for every super-reflexive Banach space  $X$  (where  $\gamma$  depends only on  $X$ ) — such graphs are called *super-expanders*.

We state at the outset that it is a major open question whether every expander satisfies (5) for every uniformly convex normed space  $X$ . It is also unknown whether there exist graph families of bounded degree and logarithmic girth which do not admit a coarse embedding into any super-reflexive normed space—this question is of particular interest in the context of the potential application to the Novikov conjecture that was suggested by Kasparov and Yu in [15]. Note that some geometric restriction on the target space  $X$  must be imposed, since the relation between non-linear spectral gaps and coarse non-embeddability, in conjunction with the fact that every finite metric space embeds isometrically into  $\ell_\infty$ , shows that (for example)  $X = \ell_\infty$  can never satisfy (5) for bounded degree family of graphs.

Our combinatorial approach can be used to show that there exist bounded degree graph sequences which do not admit a coarse embedding into any  $K$ -convex normed space. A normed space  $X$  is  $K$ -convex<sup>1</sup> if there exists  $\varepsilon_0 > 0$  and  $n_0 \in \mathbb{N}$  such that any embedding of  $\ell_1^{n_0}$  into  $X$  incurs distortion at least  $1 + \varepsilon_0$ , see [33]. This question was asked by Lafforgue [17]. Recently, independently of our work, Lafforgue [18] managed to modify his argument to obtain coarse non-embeddability into  $K$ -convex spaces for his graph sequences as well.

**Acknowledgments.** Michael Langberg was involved in early discussions on the analysis of the zigzag product. Keith Ball helped in simplifying this analysis. M. M. was partially supported by ISF grant no. 221/07, BSF grant no. 2006009, and a gift from Cisco research center. A. N. was supported in part by NSF grants CCF-0635078 and CCF-0832795, BSF grant 2006009, and the Packard Foundation.

## 2. A COMBINATORIAL APPROACH TO THE EXISTENCE OF SUPER-EXPANDERS

**2.1. The compatibility of non-linear spectral gaps with combinatorial constructions.** The parameter  $\gamma(A, K)$  will reappear presently, but for the purpose of this section we need to study a variant of it which corresponds to the absolute spectral gap of a matrix (similar to the role of absolute spectral gaps in the work of Reingold-Vadhan-Wigderson [38]). Define  $\lambda(A) = \max_{2 \leq i \leq n} |\lambda_i(A)|$  and call the quantity  $1 - \lambda(A)$  the absolute spectral gap of  $A$ . Similarly to (2), the reciprocal of the absolute spectral gap of  $A$  is the smallest constant

---

<sup>1</sup> $K$ -convexity is also equivalent to  $X$  having type strictly bigger than 1, see [25, 21]. The  $K$ -convexity property is strictly weaker than super-reflexivity, see [12, 14, 13, 36].

$\gamma_+ > 0$  such that for all  $x_1, \dots, x_n, y_1, \dots, y_n \in L_2$  we have

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \|x_i - y_j\|_2^2 \leq \frac{\gamma_+}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|x_i - y_j\|_2^2. \quad (6)$$

Analogously to (3), given a kernel  $K : X \times X \rightarrow [0, \infty)$  we can then define  $\gamma_+(A, K)$  to be the the infimum over all  $\gamma_+ \geq 0$  such that for all  $x_1, \dots, x_n, y_1, \dots, y_n \in X$  we have

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n K(x_i, y_j) \leq \frac{\gamma_+}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij} K(x_i, y_j). \quad (7)$$

Note that clearly  $\gamma_+(A, K) \geq \gamma(A, K)$ .

In what follows we will often deal with regular graphs, which will always be allowed to have self loops and multiple edges. We will use the convention that each self loop contribute 1 to the degree of a vertex. The normalized adjacency matrix of a  $d$ -regular graph  $G = (V, E)$ , denoted  $A_G$ , is defined as usual by letting its  $u, v$  entry be the number of edges joining  $u, v \in V$  divided by  $d$ . When discussing Poincaré constants we will interchangeably identify  $G$  with  $A_G$ . Thus, for example, we write  $\gamma_+(G, K) = \gamma_+(A_G, K)$ .

The starting point of our work is an investigation of the behavior of the quantity  $\gamma_+(G, K)$  under certain graph operations, the most important of which (for our purposes) is the zigzag product of Reingold-Vadhan-Wigderson [38]. As we shall see below, combinatorial constructions seem to be well-adapted to controlling non-linear quantities such as  $\gamma_+(G, K)$ . This crucial fact allows us to use them in a perhaps unexpected geometric context.

Assume now that  $G_1 = (V_1, E_1)$  is an  $n_1$ -vertex graph which is  $d_1$ -regular and that  $G_2 = (V_2, E_2)$  is a  $d_1$ -vertex graph which is  $d_2$ -regular. Since the number of vertices in  $G_2$  is the same as the degree in  $G_1$ , we can identify  $V_2$  with the edges emanating from a given vertex  $u \in V_1$ . Formally, we fix for every  $u \in V_1$  a bijection  $\pi_u : \{e \in E_1 : u \in e\} \rightarrow V_2$ . Moreover, we fix for every  $a \in V_2$  a bijection between  $[d_2] = \{1, \dots, d_2\}$  and the multiset of the vertices adjacent to  $a$  in  $G_2$ ,  $\kappa_a : [d_2] \rightarrow \{b \in V_2 : \{a, b\} \in E_2\}$ .

The zigzag product  $G_1 \mathbb{Z} G_2$  is the graph whose vertices are  $V_1 \times V_2$  and  $(u, a), (v, b) \in V_1 \times V_2$  are joined by an edge if and only if there exist  $i, j \in [d_2]$  such that:

$$\{u, v\} \in E_1 \quad \text{and} \quad a = \kappa_{\pi_u(\{u, v\})}(i) \quad \text{and} \quad b = \kappa_{\pi_v(\{u, v\})}(j).$$

The schematic description of this construction is as follows: think of the vertex set of  $G_1 \mathbb{Z} G_2$  as a disjoint union of ‘‘clouds’’ which are copies of  $V_2 = \{1, \dots, d_1\}$  indexed by  $V_1$ . Thus  $(u, a)$  is the point indexed by  $a$  in the cloud labeled by  $u$ . Every edge  $\{(u, a), (v, b)\}$  of  $G_1 \mathbb{Z} G_2$  is the result of a three step walk: a ‘‘zig’’ step in  $G_2$  from  $a$  to  $\pi_u(\{u, v\})$  in  $u$ 's cloud, a ‘‘zag’’ step in  $G_1$  from  $u$ 's cloud to  $v$ 's cloud along the edge  $\{u, v\}$  and a final ‘‘zig’’ step in  $G_2$  from  $\pi_v(\{u, v\})$  to  $b$  in  $v$ 's cloud. The zigzag product is illustrated in Figure 1. The number of vertices of  $G_1 \mathbb{Z} G_2$  is  $n_1 d_1$  and its degree is  $d_2^2$ .

The zigzag product depends on the labels  $\{\pi_u\}_{u \in V_1}$ , and in fact different labels of the same graphs can produce non-isomorphic products<sup>2</sup>. However, all of our results below will be independent of the actual choice of the labeling, so while our notation should formally depend on the labeling, we will drop its explicit mention for the sake of simplicity.

---

<sup>2</sup>The labels  $\{\kappa_a\}_{a \in V_2}$  do not affect the structure of the zigzag product but they are useful in the subsequent analysis.

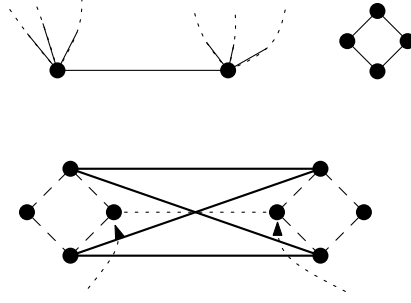


FIGURE 1. Illustration of the zigzag product. The upper part of the figure depicts *part* of a 4-regular graph  $G_1$ , and a 4-vertex cycle  $G_2$ . The bottom part of the figure depicts the edges of the zigzag product between  $u$ 's cloud and  $v$ 's cloud. The original edges of  $G_1$  and  $G_2$  are drawn as dotted and dashed lines (respectively).

Let us now examine how  $\gamma_+(G_1 \otimes G_2, K)$  is related to  $\gamma_+(G_1, K)$ , and  $\gamma_+(G_2, K)$ , where  $K : X \times X \rightarrow [0, \infty)$  is an arbitrary kernel. To this end take  $f, g : V_1 \times V_2 \rightarrow X$  and note that the definition of  $\gamma_+(G_1, K)$  implies that for all  $a, b \in V_2$  we have:

$$\frac{1}{n_1^2} \sum_{u, v \in V_1} K(f(u, a), g(v, b)) \leq \frac{\gamma_+(G_1, K)}{n_1 d_1} \sum_{\{u, v\} \in E_1} K(f(u, a), g(v, b)). \quad (8)$$

Thus:

$$\begin{aligned} \frac{1}{|V_1 \times V_2|^2} \sum_{(u, a), (v, b) \in V_1 \times V_2} K(f(u, a), g(v, b)) &= \frac{1}{d_1^2} \sum_{a, b \in V_2} \frac{1}{n_1^2} \sum_{u, v \in V_1} K(f(u, a), g(v, b)) \\ &\stackrel{(8)}{\leq} \frac{\gamma_+(G_1, K)}{n_1 d_1^3} \sum_{a, b \in V_2} \sum_{\{u, v\} \in E_1} K(f(u, a), g(v, b)). \end{aligned} \quad (9)$$

Next, the definition of  $\gamma_+(G_2, K)$  implies that for all  $u \in V_1$  and  $b \in V_2$  we have

$$\begin{aligned} \frac{1}{d_1^2} \sum_{a \in V_2} \sum_{\substack{v \in V_1: \\ \{u, v\} \in E_1}} K(f(u, a), g(v, b)) \\ \leq \frac{\gamma_+(G_2, K)}{d_1 d_2} \sum_{\substack{v \in V_1: \\ \{u, v\} \in E_1}} \sum_{i \in [d_2]} K(f(u, \kappa_{\pi_u(\{u, v\})}(i)), g(v, b)). \end{aligned} \quad (10)$$

Summing (10) over  $u \in V_1$  and  $b \in V_2$  and plugging it into (9), yields the bound:

$$\begin{aligned} \frac{1}{|V_1 \times V_2|^2} \sum_{(u, a), (v, b) \in V_1 \times V_2} K(f(u, a), g(v, b)) \\ \leq \frac{\gamma_+(G_1, K) \gamma_+(G_2, K)}{n_1 d_1^2 d_2} \sum_{v \in V_1} \sum_{i \in [d_2]} \sum_{\substack{u \in V_1: \\ \{u, v\} \in E_1}} \sum_{b \in V_2} K(f(u, \kappa_{\pi_u(\{u, v\})}(i)), g(v, b)). \end{aligned} \quad (11)$$

Another application of the definition of  $\gamma_+(G_2, K)$  implies that for all  $v \in V_1$  and  $i \in [d_2]$  we have:

$$\begin{aligned} \frac{1}{d_1^2} \sum_{\substack{u \in V_1: \\ \{u,v\} \in E_1}} \sum_{b \in V_2} K(f(u, \kappa_{\pi_u(\{u,v\})}(i)), g(v, b)) \\ \leq \frac{\gamma_+(G_2, K)}{d_1 d_2} \sum_{\substack{u \in V_1: \\ \{u,v\} \in E_1}} \sum_{j \in [d_2]} K(f(u, \kappa_{\pi_u(\{u,v\})}(i)), g(v, \kappa_{\pi_v(\{u,v\})}(j))). \end{aligned} \quad (12)$$

Summing (12) over  $v \in V_1$  and  $i \in [d_2]$  and combining the resulting inequality with (11) yields the bound:

$$\begin{aligned} \frac{1}{|V_1 \times V_2|^2} \sum_{(u,a),(v,b) \in V_1 \times V_2} K(f(u, a), g(v, b)) \\ \leq \frac{\gamma_+(G_1, K) \gamma_+(G_2, K)^2}{n_1 d_1 d_2^2} \sum_{\{u,v\} \in E_1} \sum_{i,j \in [d_2]} K(f(u, \kappa_{\pi_u(\{u,v\})}(i)), g(v, \kappa_{\pi_v(\{u,v\})}(j))) \\ = \frac{\gamma_+(G_1, K) \gamma_+(G_2, K)^2}{n_1 d_1 d_2^2} \sum_{\{u,a),(v,b) \in E(G_1 \otimes G_2)} K(f(u, a), g(v, b)). \end{aligned}$$

Hence we deduce the following theorem:

**Theorem 1** (Sub-multiplicativity). *Let  $G_1 = (V_1, E_1)$  be an  $n_1$ -vertex graph which is  $d_1$ -regular and let  $G_2 = (V_2, E_2)$  be a  $d_1$ -vertex graph which is  $d_2$ -regular. Then for every every kernel  $K : X \times X \rightarrow [0, \infty)$ ,*

$$\gamma_+(G_1 \otimes G_2, K) \leq \gamma_+(G_1, K) \cdot \gamma_+(G_2, K)^2.$$

In the special case  $X = \mathbb{R}$  and  $K(x, y) = (x - y)^2$ , Theorem 1 becomes:

$$\frac{1}{1 - \lambda(G_1 \otimes G_2)} \leq \frac{1}{1 - \lambda(G_1)} \cdot \frac{1}{(1 - \lambda(G_2))^2}. \quad (13)$$

Thus  $\lambda(G_1 \otimes G_2) \leq f(\lambda(G_1), \lambda(G_2))$  where  $f(\lambda_1, \lambda_2) < 1$  when  $\lambda_1, \lambda_2 \in (0, 1)$ . This is the main result of Reingold-Vadhan-Wigderson [38], and it coincides with the later bound of Reingold-Trevisan-Vadhan [37]. We note that in [38] an improved bound for  $\lambda(G_1 \otimes G_2)$  is obtained which is better than the bound of [37] (and hence also (13)), though this improvement has not been used (so far) in the literature. Theorem 1 shows that the fact that zigzag products preserve spectral gaps has nothing to do with the underlying Euclidean geometry (or linear algebra) that was used in [38, 37]: this is a truly non-linear phenomenon which holds in much greater generality, and simply amounts to an iteration of the Poincaré inequality (7). In the full version of this paper we present versions of the sub-multiplicativity result in Theorem 1 for other types of graph products such as tensor products, replacement products [8, 38], balanced replacement products [38], and derandomized square [40].

**2.2. An iterative construction of super-expanders.** Let us briefly recall how the Euclidean case of Theorem 1 was used by Reingold-Vadhan-Wigderson [38] to construct expanders (see also the nice exposition in Section 9.2 of the survey [11]).

For a graph  $G = (V, E)$  and for  $t \in \mathbb{N}$ , let  $G^t$  be the graph in which an edge between  $u, v \in V$  is drawn for every walk in  $G$  of length  $t$  whose endpoints are  $u, v$ . Thus  $A_{G^t} = (A_G)^t$  and if  $G$  is  $d$ -regular then  $G^t$  is  $d^t$ -regular.

Let  $H$  be an arbitrary regular graph with  $n_0$  vertices and degree  $d_0$ . Such a graph  $H$  will be called the *base graph* in what follows. Fix  $t_0 \in \mathbb{N}$ . Define  $G_1 = H^2$  and inductively  $G_{i+1} = G_i^{t_0} \otimes H$  (in [38] it sufficed to take  $t_0 = 2$ , but we will quickly realize that in the non-linear setting we need to work with higher powers, which is the reason why we are stating the construction for general  $t_0$ ). For this construction to be feasible we need to ensure that  $n_0 = d_0^{2t_0}$ , since in that case for all  $i \in \mathbb{N}$  the graph  $G_i$  is well defined and has  $n_0^i = d_0^{2it_0}$  vertices and degree  $d_0^2$ . Using the fact that  $\lambda(G^{t_0}) = \lambda(G)^{t_0}$ , we see that  $\lambda(G_1) = \lambda(H)^2$  and it follows from (13) that for  $i \in \mathbb{N}$  we have  $\lambda(G_{i+1}) \leq 1 - (1 - \lambda(G_i)^{t_0}) (1 - \lambda(H))^2$ . We thus see inductively that if  $\lambda(H) \leq 1 - (2 - 2^{1-t_0})^{-1/2}$  then  $\lambda(G_i) \leq \frac{1}{2}$  for all  $i \in \mathbb{N}$ .

Given Theorem 1, the above scheme suggests that we can repeat the Reingold-Vadhan-Wigderson iteration for every kernel  $K$  for which a sufficiently good base graph  $H$  exists. There is, however, an obstacle to this approach: we used above the identity  $\lambda(A^t) = \lambda(A)^t$  ( $t \in \mathbb{N}$ ) in order to increase the spectral gap of  $G_i$  in each step of the iteration. While this identity is a trivial corollary of spectral calculus, and was thus the “trivial part” of the construction in [38], there is no reason to expect that  $\gamma_+(A^t, K)$  decreases similarly with  $t$  for other kernels  $K : X \times X \rightarrow [0, \infty)$ . To better grasp what is happening here let us examine the asymptotic behavior of  $\gamma_+(A^t, |\cdot|^2)$  as a function of  $t$  (here and in what follows  $|\cdot|$  denotes the absolute value on  $\mathbb{R}$ ):

$$\begin{aligned} \gamma_+(A^t, |\cdot|^2) &= \frac{1}{1 - \lambda(A^t)} = \frac{1}{1 - \lambda(A)^t} \\ &= \frac{1}{1 - \left(1 - \frac{1}{\gamma_+(A, |\cdot|^2)}\right)^t} \asymp \max \left\{ 1, \frac{\gamma_+(A, |\cdot|^2)}{t} \right\}, \end{aligned} \quad (14)$$

where above, and in what follows,  $\asymp$  denotes equivalence up to universal constants (we will also use the notation  $\lesssim, \gtrsim$  to express the corresponding inequalities up to universal constants). Restating (14) in words, raising a matrix to a large power  $t \in \mathbb{N}$  corresponds to decreasing its (real) Poincaré constant by a factor of  $t$  as long as it is possible to do so (note that the Poincaré constant is necessarily at least 1).

For our scheme to work for other kernels  $K : X \times X \rightarrow [0, \infty)$  we would like  $K$  to satisfy a “spectral calculus” inequality of this type, i.e., an inequality which ensures that, if  $\gamma_+(A, K)$  is large, then  $\gamma_+(A^t, K)$  is much smaller than  $\gamma_+(A, K)$  for sufficiently large  $t \in \mathbb{N}$ . This is, in fact, not the case in general: in the Section 3.3 we construct a metric space  $(X, d_X)$  such that for each  $n \in \mathbb{N}$  there is a symmetric stochastic matrix  $A_n$  such that  $\gamma_+(A_n, d_X^2) \geq n$  yet for every  $t \in \mathbb{N}$  there is  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$  we have  $\gamma_+(A_n^t, d_X^2) \geq \frac{1}{2} \gamma_+(A_n, d_X^2)$ . The question which metric spaces satisfy the required spectral calculus inequalities thus becomes a subtle issue which we believe is of fundamental importance, beyond the particular application that we present here. A large part of the present paper is devoted to addressing this question. We obtain rather satisfactory results which allow us to carry out a zigzag type construction of super-expanders — expanders with respect to super-reflexive spaces — though we are still quite far from a complete understanding of the behavior of non-linear spectral gaps under graph powers for non-Euclidean geometries.



We will introduce in Section 2.3 a criterion on a metric space  $(X, d_X)$ , which is a bi-Lipschitz invariant, and we prove that it implies that for every  $n, m \in \mathbb{N}$  and every  $n \times n$  symmetric stochastic matrix  $A$  the Cesàro averages  $\frac{1}{m} \sum_{t=0}^{m-1} A^t$  satisfy the following spectral calculus inequality:

$$\gamma_+ \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t, d_X^2 \right) \leq C(X) \max \left\{ 1, \frac{\gamma_+(A, d_X^2)}{m^{\varepsilon(X)}} \right\}, \quad (15)$$

where  $C(X), \varepsilon(X) \in (0, \infty)$  depend only on the geometry of  $X$  but not on  $n, m, A$ . The fact that we can only prove such an inequality for Cesàro averages rather than powers does not create any difficulty in the ensuing argument, as we shall see presently. We postpone the discussion of the metric criterion which implies (15) to Section 2.3. It suffices to say here that we show that certain classes of metric spaces, including super-reflexive normed spaces, satisfy this criterion, and hence also (15).

Note that Cesàro averages have the following combinatorial interpretation in the case of graphs: given an  $n$ -vertex  $d$ -regular graph  $G = (V, E)$  let  $\mathcal{A}_m(G)$  be the graph whose vertex set is  $V$  and for every  $t \in \{0, \dots, m-1\}$  and  $u, v \in V$  we draw  $d^{m-1-t}$  edges joining  $u, v$  for every walk in  $G$  of length  $t$  which starts at  $u$  and terminates at  $v$ . With this definition  $A_{\mathcal{A}_m(G)} = \frac{1}{m} \sum_{t=0}^{m-1} A_G^t$ , and  $\mathcal{A}_m(G)$  is  $md^{m-1}$ -regular. We will slightly abuse this notation by also using the shorthand  $\mathcal{A}_m(A) = \frac{1}{m} \sum_{t=0}^{m-1} A^t$ , when  $A$  is a matrix.

Another issue which we must overcome is the existence of a sufficiently good base graph  $H$  for our zigzag iteration. This is also a subtle point which is discussed in Section 2.4 below. In order to present our construction here, we first state the following lemma, and postpone the discussion about it to Section 2.4:

**Lemma 2.1** (Existence of base graphs). *There exists an unbounded increasing sequence of integers  $\{n_i\}_{i=1}^{\infty}$  satisfying  $n_{i+1} \leq 100n_i$  with the following property: for every  $\delta \in (0, 1)$  there is a sequence of regular graphs  $\{H_i(\delta)\}_{i=1}^{\infty}$  such that  $H_i(\delta)$  has  $n_i$  vertices and degree  $d_i(\delta) \leq e^{(\log n_i)^{1-\delta}}$ . Moreover, for every super-reflexive (even just  $K$ -convex) normed space  $(X, \|\cdot\|_X)$  there exist  $\delta_0(X), \gamma(X) \in (0, 1)$  such that for all  $0 < \delta \leq \delta_0(X)$  and all  $i \in \mathbb{N}$ , we have  $\gamma_+(H_i(\delta), \|\cdot\|_X^2) \leq \gamma(X)$ .*

We briefly discuss a simple operation, called *edge completion*, which allows us to change the degree of a graph while not changing its Poincaré constant by too much. This operation will generate a lot of freedom in applying zigzag products, since we will be able to apply it in order to adjust the degrees of graphs so that their zigzag product becomes well-defined. Given a  $d$ -regular graph  $G = (V, E)$  and an integer  $d' \geq d$  define the  $d'$ -edge completion of  $G$ , denoted  $\mathcal{C}_{d'}(G) = (V, E')$ , as follows. Writing  $d' = \ell d + r$  where  $\ell \in \mathbb{N}$ , and  $r \in \{0, \dots, d-1\}$ , duplicate each edge in  $G$   $\ell$  times, and add  $r$  self-loops to each vertex of  $G$ . This makes  $\mathcal{C}_{d'}(G)$  a  $d'$ -regular graph. It is evident from the definition of  $\gamma_+(\cdot, \cdot)$  that for every kernel  $K : X \times X \rightarrow [0, \infty)$  we have  $\gamma_+(\mathcal{C}_{d'}(G), K) \leq 2\gamma_+(G, K)$ .

With these facts at hand, we can now construct an expander with respect to a given super-reflexive Banach space  $X$ .

**Lemma 2.2.** *For every super-reflexive Banach space  $X$  there exist  $C'(X) \geq 1, d = d(X) \in \mathbb{N}$  and sequence of  $d$ -regular graphs  $\{F_j(X)\}_{j=1}^{\infty}$  whose number of vertices tend to  $\infty$  with  $j$ , such that  $\gamma_+(F_j(X), \|\cdot\|_X^2) \leq C'(X)$  for every  $j \in \mathbb{N}$ .*

*Proof.* Fix  $k \in \mathbb{N}$  which will be determined presently. Write  $t_0 = (2k)^{3k}$ , and let  $i_0$  be the smallest  $i \in \mathbb{N}$  such that  $n_i \geq e^{(4t_0)^k}$ . Note that this implies that  $n_{i_0} \leq 100e^{(4t_0)^k}$ . We shall now define inductively a sequence of graphs  $\{F_j(k)\}_{j=0}^\infty$  as follows:  $F_0(k) = H_{i_0}(1/k)$  (from Lemma 2.1). The degree of  $F_0(k)$  is  $d_{i_0} \leq e^{(\log n_{i_0})^{1-1/k}}$  and it has  $n_{i_0}$  vertices. Define  $F_1(k) = \mathcal{C}_{d_{i_0}^2}(F_0(k))$ , so that the degree of  $F_1(k)$  is  $d_{i_0}^2$ . Assume inductively that we defined  $F_j(k)$  to be a graph with  $n_{i_0}^j$  vertices and degree  $d_{i_0}^{2^j}$ . Then  $\mathcal{A}_{t_0}(F_j(k))$  has  $n_{i_0}^j$  vertices and degree  $t_0 d_{i_0}^{2(t_0-1)} \leq n_{i_0}$  (where we used our choice of  $i_0$ ). It follows that we can form the edge completion  $\mathcal{C}_{n_{i_0}}(\mathcal{A}_{t_0}(F_j(k)))$ , which has degree  $n_{i_0}$ , and therefore it is possible to form the zigzag product

$$F_{j+1}(k) = (\mathcal{C}_{n_{i_0}}(\mathcal{A}_{t_0}(F_j(k)))) \circledast F_0(k),$$

which has degree  $d_{i_0}^{2^{j+1}}$  and  $n_{i_0}^{j+1}$  vertices.

Let  $(X, \|\cdot\|_X)$  be a super-reflexive normed space and let  $\delta_0(X), \gamma(X)$  and  $\varepsilon(X), C(X)$  be as in Lemma 2.1 and (15), respectively. Assume that  $k \geq 1/\delta_0(X)$ . Then by Lemma 2.1 we have  $\gamma_+(F_0(k), \|\cdot\|_X^2) \leq \gamma(X)$ , and therefore also by the basic property of edge completion,  $\gamma_+(F_1(k), \|\cdot\|_X^2) \leq 2\gamma(X)$ . We can now apply Theorem 1 to get the following recursive estimate:

$$\begin{aligned} \gamma_+(F_{j+1}(k), \|\cdot\|_X^2) &\leq \gamma_+(\mathcal{C}_{n_{i_0}}(\mathcal{A}_{t_0}(F_j(k))), \|\cdot\|_X^2) \gamma_+(F_0(k), \|\cdot\|_X^2)^2 \\ &\leq 2\gamma_+(\mathcal{A}_{t_0}(F_j(k)), \|\cdot\|_X^2) \gamma(X)^2 \\ &\stackrel{(15)}{\leq} 2C(X)\gamma(X)^2 \max\left\{1, \frac{\gamma_+(F_j(k), \|\cdot\|_X^2)}{t_0^{\varepsilon(X)}}\right\} \\ &= 2C(X)\gamma(X)^2 \max\left\{1, \frac{\gamma_+(F_j(k), \|\cdot\|_X^2)}{(2k)^{3k\varepsilon(X)}}\right\}. \end{aligned}$$

It follows by induction that if  $k \geq \max\{C(X), \gamma(X), 1/\varepsilon(X), 1/\delta_0(X)\}$  then for all  $j \in \mathbb{N}$  we have  $\gamma_+(F_j(k), \|\cdot\|_X^2) \leq 2C(X)\gamma(X)^2$ .  $\square$

Lemma 2.2 is not the final word. Our goal is to construct a super-expander, i.e., one sequence  $\{H_i\}_{i=1}^\infty$  of bounded degree graphs which is expander family with respect to *all* super-reflexive Banach spaces. In order to do so, we begin by examining more closely the proof of Theorem 2.2 and the parameters in Lemma 2.1 and in (15).

The proof of Theorem 2.2 gives (i) for each  $k \in \mathbb{N}$  a sequence of graphs  $\{F_j(k)\}_{j=0}^\infty$ , such that  $F_j(k)$  has degree  $d_k$  and  $n_j(k)$  vertices, where  $\{n_j(k)\}_{j=0}^\infty$  is a strictly increasing sequence of integers; (ii) a nondecreasing sequence  $\{C'_k\}_{k=1}^\infty$  such that for every super-reflexive Banach space  $X$ , there exists  $k_1(X) \in \mathbb{N}$  such that for every integer  $k \geq k_1(X)$  and  $j \in \mathbb{N}$ ,

$$\gamma_+(F_j(k), \|\cdot\|_X^2) \leq C'_{k_1(X)}. \quad (16)$$

From the proof of (15) in Section 2.3 we infer that there exist a nondecreasing sequence  $\{C''_k\}_{k=1}^\infty \subseteq (0, \infty)$  and a nonincreasing sequence  $\{\varepsilon_k\}_{k=1}^\infty \subseteq (0, 1)$ , such that for any super-reflexive Banach space  $X$  there exists  $k_2(X) \in \mathbb{N}$  such that for any regular graph  $G$  and  $t \in \mathbb{N}$ ,

$$\gamma_+(\mathcal{A}_t(G), \|\cdot\|_X^2) \leq C''_{k_2(X)} \cdot \max\left\{1, \frac{\gamma_+(G, \|\cdot\|_X^2)}{t^{\varepsilon_{k_2(X)}}}\right\}. \quad (17)$$

The analogues of properties (16) and (17) for a general family of kernels suffice to prove the following lemma:

**Lemma 2.3.** *Let  $\{F_j(k)\}_{j,k \in \mathbb{N}}$  be graphs as described before (16), i.e.,  $F_j(k)$  has degree  $d_k$  and  $n_j(k)$  vertices, where  $\{n_j(k)\}_{j=0}^\infty$  is strictly increasing. Assume that  $\mathcal{K}$  is a family of kernels such that  $\gamma_+(F_j(k), K) < \infty$  for all  $K \in \mathcal{K}$  and  $j, k \in \mathbb{N}$ . Assume also that  $\mathcal{K}$  satisfies (16) and (17), i.e.,*

- *There is a nondecreasing sequence  $\{C'_k\}_{k=1}^\infty$  such that for every  $K \in \mathcal{K}$  there exists  $k_1(K) \in \mathbb{N}$  such that for every integer  $k \geq k_1(K)$  and  $j \in \mathbb{N}$ ,*

$$\gamma_+(F_j(k), K) \leq C'_{k_1(K)}. \quad (18)$$

- *There exist a nondecreasing sequence  $\{C''_k\}_{k=1}^\infty \subseteq (0, \infty)$  and a nonincreasing sequence  $\{\varepsilon_k\}_{k=1}^\infty \subseteq (0, 1)$ , such that for any  $K \in \mathcal{K}$  there exists  $k_2(K) \in \mathbb{N}$  such that for any regular graph  $G$  and  $t \in \mathbb{N}$ ,*

$$\gamma_+(\mathcal{A}_t(G), K) \leq C''_{k_2(K)} \cdot \max \left\{ 1, \frac{\gamma_+(G, K)}{t^{\varepsilon_{k_2(K)}}} \right\}. \quad (19)$$

Then there exists  $d \in \mathbb{N}$  and a sequence of  $d$ -regular graphs  $\{H_i\}_{i=1}^\infty$  whose number of vertices tends to  $\infty$  with  $i$ , such that for every  $K \in \mathcal{K}$  there exists  $C(K) \in (0, \infty)$  satisfying  $\gamma_+(H_i, K) \leq C(K)$  for all  $i \in \mathbb{N}$ .

*Proof.* Set  $C_k = \max \{C'_k, C''_k\}$  and  $m_k = \lceil (2C_k^3)^{1/\varepsilon_k} \rceil$ . Let  $j(k)$  be the smallest  $j \in \mathbb{N}$  such that  $n_j(k) > \max \{k, m_{k+1} d_{k+1}^{2m_{k+1}}\}$ . We define  $L_k = F_{j(k)}(k)$ . While the sequence  $\{\gamma_+(L_k, K)\}_{k=1}^\infty$  is bounded for any  $K \in \mathcal{K}$ , this is not yet the desired graph sequence since its degrees are unbounded.

We next define for every  $k \in \mathbb{N}$  a sequence of graphs  $L_{k,0}, L_{k,1}, \dots, L_{k,\ell(k)}$  as follows. Along with  $\{L_{k,j}\}_{j=0}^{\ell(k)}$  we shall also define an auxiliary sequence  $\{h_j(k)\}_{j=0}^{\ell(k)-1} \subseteq \mathbb{N}$ . Set  $L_{k,0} = L_k$  and  $h_0(k) = k$ . Define  $h_1(k)$  to be the smallest  $h \in \mathbb{N}$  such that  $n_{j(h)}(h) > d_{h_0(k)}$ . From the definition of  $j(\cdot)$ , we have that  $h_1(k) < h_0(k)$ . Define

$$L_{k,1} = \mathcal{A}_{m_{h_1(k)}} \left( \mathcal{C}_{n_{j(h_1(k))}(h_1(k))}(L_{k,0}) \otimes L_{h_1(k)} \right).$$

Note that the degree of  $L_{k,1}$  is  $m_{h_1(k)} d_{h_1(k)}^{2m_{h_1(k)}}$ .

Inductively, assume we have already defined  $L_{k,i-1}$ , and  $h_{i-1}(k)$ , and that the degree of  $L_{k,i-1}$  is  $m_{h_{i-1}(k)} d_{h_{i-1}(k)}^{2m_{h_{i-1}(k)}}$ . Define  $h_i(k)$  to be the smallest  $h \in \mathbb{N}$  such that  $n_{j(h)}(h) > m_{h_{i-1}(k)}^2 d_{h_{i-1}(k)}^{2m_{h_{i-1}(k)}}$ . From the definition of  $j(\cdot)$ , we have that  $h_i(k) < h_{i-1}(k)$ . Define

$$L_{k,i} = \mathcal{A}_{m_{h_i(k)}} \left( \mathcal{C}_{n_{j(h_i(k))}(h_i(k))}(L_{k,i-1}) \otimes L_{h_i(k)} \right).$$

Note that the degree of  $L_{k,i}$  is  $m_{h_i(k)} d_{h_i(k)}^{2m_{h_i(k)}}$ .

Continue this way, until  $h_i(k) = 1$ . At that point, set  $\ell(k) = i + 1$ , and define

$$L_{k,\ell(k)} = \mathcal{A}_{m_1}(\mathcal{C}_{n_{j(1)}(1)}(L_{k,\ell(k)-1}) \otimes L_1).$$

Hence  $L_{k,\ell(k)}$  has degree  $m_1 d_1^{2m_1}$  which is a universal constant. Define  $H_k = L_{k,\ell(k)}$ . We shall now show that  $\{H_k\}_{k=1}^\infty$  is the desired graph sequence.

Fix  $K \in \mathcal{K}$  and set  $k_0(K) = \max\{k_1(K), k_2(K)\}$ . Now fix  $k > k_0(K)$ . We shall first estimate  $\gamma_+(L_{k,i}, K)$ , when  $h_i(k) > k_0(K)$ . By induction on  $i$  we shall show that as long as  $h_i(k) > k_0(K)$ , we have  $\gamma_+(L_{k,i}, K) \leq C_{k_0(K)}$ . If  $h_0(k) = k > k_0(K) \geq k_1(K)$ , then by our assumption (18),  $\gamma_+(L_{k,0}, K) = \gamma_+(F_{j(k)}(k), K) \leq C_{k_0(K)}$ . Assume next that  $h_i(k) > k_0(K)$ , and inductively that  $\gamma_+(L_{k,i-1}, K) \leq C_{k_0(K)}$ . Then using Theorem 1 and (19),

$$\begin{aligned} \gamma_+(L_{k,i}, K) &\leq C_{k_0(K)} \cdot \max \left\{ 1, \frac{2\gamma_+(L_{k,i-1}, K) \cdot \gamma_+(L_{h_i(k)}, K)^2}{m_{h_i(k)}^{\varepsilon_{h_i(k)}}} \right\} \\ &\stackrel{(18)}{\leq} C_{k_0(K)} \max \left\{ 1, \frac{2C_{k_0(K)}^3}{m_{h_i(k)}^{\varepsilon_{h_i(k)}}} \right\} \leq C_{k_0(K)}. \end{aligned}$$

If we let  $i_0(k)$  be the largest  $i$  such that  $h_i(k) > k_0(K)$ , then we have just proved that  $\gamma_+(L_{k,i_0(k)}, K) \leq C_{k_0(K)}$ . For  $i > i_0(k)$ , we use shall only rough bounds:

$$\begin{aligned} \gamma_+(L_{k,i}, K) &\leq C_{k_0(K)} \cdot \max \left\{ 1, \frac{2\gamma_+(L_{k,i-1}, K) \cdot \gamma_+(L_{h_i(k)}, K)^2}{m_{h_i(k)}^{\varepsilon_{h_i(k)}}} \right\} \\ &\leq 2C_{k_0(K)} \gamma_+(L_{h_i(k)}, K)^2 \gamma_+(L_{k,i-1}, K). \quad (20) \end{aligned}$$

By iterating (20) we obtain the crude bound:

$$\gamma_+(H_k, K) = \gamma_+(L_{k,\ell(k)}, K) \leq (2C_{k_0(K)})^{k_0(K)} \prod_{i=1}^{k_0(K)} \gamma_+(L_i, K)^2. \quad (21)$$

The right hand side of (21) depends on  $K$ , but not on  $k$ . Hence  $\{\gamma_+(H_k, K)\}_{k=1}^\infty$  is bounded for every  $K \in \mathcal{K}$ . It remains to argue that the number of vertices of  $H_k$  tends to  $\infty$  with  $k$ . Indeed, the number of vertices in  $H_k$  is at least as the number of vertices of  $L_k = F_{j(k)}(k)$  which is  $n_{j(k)}(k)$ , and from the definition of  $j(k)$ , we have  $n_{j(k)}(k) > k$ .  $\square$

**Theorem 2** (Existence of super-expanders). *There exists a sequence of 9-regular graphs  $\{G_i\}_{i=1}^\infty$ , whose number of vertices tends to  $\infty$  with  $i$ , such that for any super-reflexive Banach space  $(X, \|\cdot\|_X)$  there is  $C(X) \in (0, \infty)$ , such that  $\gamma_+(G_i, \|\cdot\|_X^2) \leq C(X)$  for every  $i \in \mathbb{N}$ .*

*Proof.* From the discussion before Lemma 2.3, and by Lemma 2.3 itself, we conclude that there exists a universal constant  $d \in \mathbb{N}$ , and  $d$ -regular super-expander family  $\{H_k\}_{k=1}^\infty$ , with  $\gamma_+(H_k, \|\cdot\|_X^2) \leq C(X)$ .

Let  $C_m^\circ$  be the  $m$ -vertex cycle with self-loops. It is an easy consequence of the triangle inequality that for every metric space  $(Y, d_Y)$  we have  $\gamma_+(C_m^\circ, d_Y^2) \leq 4m^2$ . Thus, using Theorem 1 once more we see that for all  $k \in \mathbb{N}$  we have

$$\gamma_+(H_k \otimes C_d^\circ, \|\cdot\|_X^2) \leq 16d^4 C(X),$$

which is a constant depending only on  $X$  but not on  $k$ . Hence  $\{H_k \otimes C_d^\circ\}_{k=1}^\infty$  is the required sequence of 9-regular super-expanders.  $\square$

*Remark 2.1.* Similarly, the sequence  $\{(H_k \otimes C_d^\circ) \oplus C_9\}_{k=1}^\infty$  is a 3-regular super-expander, where,  $C_9$  is the cycle on 9 vertices (without self-loops), and “ $\oplus$ ” is the replacement (graph) product (cf. [38]). In the full version of this extended abstract we analyze the replacement product similarly to the zigzag product analysis presented here.

*Remark 2.2.* Lafforgue [17] asked whether there exists a sequence of bounded degree graphs  $\{G_k\}_{k=1}^\infty$  which does not admit a coarse embedding (with uniform moduli) into any  $K$ -convex Banach space. An examination of Lafforgue's argument shows that his method produces graphs  $\{H_j(k)\}_{j,k \in \mathbb{N}}$  such that for each  $k \in \mathbb{N}$  the graphs  $\{H_j(k)\}_{j \in \mathbb{N}}$  have degree  $d_k$ , their cardinalities are unbounded, and for every  $K$ -convex Banach space  $(X, \|\cdot\|_X)$  there is some  $k \in \mathbb{N}$  for which  $\sup_{j \in \mathbb{N}} \gamma_+(H_j(k), \|\cdot\|_X^2) < \infty$ . The problem is that the degrees  $\{d_k\}_{k \in \mathbb{N}}$  are unbounded, but this can be overcome as above by applying the zigzag product with a cycle with self-loops. Indeed, define  $G_j(k) = H_j(k) \textcircled{Z} C_{d_k}^\circ$ . Then  $G_j(k)$  is 9-regular, and as argued in the proof of Theorem 2, we still have  $\sup_{j \in \mathbb{N}} \gamma_+(G_j(k), \|\cdot\|_X^2) < \infty$ . To get a single sequence of graphs which does not admit a coarse embedding into any  $K$ -convex Banach space, fix a bijection  $\psi = (a, b) : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ , and define  $G_m = G_{a(m)}(b(m))$ . The graphs  $G_m$  all have degree 9. If  $X$  is  $K$ -convex then choose  $k \in \mathbb{N}$  as above. If we let  $m_j \in \mathbb{N}$  be such that  $\psi(m_j) = (j, k)$  then we have shown that the graphs  $\{G_{m_j}\}_{j=1}^\infty$  are arbitrarily large, have bounded degree, and satisfy  $\sup_{j \in \mathbb{N}} \gamma_+(G_{m_j}, \|\cdot\|_X^2) < \infty$ . The argument that was presented in Section 1.1 implies that  $\{G_m\}_{m=1}^\infty$  do not embed coarsely into  $X$ .

**2.3. Cesàro averages of matrices: coping with new issues that arise from non-linearity.** The goal in this section is to prove (15), i.e., the spectral calculus inequality for Cesàro averages in super-reflexive spaces.

Recall that we use the following notation  $\mathcal{A}_m(A) = \frac{1}{m} \sum_{t=0}^{m-1} A^t$ . Fix  $\varepsilon > 0$  and  $C \in (0, \infty)$ . Assume that a metric space  $(X, d_X)$  satisfies the following property, which we call *metric Markov cotype  $\frac{2}{\varepsilon}$  with constant  $C$  and exponent 2*: for every  $m, n \in \mathbb{N}$ , every symmetric stochastic  $n \times n$  matrix  $A = (a_{ij})$ , and every  $x_1, \dots, x_n \in X$  there exist  $y_1, \dots, y_n \in X$  which satisfy the following inequality:

$$\sum_{i=1}^n d_X(x_i, y_i)^2 + m^\varepsilon \sum_{i=1}^n \sum_{j=1}^n a_{ij} d_X(y_i, y_j)^2 \leq C^2 \sum_{i=1}^n \sum_{j=1}^n \mathcal{A}_m(A)_{ij} d_X(x_i, x_j)^2. \quad (22)$$

The origin of this (admittedly cumbersome) name comes from a key *linear* property of normed spaces that was introduced by Ball [1] under the name of *Markov cotype* in his profound work on the Lipschitz extension problem. We will see below that any super-reflexive normed space has metric Markov cotype  $\frac{2}{\varepsilon}$  with constant  $C$  and exponent 2 for some  $\varepsilon \in (0, 1]$  and  $C \in (0, \infty)$  (our proof uses Ball's insights in [1]).

Informally, (22) means that the average square distance of the edges in an embedding of the graph  $\frac{1}{m} \sum_{t=0}^{m-1} A^t$  into  $X$  is larger than the average square distance of the edges in the graph  $A$  by a factor of  $m^\varepsilon$ , in a different embedding of the graph  $A$  into  $X$ , which is near the original embedding. This is formalized in the following claim whose proof is in Section 3.1.

**Claim 2.4.** *Assuming that  $(X, d_X)$  has metric Markov cotype  $\frac{2}{\varepsilon}$  with constant  $C$  and exponent 2, we have:*

$$\gamma(\mathcal{A}_m(A), d_X^2) \leq 12C^2 \max \left\{ 1, \frac{\gamma(A, d_X^2)}{m^\varepsilon} \right\}. \quad (23)$$

In order to deduce (15), which is the variant of (23) for absolute spectral gaps, we need the following two easy consequences of the triangle inequality. The proofs appear in Section 3.1.

**Claim 2.5.** *For every doubly stochastic  $n \times n$  matrix  $A$ , and any metric space  $X$ ,*

$$\frac{2}{5} \gamma \left( \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2 \right) \leq \gamma_+(A, d_X^2) \leq 2\gamma \left( \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2 \right). \quad (24)$$

**Claim 2.6.** For every doubly stochastic  $n \times n$  matrix  $A$ , and any metric space  $X$ ,

$$\gamma \left( \begin{pmatrix} 0 & \mathcal{A}_m(A) \\ \mathcal{A}_m(A) & 0 \end{pmatrix}, d_X^2 \right) \leq 9\gamma \left( \mathcal{A}_m \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2 \right). \quad (25)$$

By combining (23) with (24) and (25) we deduce (15) with  $C(X) \leq 540C^2$  whenever  $(X, d_X)$  satisfies (22).

It remains to explain why any super-reflexive normed space  $(X, \|\cdot\|_X)$  has metric Markov cotype  $\frac{2}{\varepsilon}$  with constant  $C$  and exponent 2 for some  $\varepsilon, C \in (0, \infty)$ . An important theorem of Pisier [31] says that  $X$  has an equivalent norm whose modulus of uniform convexity has power type  $p$  for some  $p \in [2, \infty)$ . From this fact, together with results from [6, 2], we deduce the following variant of Pisier's martingale inequality [31], which is proved in Section 3.2:

$$\sum_{k=1}^m \mathbb{E} [\|M_k - M_{k-1}\|_X^2] \lesssim m^{1-\frac{2}{p}} \mathbb{E} [\|M_m - M_0\|_X^2], \quad (26)$$

for every square integrable martingale  $\{M_k\}_{k=0}^m \subseteq X$ , where the implied constant in (26) depends on  $X$ .

In order to deduce (22) from (26) fix  $x_1, \dots, x_n \in X$  and define  $f : \{1, \dots, n\} \rightarrow X$  by  $f(i) = x_i$ . For every  $\ell \in \{1, \dots, n\}$  let  $\{Z_t^{(\ell)}\}_{t=0}^m$  be the Markov chain on  $\{1, \dots, n\}$  which starts at  $\ell$  and has transition matrix  $A$ . In other words  $Z_0^{(\ell)} = \ell$  with probability 1 and for  $t \in \{1, \dots, m\}$ ,  $i, j \in \{1, \dots, n\}$  we have  $\Pr[Z_t^{(\ell)} = j | Z_{t-1}^{(\ell)} = i] = a_{ij}$ . For  $t \in \{0, \dots, m\}$  define  $f_t : \{1, \dots, n\} \rightarrow X$  by  $f_t(i) = \sum_{j=1}^n (A^{m-t})_{ij} f(j)$ . A simple computation shows that if we set  $M_t^{(\ell)} = f_t(Z_t^{(\ell)})$  then  $\{M_t^{(\ell)}\}_{t=0}^m$  is a martingale with respect to the filtration induced by  $\{Z_t^{(\ell)}\}_{t=0}^m$ . We can therefore apply (26) to this martingale, and then average over  $\ell \in \{1, \dots, n\}$ , to get the inequality:

$$\sum_{t=1}^m \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|f_t(i) - f_{t-1}(j)\|_X^2 \lesssim m^{1-\frac{2}{p}} \sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \left\| x_i - \sum_{r=1}^n (A^m)_{jr} x_r \right\|_X^2. \quad (27)$$

If we choose  $y_i = \frac{1}{m} \sum_{s=0}^{m-1} (A^s)_{ij} x_j$  for  $i \in \{1, \dots, n\}$ , then we can use convexity to deduce the bound:

$$\begin{aligned} \sum_{t=1}^m \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|f_t(i) - f_{t-1}(j)\|_X^2 &\geq m \sum_{i=1}^n \sum_{j=1}^n a_{ij} \left\| \frac{1}{m} \sum_{t=1}^m (f_t(i) - f_{t-1}(j)) \right\|_X^2 \\ &= m \sum_{i=1}^n \sum_{j=1}^n a_{ij} \left\| y_i - y_j - \frac{1}{m} \sum_{r=1}^n (A^m)_{jr} (x_j - x_r) \right\|_X^2 \\ &\geq \frac{m}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|y_i - y_j\|_X^2 - \frac{1}{m} \sum_{j=1}^n \sum_{r=1}^n (A^m)_{jr} \|x_j - x_r\|_X^2. \end{aligned} \quad (28)$$

At the same time we can bound the right-hand side of (27) as follows:

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \left\| x_i - \sum_{r=1}^n (A^m)_{jr} x_r \right\|_X^2 &\leq \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^n (A^m)_{ij} (A^m)_{jr} \|x_i - x_r\|_X^2 \\
&\leq 2 \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^n (A^m)_{ij} (A^m)_{jr} (\|x_i - x_j\|_X^2 + \|x_j - x_r\|_X^2) \\
&= 4 \sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \|x_i - x_j\|_X^2. \tag{29}
\end{aligned}$$

We note that:

$$\begin{aligned}
\sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \|x_i - x_j\|_X^2 &= \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t A^{m-t} \right)_{ij} \|x_i - x_j\|_X^2 \\
&\leq \frac{2}{m} \sum_{i=1}^n \sum_{j=1}^n \sum_{r=1}^n \sum_{t=0}^{m-1} (A^t)_{ir} (A^{m-t})_{rj} (\|x_i - x_r\|_X^2 + \|x_r - x_j\|_X^2) \\
&= 2 \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t \right)_{ij} \|x_i - x_j\|_X^2 + 2 \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=1}^m A^t \right)_{ij} \|x_i - x_j\|_X^2 \\
&\leq 4 \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t \right)_{ij} \|x_i - x_j\|_X^2 + \frac{2}{m} \sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \|x_i - x_j\|_X^2,
\end{aligned}$$

which implies (by separating the cases  $m \geq 4$  and  $m < 4$ ) the bound:

$$\sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \|x_i - x_j\|_X^2 \lesssim \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t \right)_{ij} \|x_i - x_j\|_X^2. \tag{30}$$

Substituting (28) and (29) into (27) yields the bound:

$$\begin{aligned}
m^{2/p} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \|y_i - y_j\|_X^2 &\lesssim \sum_{i=1}^n \sum_{j=1}^n (A^m)_{ij} \|x_i - x_j\|_X^2 \\
&\stackrel{(30)}{\lesssim} \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t \right)_{ij} \|x_i - x_j\|_X^2. \tag{31}
\end{aligned}$$

At the same time,

$$\begin{aligned}
\sum_{i=1}^n \|x_i - y_i\|_X^q &= \sum_{i=1}^n \sum_{j=1}^n \left\| \frac{1}{m} \sum_{t=0}^{m-1} (A^t)_{ij} (x_i - x_j) \right\|_X^q \\
&\leq \sum_{i=1}^n \sum_{j=1}^n \left( \frac{1}{m} \sum_{t=0}^{m-1} A^t \right)_{ij} \|x_i - x_j\|_X^q. \tag{32}
\end{aligned}$$

Inequalities (31) and (32) imply (22) with  $\varepsilon = \frac{2}{p}$ , as required.

**2.4. Constructing the base graph: the heat semigroup on the tail space.** In this section we outline the proof of Lemma 2.1 — the existence of “sufficiently good” base graphs for the class of  $K$ -convex spaces (which contain the class of super-reflexive spaces as a subclass). Some of the details are deferred to Section 3.4.

We begin with few definitions. Let  $\mathbb{F}_2 = \{0, 1\}$  be the field of two elements, and  $(X, \|\cdot\|_X)$  a given normed space.  $L_p(X)$  denotes the normed space of functions  $f : V \rightarrow X$ , where  $V$  is an (implicit) finite set (in this paper,  $V$  is the vertex set of some graph, and in this section it is  $V = \mathbb{F}_2^n$ ) and the norm is  $\|f\|_{L_p(X)} = \left(\frac{1}{|V|} \sum_{x \in V} \|f(x)\|_X^p\right)^{1/p}$ . Given a  $V \times V$  symmetric stochastic matrix  $A$ , we view  $A$  as a linear operator over  $L_p(X)$ : For  $f \in L_p(X)$  we define  $Af$  as  $(Af)(i) = \sum_j A_{ij}f(j)$  — essentially we identify  $A$  with  $A \otimes I_X$ . For  $f \in L_2(X)$ , let  $\widehat{f}(A) = \mathbb{E}[f(x)W_A(x)]$  be the Fourier coefficient corresponding to the Walsh function  $W_A(x) = (-1)^{\sum_{i \in A} x_i}$ , where the expectation is with respect to the uniform probability measure on  $\mathbb{F}_2^n$ . Let  $\partial_i f(x) = \frac{f(x+e_i) - f(x)}{2}$  (where  $\{e_i\}_{i=1}^n$  is the standard basis) and  $\Delta f = \sum_{i=1}^n \partial_i f$ .

The starting point of our construction is the example of Khot and Naor [16] of quotients of the hypercube by good codes, as examples of metric spaces for which Bourgain’s embedding theorem [4] is asymptotically sharp. It would be instructive to first recall the argument from [16]. Thinking of the cube  $\mathbb{F}_2^n$  as an  $n$ -dimensional vector space over  $\mathbb{F}_2$ , let  $C \subseteq \mathbb{F}_2^n$  be a linear subspace of dimension at least  $\frac{n}{10}$  such that the minimum number of 1’s in any non-zero element of  $C$  is  $m \geq \frac{n}{10}$  (a good code). Given a function  $f : \mathbb{F}_2^n/C^\perp \rightarrow X$  we will think of  $f$  as a function defined on all of  $\mathbb{F}_2^n$  which is constant on cosets of  $C^\perp$ . The following claim is a simple observation of [16]:

**Claim 2.7.** *Let  $C \subseteq \mathbb{F}_2^n$  be a linear code with minimum distance of  $m$ , and fix  $f : \mathbb{F}_2^n \rightarrow X$  which is constant on cosets of  $C^\perp$ . Then  $\widehat{f}(A) = 0$  for all nonempty  $A \subseteq \{1, \dots, n\}$  of cardinality less than  $m$ .*

Since  $\Delta W_A = |A|W_A$  and the (nonempty) Fourier spectrum of  $f$  is supported on sets of size at least  $m$ , it follows from Parseval’s identity that if  $X$  is a Hilbert space then  $\|\Delta f\|_{L_2(X)} \geq m\|f - \mathbb{E}f\|_{L_2(X)}$ . But  $\|\Delta f\|_{L_2(X)} = \|\sum_{i=1}^n \partial_i f\|_{L_2(X)} \leq \sum_{i=1}^n \|\partial_i f\|_{L_2(X)}$  and  $\|f - \mathbb{E}f\|_{L_2(X)}^2 = \frac{1}{2} \mathbb{E}_{x,y} [\|f(x) - f(y)\|_X^2]$ . An application of Cauchy-Schwarz and the fact that  $m \asymp n$  now implies that  $\mathbb{E}_{x,y} [\|f(x) - f(y)\|_X^2] \lesssim \frac{1}{n} \sum_{i=1}^n \|\partial_i f\|_{L_2(X)}^2$ . A moment of thought reveals that this is the desired Poincaré inequality (in the case of Hilbert space), where the graph in question is on the vertex set  $\mathbb{F}_2^n/C^\perp$ , and each coset  $x + C^\perp$  is joined by an edge to the cosets  $\{x + e_i + C^\perp\}_{i=1}^n$ . In fact, this graph has degree logarithmic in the number of vertices, which is much better than the assertion in Lemma 2.1.

In order to make this idea work for a non-Hilbertian normed space  $X$  it would be desirable to prove that if  $f : \mathbb{F}_2^n \rightarrow X$  is in the  $m$ -tail space, i.e., the subspace of  $L_2(X)$ , denoted  $L_2^{\geq m}(X)$ , consisting of functions  $f : \mathbb{F}_2^n \rightarrow X$  with  $\widehat{f}(A) = 0$  whenever  $|A| < m$ , then  $\|\Delta f\|_{L_2(X)} \gtrsim m\|f\|_{L_2(X)}$ . This fact is false without some additional assumption on the geometry of  $X$  — in the full version of this extended abstract we observe that it fails when  $X = L_1$ . Recall from Section 1.1 that a Banach space  $X$  is called  $K$ -convex if there exists  $\varepsilon_0 > 0$  and  $n_0 \in \mathbb{N}$  such that any embedding of  $\ell_1^{n_0}$  into  $X$  incurs distortion at least  $1 + \varepsilon_0$ . Here we will show that if  $X$  is  $K$ -convex then there exists  $\delta = \delta(X) > 0$  such that if



$f \in L_2^{\geq m}(X)$  then

$$\|\Delta f\|_{L_2(X)} \gtrsim m^\delta \|f\|_{L_2(X)}. \quad (33)$$

While this bound is insufficient for our purpose, from the the proof of (33) we can extract a graph which is more complicated than  $\mathbb{F}_2^n/C^\perp$  (but is of independent interest), and whose degree bound to what is claimed in Lemma 2.1, which will serve as the base graph(s) for  $K$ -convex spaces.

The key idea is to consider the heat semigroup  $\{T_t\}_{t>0}$  of operators given by

$$(T_t f)(x) = \sum_{A \subseteq [n]} e^{-t|A|} \hat{f}(A) W_A(x).$$

and make use of the following lemma, which is the heart of our argument.

**Lemma 2.8.** *If  $X$  is a  $K$ -convex Banach Space, and  $p > 1$ , then there exist  $\alpha \geq 1$ ,  $a > 0$ , such that for every  $n, m \in \mathbb{N}$ ,  $m \leq n$ , every  $t \geq 0$ , and every  $f : \mathbb{F}_2^n \rightarrow X$ ,  $f \in L_p^{\geq m}(X)$ ,*

$$\|T_t f\|_{L_p(X)} \lesssim e^{-am \min\{t, t^\alpha\}} \|f\|_{L_p(X)}, \quad (34)$$

The proof of Lemma 2.8 is based on deep results of Pisier on holomorphic extensions of the heat semigroup on  $K$ -convex spaces [33], and a quantitative version of the proof of a recent factorization theorem of Pisier [34]. The proof is contained in Section 3.4.1.

Note that when  $m \geq 1$ ,  $\Delta$  is invertible on  $L_p^{\geq m}(X)$ , and moreover we have the identity  $\Delta^{-1} = \int_0^\infty T_t dt$ , since  $\Delta^{-1} W_A = |A|^{-1} W_A$ , and

$$\left( \int_0^\infty T_t dt \right) W_A = \int_0^\infty T_t W_A dt = \int_0^\infty e^{-t|A|} W_A dt = |A|^{-1} W_A.$$

Hence by integrating (34) over  $t$ , it immediately implies (33).

Inequality (34) will serve as the basis for our base graph. This is made transparent by the observation that the heat semigroup can be viewed as a noise operator with noise rate  $\frac{1-e^{-t}}{2}$ . The following standard claim (whose proof is recalled in Section 3.4) formalizes this statement.

**Claim 2.9.** *For every  $f : \mathbb{F}_2^n \rightarrow X$ ,  $x \in \mathbb{F}_2^n$ , and  $t \in (0, \infty)$ ,*

$$(T_t f)(x) = \sum_{y \in \mathbb{F}_2^n} \left( \frac{1-e^{-t}}{2} \right)^{\|x-y\|_1} \cdot \left( \frac{1+e^{-t}}{2} \right)^{n-\|x-y\|_1} f(y). \quad (35)$$

Hence  $T_t$  induces a natural weighted graph structure on  $\mathbb{F}_2^n$ : the weight of the edge joining  $x, y \in \mathbb{F}_2^n$  is  $\left( \frac{1-e^{-t}}{2} \right)^{\|x-y\|_1} \left( \frac{1+e^{-t}}{2} \right)^{n-\|x-y\|_1}$ . Our base graph is ‘‘morally’’ the quotient of this weighted cube by  $C^\perp$ . When  $t \approx n^{-1/\alpha}$ , the factor in the right hand side of (34) becomes a constant, and this implies a Poincaré inequality because of the following general estimate (proved in Section 3.4).

**Proposition 2.10.** *Fix a Banach space  $X$ , a symmetric stochastic  $N \times N$  matrix  $A$ , and  $p \geq 1$ . Assume that there exists  $1 > \lambda > 0$  such that for every  $f : [N] \rightarrow X$  with  $\sum_{i=1}^N f(i) = 0$  we have,*

$$\|A f\|_{L_p(X)} \leq \lambda \|f\|_{L_p(X)}. \quad (36)$$

Then,

$$\gamma_+(A, \|\cdot\|_X^p) \leq 8^p (1-\lambda)^{-p}.$$

Still, “ $T_t/C^\perp$ ” is not a low degree unweighted graph. To make it an unweighted graph we first truncate the weights of  $T_t$  below an appropriate threshold, and approximate the remaining (large enough) weights using multiple parallel edges. This process is summarized in the following lemma (whose proof is deferred to the full version due to lack of space).

**Lemma 2.11.** *Let  $\tau = (1 - e^{-t})/2$ . There exists a Cayley graph on  $\mathbb{F}_2^n$ ,  $G = (\mathbb{F}_2^n, E)$ , having degree at most  $\tau^{-4\tau n}(1 - \tau)^{-(1-4\tau)n}$  such that for every metric space  $(X, d_X)$ ,  $p \geq 1$  and  $f, g : \mathbb{F}_2^n \rightarrow X$ ,*

$$\begin{aligned} \frac{1}{3|E|} \sum_{(x,y) \in E} d_X(f(x), g(y))^p \\ \leq \frac{1}{2^n} \sum_{x,y \in \mathbb{F}_2^n} (T_t)_{xy} d_X(f(x), g(y))^p \leq \frac{3}{|E|} \sum_{(x,y) \in E} d_X(f(x), g(y))^p, \end{aligned} \quad (37)$$

as long as  $18\tau^2 n \geq 2p \log n + \log 4$ .

Only then we pass to the quotient  $\mathbb{F}_2^n/C^\perp$  by identifying vertices within each coset and dividing the number of the resulting “quotient edges” between each coset by  $|C^\perp|$ . This is formally summarized as the proof of Lemma 2.1.

*Proof of Lemma 2.1.* The argument works for any power  $p \in (1, \infty)$  and not just  $p = 2$ . Fix a  $K$ -convex Banach space  $X$ , and let  $\alpha \geq 1$  and  $a > 0$  as in Lemma 2.8.

Fix  $n \in \mathbb{N}$ . Set  $t = \min\{n^{-1/\alpha}/a, 0.1\}$ ,  $\tau = (1 - e^{-t})/2$  and apply the discretization process — Lemma 2.11 — on  $T_t$ , from which we obtain a Cayley graph  $G = G_n = (\mathbb{F}_2^n, E)$  whose degree is at most

$$\tau^{-4\tau n}(1 - \tau)^{(1-4\tau)n} \leq \tau^{-8\tau n} \leq \max\left\{40^{320 \cdot 10^{10\alpha}}, n^{2n^{1-1/\alpha}/\alpha/a}\right\} \lesssim e^{n^{1-\delta}},$$

for some  $\delta = \delta(X) \in (0, 1)$ , depending only on  $\alpha$  and  $a$  (and hence on  $X$ ), but not on  $n$ .

Let  $C \subseteq \mathbb{F}_2^n$  be a “good” linear code of dimension at least  $n/10$  and minimum weight  $m \geq n/10$ . We define  $H = H_n = (\mathbb{F}_2^n/C^\perp, E/C^\perp)$ : the quotient of  $G$  by  $C^\perp$ , whose vertex-set is the set of cosets of  $C^\perp$  and in the edge (multi)set  $E/C^\perp$  the number of edges between two cosets  $x + C^\perp$  and  $y + C^\perp$ , is the number of edges in  $G$  with one endpoint in  $x + C^\perp$  and the other in  $y + C^\perp$  divided by  $|C^\perp|$ . Since  $C^\perp$  is a linear subspace in a vector space, the degree of  $H$  is the same as the degree of  $G$ .

Next, fix  $f : \mathbb{F}_2^n/C^\perp \rightarrow X$  with  $\sum_{x \in \mathbb{F}_2^n/C^\perp} f(x) = 0$ . Thinking on  $f$  as a function on  $\mathbb{F}_2^n$  which is constant on cosets of  $C^\perp$ , we have by Claim 2.7 that  $\widehat{f}(A) = 0$  for every  $\emptyset \neq A \subseteq \{1, \dots, n\}$ ,  $|A| < m$ . Hence by Lemma 2.8,

$$\|T_t f\|_{L_p(X)} \lesssim e^{-amt^\alpha} \|f\|_{L_p(X)} \leq \lambda \|f\|_{L_p(X)},$$

where  $0 < \lambda < 1$  is a universal constant. Since this is true for any  $f : \mathbb{F}_2^n/C^\perp \rightarrow X$  with  $\sum_{x \in \mathbb{F}_2^n/C^\perp} f(x) = 0$ , applying Proposition 2.10 we have that for every  $f, g : \mathbb{F}_2^n/C^\perp \rightarrow X$ ,

$$\begin{aligned} \frac{|C^\perp|^2}{2^{2n}} \sum_{x,y \in \mathbb{F}_2^n/C^\perp} \|f(x) - g(y)\|_X^2 \\ \lesssim \frac{|C^\perp|}{2^n} \sum_{x,y \in \mathbb{F}_2^n/C^\perp} \frac{1}{|C^\perp|} \sum_{a,b \in C^\perp} (T_t)_{x+a,y+b} \|f(x+a) - g(y+b)\|_X^2 \end{aligned} \quad (38)$$

$$\begin{aligned}
&= \frac{1}{2^n} \sum_{x,y \in \mathbb{F}_2^n} (T_t)_{x,y} \|f(x) - g(y)\|_X^2 \\
&\leq \frac{3}{|E|} \sum_{(x,y) \in E} \|f(x) - g(y)\|_X^2 \\
&= \frac{3}{|E/C^\perp|} \sum_{(x,y) \in E/C^\perp} \|f(x) - g(y)\|_X^2.
\end{aligned} \tag{39}$$

The estimate (38) follows from Proposition 2.10, and (39) follows from (37), and is valid since  $18\tau^2 n \geq 4 \log n + \log 4$ .  $\square$

**2.5. Discussion.** In the full version of this extended abstract we will also establish the metric Markov cotype property (22) for other classes of metric spaces, including  $CAT(0)$  metric spaces, and in particular simply connected manifolds with non-positive sectional curvature (in which case we can take  $C = O(1)$  and  $\varepsilon = 1$ ). We will discuss this notion in detail and show that it implies a slight variant of Ball’s Markov cotype. This fact is new for  $CAT(0)$  metric spaces, and hence, in conjunction with Ball’s extension theorem [1], our work implies a new Lipschitz extension theorem for  $CAT(0)$  targets: any Lipschitz function from a subset  $U$  of a metric space  $(X, d_X)$  which has Markov type 2 and takes values in a  $CAT(0)$  metric space  $(Y, d_Y)$  can be extended to a Lipschitz function defined on all of  $X$  whose Lipschitz constant is larger by at most a constant factor (depending on  $X$ ). The definition of the notion of Markov type is beyond the scope of this extended abstract: it suffices to say that we can take  $X$  to be Hilbert space [1], or  $L_p$  for  $p \in [2, \infty)$  and even more generally a normed spaces whose modulus of smoothness has power type 2 [26], a tree, the word metric on a hyperbolic group or a simply connected manifold with pinched sectional curvature [26], series parallel graphs [5], or Alexandrov spaces (in particular manifolds) of non-negative curvature [29]. While we believe that this extension theorem is a key consequence of our work, for lack of space we defer the discussion about it to the full version of this paper.

Another major addition to the full version of this extended abstract will be a different proof of the decay of the Poincaré constant of powers of symmetric stochastic matrices in super-reflexive spaces and in  $CAT(0)$  spaces. This proof is based on using *norm bounds*. The norm bound  $\lambda(A, X, p)$  for a doubly stochastic  $n \times n$  matrix  $A$  in a normed space  $X$  is the smallest  $\lambda > 0$  satisfying (36). The quantity  $\lambda(A, \mathbb{R}, 2)$  is the second absolute eigenvalue of  $A$ , and Proposition 2.10 shows that it always controls  $\gamma_+(A, \|\cdot\|_X^p)$  from above. We show in the full version that when  $X$  is  $p$ -convex,  $\lambda(A, X, p)$  also controls  $\gamma_+(A, \|\cdot\|_X^p)$  from below, and this allows us to obtain a short proof of the decay of  $\gamma_+(A^t, \|\cdot\|_X^p)$  as a function of  $t$  when  $X$  is a  $p$ -convex space. This alternative approach has the advantage that it proves the decay of the Poincaré constant of the power of the matrix instead of its Cesàro average, and it is also shorter than the Markov cotype approach. However, in many respects it is less “robust” than the Markov cotype approach taken here, it yields worse bounds, and most importantly, it yields a Poincaré inequality which is insufficient for our construction of zigzag super-expanders. All of these issues will be discussed in the full version.

A major question left open is what are the geometric conditions which allows for “sufficient” decay (for the purposes of our zigzag iteration) of the Poincaré constant. We give an example in Section 3.3 for a metric space in which the decay is insufficient. However, there

are natural spaces like  $L_\infty$  and  $L_1$  for which we do not know the answer. Indeed, it may be the case that any normed space has a sufficient decay.

Another question is what are the geometric conditions on metric spaces for the existence of expanders with respect to those spaces. It is tempting to conjecture that for normed spaces, having such an expander is equivalent to having finite cotype (i.e.,  $\exists \varepsilon_0 > 0, n_0 \in \mathbb{N}$  such that  $\ell_\infty^{n_0}$  does not embed with distortion less than  $1 + \varepsilon_0$ ).

As was mentioned in Section 1.1, it is an open question whether every ‘‘classical’’ expander is also super-expander. In the full version we rule out the most obvious approach to prove such a result: coarsely embedding any expander family in any other expander family. We give an example of two expander families  $\mathcal{F}_1$  and  $\mathcal{F}_2$  such that  $\mathcal{F}_1$  does not coarsely embed in  $\mathcal{F}_2$ .

It seems to be a challenging problem whether in (33) we can take  $\delta = 1$ . Even in the case  $X = L_p, p \neq 2$ , this seems to be unknown. A classical result of Meyer [24] implies that for  $p \geq 2$  we can take  $\delta = \frac{1}{2}$ , but even this improved bound is insufficient for our purpose of constructing a ‘‘sufficiently good’’ base graph.

### 3. SOME DETAILS

#### 3.1. Markov cotype and the decay of $\gamma$ .

*Proof of Claim 2.4.* Write  $B = (b_{ij}) = \frac{1}{m} \sum_{t=0}^{m-1} A^t$ . We may assume that  $\gamma(B, d_X^2) > 12C^2$ , since otherwise we are done. This assumption implies that there exist  $x_1, \dots, x_n \in X$  such that:

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2 > \frac{12C^2}{n} \sum_{i=1}^n \sum_{j=1}^n b_{ij} d_X(x_i, x_j)^2. \quad (40)$$

Let  $y_1, \dots, y_n \in X$  be as in (22). Note that for all  $i, j \in \{1, \dots, n\}$  we have:

$$\begin{aligned} d_X(x_i, x_j)^2 &\leq [d_X(x_i, y_i) + d_X(y_i, y_j) + d_X(y_j, x_j)]^2 \\ &\leq 3d_X(x_i, y_i)^2 + 3d_X(y_i, y_j)^2 + 3d_X(y_j, x_j)^2. \end{aligned} \quad (41)$$

By averaging (41) we get the bound:

$$\begin{aligned} \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(y_i, y_j)^2 &\geq \frac{1}{3n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2 - \frac{2}{n} \sum_{i=1}^n d_X(x_i, y_i)^2 \\ &= \frac{1}{6n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2 + \frac{1}{6n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2 - \frac{2}{n} \sum_{i=1}^n d_X(x_i, y_i)^2 \\ &\stackrel{(40)}{>} \frac{1}{6n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2 + \frac{2C^2}{n} \sum_{i=1}^n \sum_{j=1}^n b_{ij} d_X(x_i, x_j)^2 - \frac{2}{n} \sum_{i=1}^n d_X(x_i, y_i)^2 \\ &\stackrel{(22)}{\geq} \frac{1}{6n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2. \end{aligned} \quad (42)$$

Using (22) once more, in conjunction with the definition of  $\gamma(A, d_X^2)$ , we see that:

$$\begin{aligned} C^2 \sum_{i=1}^n \sum_{j=1}^n b_{ij} d_X(x_i, x_j)^2 &\geq \frac{m^\varepsilon}{\gamma(A, d_X^2)} \cdot \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(y_i, y_j)^2 \\ &\stackrel{(42)}{>} \frac{m^\varepsilon}{6\gamma(A, d_X^2)} \cdot \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_X(x_i, x_j)^2. \end{aligned}$$

This concludes the proof of (23). □

*Proof of Claim 2.5.* Fix  $f, g : [n] \rightarrow X$ , and define  $h : [2n] \rightarrow X$ ,

$$h(x) = \begin{cases} f(x) & x \leq n \\ g(x - n) & x > n. \end{cases}$$

Then

$$\begin{aligned} \frac{1}{n^2} \sum_{x, y \in [n]} d_X(f(x), g(y))^2 &= \frac{1}{n^2} \sum_{x, y \in [n]} d_X(h(x), h(y + n))^2 \\ &\leq \frac{2}{(2n)^2} \sum_{x, y \in [2n]} d_X(h(x), h(y))^2 \\ &\leq \gamma\left(\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2\right) \frac{2}{2n} \sum_{x, y \in [n]} 2A_{xy} d_X(h(x), h(y + n))^2 \\ &= \frac{2\gamma\left(\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2\right)}{n} \sum_{x, y \in [n]} A_{xy} d_X(f(x), g(y))^2. \end{aligned}$$

We next prove the lower bound. Fix  $h : [2n] \rightarrow X$ , and define  $f, g : [n] \rightarrow X$ ,  $f(x) = h(x)$ , and  $g(x) = h(x + n)$ . Then,

$$\begin{aligned} \frac{1}{n^2} \sum_{x, y \in [n]} d_X(h(x), h(y))^2 &\leq \frac{1}{n^2} \sum_{x, y \in [n]} 2 \sum_{z \in [n]} (d_X(h(x), h(z + n))^2 + d_X(h(y), h(z + n))^2) \\ &= \frac{4}{n^2} \sum_{x, y \in [n]} d_X(f(x), g(y))^2. \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{1}{n^2} \sum_{x, y \in [n]} d_X(h(x + n), h(y + n))^2 \\ &\leq \frac{1}{n^2} \sum_{x, y \in [n]} 2 \sum_{z \in [n]} (d_X(h(x + n), h(z))^2 + d_X(h(y + n), h(z))^2) \\ &= \frac{4}{n^2} \sum_{x, y \in [n]} d_X(f(x), g(y))^2. \end{aligned}$$

Hence, by summing the above two inequalities,

$$\begin{aligned}
& \frac{1}{(2n)^2} \sum_{x,y \in [2n]} d_X(h(x), h(y))^2 \\
&= \frac{1}{(2n)^2} \sum_{x,y \in [n]} d_X(h(x), h(y))^2 + \frac{1}{(2n)^2} \sum_{x,y \in [n]} d_X(h(x+n), h(y+n))^2 \\
&\quad + \frac{1}{(2n)^2} \sum_{x,y \in [n]} d_X(h(x), h(y+n))^2 + \frac{1}{(2n)^2} \sum_{x,y \in [n]} d_X(h(x+n), h(y))^2 \\
&\leq \frac{1+1+0.25+0.25}{n^2} \sum_{x,y \in [n]} d_X(f(x), g(y))^2 \\
&\leq \frac{2.5\gamma_+(G, d_X^2)}{n} \sum_{x,y \in [n]} A_{xy} d_X(f(x), g(y))^2 \\
&= \frac{2.5\gamma_+(G, d_X^2)}{n} \sum_{x,y \in [n]} A_{xy} d_X(h(x), h(y+n))^2. \quad \square
\end{aligned}$$

*Proof of Claim 2.6.* We observe that for odd  $t$ ,  $\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}^t = \begin{pmatrix} 0 & A^t \\ A^t & 0 \end{pmatrix}$ . What complicates matters is that for even  $t$ ,  $\begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}^t = \begin{pmatrix} A^t & 0 \\ 0 & A^t \end{pmatrix}$ . However, we can write every even  $t > 0$  as a sum of two almost unique odd numbers by defining for even  $t$ ,  $o_1(t) \leq o_2(t)$  the unique pair of odd numbers such that  $o_2(t) - o_1(t) \leq 2$ , and  $o_1(t) + o_2(t) = t$ . Note that the multiset

$$\{o_1(t), o_2(t) : t \text{ is even, and } t \leq m\}$$

contains all the odd numbers in the range  $\{1, \dots, \lfloor m/2 \rfloor + 1\}$ , and at most four items for each value. Let  $B = \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}$ . We therefore can bound for every  $h : [2n] \rightarrow X$ ,

$$\begin{aligned}
& \frac{1}{2n} \sum_{x,y \in [2n]} (\mathcal{A}_m(B))_{xy} d_X(h(x), h(y))^2 = \frac{1}{2nm} \sum_{t=0}^{m-1} \sum_{x,y \in [2n]} (B^t)_{xy} d_X(h(x), h(y))^2 \\
&\leq \frac{1}{2nm} \sum_{\substack{t \in [m] \\ t \text{ is odd}}} \sum_{x,y \in [2n]} (B^t)_{xy} d_X(h(x), h(y))^2 + \frac{1}{2nm} \sum_{x \in [2n]} d_X(h(x), h(x))^2 \\
&\quad + \frac{1}{2nm} \sum_{\substack{t \in \{2, \dots, m-1\} \\ t \text{ is even}}} 2 \left( \sum_{x,y \in [2n]} (B^{o_1(t)})_{xy} d_X(h(x), h(y))^2 + \sum_{x,y \in [2n]} (B^{o_2(t)})_{xy} d_X(h(x), h(y))^2 \right) \\
&\leq \frac{9}{2nm} \sum_{\substack{t \in [m] \\ t \text{ is odd}}} \sum_{x,y \in [2n]} (B^t)_{xy} d_X(h(x), h(y))^2 \\
&= \frac{9}{2nm} \sum_{\substack{t \in [m] \\ t \text{ is odd}}} \sum_{x,y \in [2n]} \begin{pmatrix} 0 & A^t \\ A^t & 0 \end{pmatrix}_{xy} d_X(h(x), h(y))^2 \\
&\leq \frac{9}{2n} \sum_{x,y \in [2n]} \begin{pmatrix} 0 & \mathcal{A}_m(A) \\ \mathcal{A}_m(A) & 0 \end{pmatrix} d_X(h(x), h(y))^2.
\end{aligned}$$

Hence

$$\begin{aligned} \frac{1}{(2n)^2} \sum_{x,y \in [2n]} d_X(h(x), h(y))^2 &\leq \frac{\gamma(\mathcal{A}_m \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2)}{2n} \cdot \sum_{x,y \in [2n]} (\mathcal{A}_m \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix})_{xy} d_X(h(x), h(y))^2 \\ &\leq \frac{9\gamma(\mathcal{A}_m \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2)}{2n} \cdot \sum_{x,y \in [2n]} (\mathcal{A}_m \begin{pmatrix} 0 & \mathcal{A}_m(A) \\ \mathcal{A}_m(A) & 0 \end{pmatrix}) d_X(h(x), h(y))^2, \end{aligned}$$

which means that

$$\gamma\left(\begin{pmatrix} 0 & \mathcal{A}_m(A) \\ \mathcal{A}_m(A) & 0 \end{pmatrix}, d_X^2\right) \leq 9\gamma\left(\mathcal{A}_m \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}, d_X^2\right). \quad \square$$

**3.2. Proof of (26).** Our methods here are inspired by Ball's proof [1] of the Markov cotype property. Let  $(X, \|\cdot\|_X)$  be a normed space. The *modulus of uniform convexity* of  $X$  is defined for  $\varepsilon \in [0, 2]$  as

$$\delta_X(\varepsilon) = \inf \left\{ 1 - \frac{\|x+y\|_X}{2} : x, y \in X, \|x\|_X = \|y\|_X = 1, \|x-y\|_X = \varepsilon \right\}. \quad (43)$$

A normed space  $X$  is said to be **uniformly convex** if  $\delta_X(\varepsilon) > 0$  for all  $\varepsilon \in (0, 2]$ . Furthermore,  $X$  is said to have modulus of convexity of power type  $p$  if there exists a constant  $c$  such that  $\delta(\varepsilon) \geq c\varepsilon^p$  for all  $\varepsilon \in [0, 2]$ . It is straightforward to check that in this case necessarily  $p \geq 2$ . By Proposition 7 in [2] (see also [6]),  $X$  has modulus of convexity of power type  $p$  if and only if there exists a constant  $K > 0$  such that for every  $x, y \in X$

$$2\|x\|_X^p + \frac{2}{K^p} \|y\|_X^p \leq \|x+y\|_X^p + \|x-y\|_X^p. \quad (44)$$

The least  $K$  for which (44) holds is called the  $p$ -convexity constant of  $X$ , and is denoted  $K_p(X)$ .

The following lemma is stated and proved in [1] when  $p = 2$ .

**Lemma 3.1.** *Let  $X$  be a normed space and  $U$  a random vector in  $X$  with  $\mathbb{E}\|U\|_X^p < \infty$ . Then*

$$\|\mathbb{E}U\|_X^p + \frac{\mathbb{E}\|U - \mathbb{E}U\|_X^p}{(2^{p-1} - 1)K_p(X)^p} \leq \mathbb{E}\|U\|_X^p. \quad (45)$$

*Proof.* We repeat here the  $p > 2$  variant of the argument from [1] for the sake of completeness. Define

$$\theta = \inf \left\{ \frac{\mathbb{E}\|V\|_X^p - \|\mathbb{E}V\|_X^p}{\mathbb{E}\|V - \mathbb{E}V\|_X^p} : V \in L_p(X) \wedge \mathbb{E}\|V - \mathbb{E}V\|_X^p > 0 \right\}. \quad (46)$$

Then  $\theta \geq 0$ . Our goal is to show that

$$\theta \geq \frac{1}{(2^{p-1} - 1)K_p(X)^p}. \quad (47)$$

Fix  $\phi > \theta$ . Then there exists a random vector  $V_0 \in L_p(X)$  for which

$$\phi \mathbb{E}\|V_0 - \mathbb{E}V_0\|_X^p > \mathbb{E}\|V_0\|_X^p - \|\mathbb{E}V_0\|_X^p. \quad (48)$$

Fix  $K > K_p(X)$ . Apply the inequality (44) point-wise to the vectors  $x = \frac{1}{2}V_0 + \frac{1}{2}\mathbb{E}V_0$  and  $y = \frac{1}{2}V_0 - \frac{1}{2}\mathbb{E}V_0$ , to get that

$$2 \left\| \frac{1}{2}V_0 + \frac{1}{2}\mathbb{E}V_0 \right\|_X^p + \frac{2}{K^p} \left\| \frac{1}{2}V_0 - \frac{1}{2}\mathbb{E}V_0 \right\|_X^p \leq \|V_0\|_X^p + \|\mathbb{E}V_0\|_X^p. \quad (49)$$

Hence

$$\begin{aligned}
\phi \mathbb{E} \|V_0 - \mathbb{E} V_0\|_X^p &\stackrel{(48)}{>} \mathbb{E} \|V_0\|_X^p - \|\mathbb{E} V_0\|_X^p \\
&\stackrel{(49)}{\geq} 2 \left( \mathbb{E} \left\| \frac{1}{2} V_0 + \frac{1}{2} \mathbb{E} V_0 \right\|_X^p - \left\| \mathbb{E} \left( \frac{1}{2} V_0 + \frac{1}{2} \mathbb{E} V_0 \right) \right\|_X^p \right) + \frac{2}{K^p} \mathbb{E} \left\| \frac{1}{2} V_0 - \frac{1}{2} \mathbb{E} V_0 \right\|_X^p \\
&\stackrel{(46)}{\geq} 2 \mathbb{E} \left\| \left( \frac{1}{2} V_0 + \frac{1}{2} \mathbb{E} V_0 \right) - \mathbb{E} \left( \frac{1}{2} V_0 + \frac{1}{2} \mathbb{E} V_0 \right) \right\|_X^p + \frac{2}{K^p} \mathbb{E} \left\| \frac{1}{2} V_0 - \frac{1}{2} \mathbb{E} V_0 \right\|_X^p \\
&= \left( \frac{\theta}{2^{p-1}} + \frac{1}{2^{p-1} K^p} \right) \mathbb{E} \|V_0 - \mathbb{E} V_0\|_X^p.
\end{aligned}$$

Thus

$$\phi \geq \frac{\theta}{2^{p-1}} + \frac{1}{2^{p-1} K^p}. \quad (50)$$

Since (50) holds for all  $\phi > \theta$  and  $K > K_p(X)$ , the desired lower bound (47) follows.  $\square$

Let  $X$  be a Banach space with  $K_p(X) < \infty$ . Assume that  $\{M_k\}_{k=0}^n \subseteq X$  is a martingale with respect to the filtration  $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_{n-1}$ ; that is,  $\mathbb{E}(M_{i+1} | \mathcal{F}_i) = M_i$  for every  $i \in \{0, 1, \dots, n-1\}$ . Lemma 3.1 implies that

$$\begin{aligned}
&\mathbb{E} [\|M_n - M_0\|_X^p | \mathcal{F}_{n-1}] \\
&\geq \|\mathbb{E} [M_n - M_0 | \mathcal{F}_{n-1}]\|_X^p + \frac{\mathbb{E} [\|M_n - M_0 - \mathbb{E} [M_n - M_0 | \mathcal{F}_{n-1}]\|_X^p | \mathcal{F}_{n-1}]}{(2^{p-1} - 1) K_p(X)^p} \\
&= \|M_{n-1} - M_0\|_X^p + \frac{\mathbb{E} [\|M_n - M_{n-1}\|_X^p | \mathcal{F}_{n-1}]}{(2^{p-1} - 1) K_p(X)^p}. \quad (51)
\end{aligned}$$

Taking expectation in (51) implies that

$$\mathbb{E} [\|M_n - M_0\|_X^p] \geq \mathbb{E} [\|M_{n-1} - M_0\|_X^p] + \frac{\mathbb{E} [\|M_n - M_{n-1}\|_X^p]}{(2^{p-1} - 1) K_p(X)^p}.$$

Iterating this argument we obtain the following famous inequality of Pisier [31].

**Theorem 3** (Pisier [31]). *Let  $X$  be a Banach space with  $K_p(X) < \infty$ . Assume that  $\{M_k\}_{k=0}^n \subseteq X$  is a martingale (with respect some filtration). Then*

$$\mathbb{E} [\|M_n - M_0\|_X^p] \geq \frac{\sum_{k=1}^n \mathbb{E} [\|M_k - M_{k-1}\|_X^p]}{(2^{p-1} - 1) K_p(X)^p}.$$

We shall also need the following variant of Pisier's inequality:

**Corollary 3.2.** *For every  $p, q \in (1, \infty)$  there is a constant  $c(p, q) \in (0, \infty)$  with the following properties. Let  $X$  be a normed space with  $K_p(X) < \infty$ . Then for every martingale  $\{M_k\}_{k=0}^n \subseteq X$ , if  $p \leq q$ , then*

$$\mathbb{E} [\|M_n - M_0\|_X^q] \geq \frac{c(p, q)}{K_p(X)^p} \sum_{k=1}^n \mathbb{E} [\|M_k - M_{k-1}\|_X^q]. \quad (52)$$

and if  $p \geq q$ , then

$$\mathbb{E} [\|M_n - M_0\|_X^q] \geq \frac{c(p, q)}{n^{1-\frac{q}{p}} K_p(X)^q} \sum_{k=1}^n \mathbb{E} [\|M_k - M_{k-1}\|_X^q]. \quad (53)$$



*Proof.* Here we only prove the case  $q \leq p$ , since it is the only one needed in the sequel. By a theorem of Figiel [6] (combined with [2, Proposition 7]) we have  $K_p(L_q(X)) \leq CK_p(X)$ , where  $C$  depends only on  $p, q$ . We may therefore apply (44) to the vectors

$$\begin{aligned} x &= M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2} \in L_q(X), \\ y &= \frac{M_n - M_{n-1}}{2} \in L_q(X) \end{aligned}$$

to get

$$\begin{aligned} 2 \left( \mathbb{E} \left[ \left\| M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2} \right\|_X^q \right] \right)^{p/q} &+ \frac{2}{2^p CK_p(X)^p} (\mathbb{E} [\|M_n - M_{n-1}\|_X^q])^{p/q} \\ &\leq (\mathbb{E} [\|M_n - M_0\|_X^q])^{p/q} + (\mathbb{E} [\|M_{n-1} - M_0\|_X^q])^{p/q}. \end{aligned} \quad (54)$$

Assume that  $\{M_k\}_{k=0}^n \subseteq X$  is a martingale with respect to the filtration  $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_{n-1}$ . Then

$$\mathbb{E} [\|M_{n-1} - M_0\|_X^q] = \mathbb{E} [\|M_{n-1} - M_0 + \mathbb{E}[M_n - M_{n-1} | \mathcal{F}_{n-1}]\|_X^q] \leq \mathbb{E} [\|M_n - M_0\|_X^q],$$

and

$$\begin{aligned} \mathbb{E} [\|M_{n-1} - M_0\|_X^q] &= \mathbb{E} \left[ \left\| M_{n-1} - M_0 + \mathbb{E} \left[ \frac{M_n - M_{n-1}}{2} \middle| \mathcal{F}_{n-1} \right] \right\|_X^q \right] \\ &\leq \mathbb{E} \left[ \left\| M_{n-1} - M_0 + \frac{M_n - M_{n-1}}{2} \right\|_X^q \right]. \end{aligned}$$

Thus (54) implies that

$$(\mathbb{E} [\|M_{n-1} - M_0\|_X^q])^{p/q} + \frac{(\mathbb{E} [\|M_n - M_{n-1}\|_X^q])^{p/q}}{2^p CK_p(X)^p} \leq (\mathbb{E} [\|M_n - M_0\|_X^q])^{p/q}. \quad (55)$$

Applying (55) inductively we get the lower bound

$$\begin{aligned} 2^p CK_p(X)^p (\mathbb{E} [\|M_n - M_0\|_X^q])^{p/q} &\geq \sum_{k=1}^n (\mathbb{E} [\|M_k - M_{k-1}\|_X^q])^{p/q} \\ &\geq \frac{1}{n^{q/p-1}} \left( \sum_{k=1}^n \mathbb{E} [\|M_k - M_{k-1}\|_X^q] \right)^{p/q}, \end{aligned}$$

which is precisely (53).  $\square$

Note that Corollary 3.2 specialized to  $q = 2$  is precisely (26).

**3.3. A metric space without uniform decay of the Poincaré constant.** Observe that as long as a metric space  $(X, d_X)$  contains at least two points, the fact that  $\gamma_+(A, d_X^p) < \infty$  implies that  $A$  is ergodic, and therefore  $\lim_{t \rightarrow \infty} \gamma_+(A^t, d_X^p) = \lim_{t \rightarrow \infty} \gamma_+(\mathcal{A}_t(A), d_X^p) = 1$ . However this weak decay is insufficient for our purposes. For the iterative construction in

this paper we need at the very least the following type of *uniform decay* of the Poincaré constant:

$\forall M > 1 \exists t \in \mathbb{N} \exists s_0 \in [1, \infty) \forall n \in \mathbb{N} \forall A \in M_n(\mathbb{R})$  symmetric & stochastic,

$$\gamma_+(A, d_X^p) \geq s_0 \implies \gamma_+(\mathcal{A}_t(A), d_X^p) \leq \frac{\gamma_+(A, d_X^p)}{M}.$$

Here we show that there exists a metric space with no such uniform decay.

**Proposition 3.3.** *There exists a metric space  $(X, \rho)$ , and graphs  $\{G_n = (V_n, E_n)\}_{n=1}^\infty$ , where  $|V_n| = n$ , such that  $\lim_{n \rightarrow \infty} \gamma_+(G_n, \rho^2) = \infty$ , and for every  $t \in \mathbb{N}$ , there exists  $N_0$  such that for every  $n > N_0$ ,  $\gamma_+(\mathcal{A}_t(G_n), \rho^2) \geq \frac{1}{2} \gamma_+(G_n, \rho)$ .*

*Proof.* Let  $G_n = (V_n, E_n)$  be an arbitrary family of constant degree expanders, i.e.,  $G_n$  is a degree  $d$  (say  $d = 4$ ),  $n$ -vertex graph satisfying  $\gamma_+(G_n, \|\cdot\|_2^2) \leq C$ , for some  $C > 0$ . Let  $(X, \rho)$  be the metric space defined as follows: as a set,  $X = \ell_\infty \cap \mathbb{Z}^{\mathbb{N}_0}$ , i.e., the set of all finite integer-valued sequences. The metric  $\rho$  on  $X$  is given by  $\rho(x, y) = \log(1 + \|x - y\|_\infty)$ . Note that  $\rho$  is indeed a metric since  $T(s) = \log(1 + s)$  is concave, increasing and  $T(0) = 0$  (i.e.,  $\rho = T \circ \|\cdot\|_\infty$  is a metric transform of the  $\ell_\infty$  norm).

First we prove that  $\gamma_+(G_n, \rho^2) \lesssim (\log(1 + \log n))^2$ . Let  $f, g : G_n \rightarrow X$  be arbitrary mappings. Our goal is to prove that

$$\frac{1}{n^2} \sum_{u, v \in V} \rho(f(u), g(v))^2 \lesssim \frac{(\log(1 + \log n))^2}{nd} \sum_{\{u, v\} \in E} \rho(f(u), f(v))^2.$$

Write  $S = f(V) \cup g(V) \subseteq \mathbb{Z}^\infty$ . Let  $B : (S, \|\cdot\|_\infty) \rightarrow \ell_2$  be the Bourgain embedding [4] of  $S$  equipped the  $\ell_\infty$  norm into  $\ell_2$ , i.e., for every  $u, v \in V$ ,

$$\|f(u) - g(v)\|_\infty \leq \|B(f(u)) - B(g(v))\|_2 \leq c(1 + \log n) \|f(u) - g(v)\|_\infty. \quad (56)$$

for some constant  $c \geq 1$ .

We next apply the metric transform  $T(x) = \log(1 + x)$  to (56), and obtain for  $f(u) \neq g(v)$ ,

$$\begin{aligned} \rho(f(u), g(v)) &= T(\|f(u) - g(v)\|_\infty) \\ &\leq T(\|B(f(u)) - B(g(v))\|_2) \\ &\leq \log(1 + c(1 + \log n) \|f(u) - g(v)\|_\infty) \\ &\leq \log(c(1 + \log n)) + \log(2 \|f(u) - g(v)\|_\infty) \\ &\lesssim \log(1 + \log n) \cdot \rho(f(u), g(v)). \end{aligned} \quad (57)$$

The second and last inequalities above follow from the fact that  $\|f(u) - g(v)\|_\infty \geq 1$ .

As was shown in [22, Remark 5.4], there exists a universal constant  $c' > 1$  and a mapping  $\phi : \ell_2 \rightarrow \ell_2$  such that for all  $x, y \in \ell_2$  we have  $T(\|x - y\|_2) \leq \|\phi(x) - \phi(y)\|_2 \leq c' T(\|x - y\|_2)$ . This fact, together with (57), implies that the mapping  $\psi = \phi \circ B : S \rightarrow \ell_2$  satisfies:

$$\rho(f(u), g(v)) \leq \|\psi(f(u)) - \psi(g(v))\|_2 \lesssim \log(1 + \log n) \cdot \rho(f(u), g(v)),$$

for every  $u, v \in V$ . Applying the Poincaré inequality of  $G$  in  $\ell_2$  we conclude that

$$\frac{1}{n^2} \sum_{u, v \in V} \rho(f(u), g(v))^2 \leq \frac{1}{n^2} \sum_{u, v \in V} \|\psi(f(u)) - \psi(g(v))\|_2^2$$

$$\lesssim \frac{1}{nd} \sum_{(u,v) \in E} \|\psi(f(u)) - \psi(g(v))\|_2^2 \lesssim \frac{(\log(1 + \log n))^2}{nd} \sum_{(u,v) \in E} \rho(f(u), g(v))^2.$$

This proves that  $\gamma_+(G_n, \rho^2) \lesssim (\log(1 + \log n))^2$ .

Next, we show a lower bound on  $\gamma_+(\mathcal{A}_t(G), \rho)$ . For this purpose it is sufficient to examine a specific embedding of  $\mathcal{A}_t(G)$  in  $X$ . Let  $g = f : V \rightarrow \mathbb{Z}^{\mathbb{N}_0}$  which is an isometric embedding of the shortest path metric on  $\mathcal{A}_t(G)$  into  $(\mathbb{Z}^{\mathbb{N}_0}, \|\cdot\|_\infty)$ . In this case for every  $\{u, v\} \in E(\mathcal{A}_t(G))$ ,  $\rho(f(u), f(v)) = T(\|f(u) - f(v)\|_\infty) = T(1) = 1$ . On the other hand, since the degree of  $\mathcal{A}_t(G)$  is  $td^t$ , at least half of the pairs in  $V \times V$  are at distance  $\gtrsim \frac{\log n}{t \log d}$  in the shortest path metric on  $\mathcal{A}_t(G)$ , and hence their images under  $f$  are at  $\rho$ -distance  $\geq \log\left(1 + c'' \frac{\log n}{t \log d}\right)$ , where  $c'' > 0$  is a universal constant. Hence, provided  $t \leq c''' \log n$  for some sufficiently small constant  $c'''$ , we have  $\gamma_+(\mathcal{A}_t(G), \rho^2) \gtrsim \left(\log\left(1 + \frac{\log n}{t}\right)\right)^2$ .  $\square$

### 3.4. Construction of the base graph.

*Proof of Claim 2.9.* The claim easily follows by computing the right hand side in (35) when  $f = W_A$ . We write  $x = x_A x_{A^c}$ , and

$$\begin{aligned} & \sum_{y \in \mathbb{F}_2^n} \left(\frac{1 - e^{-t}}{2}\right)^{\|x-y\|_1} \left(\frac{1 + e^{-t}}{2}\right)^{n - \|x-y\|_1} \prod_{i \in A} (-1)^{y_i} \\ &= \sum_{y_A \in \mathbb{F}_2^A} \left(\frac{1 - e^{-t}}{2}\right)^{\|x_A - y_A\|_1} \left(\frac{1 + e^{-t}}{2}\right)^{|A| - \|x_A - y_A\|_1} \prod_{i \in A} (-1)^{y_i} \\ & \quad \cdot \sum_{y_{A^c} \in \mathbb{F}_2^{[n] \setminus A}} \left(\frac{1 - e^{-t}}{2}\right)^{\|x_{A^c} - y_{A^c}\|_1} \left(\frac{1 + e^{-t}}{2}\right)^{n - |A| - \|x_{A^c} - y_{A^c}\|_1} \\ &= \sum_{\ell=0}^{|A|} \binom{|A|}{\ell} \left(\frac{1 - e^{-t}}{2}\right)^\ell \left(\frac{1 + e^{-t}}{2}\right)^{|A| - \ell} (-1)^\ell W_A(x) \\ &= e^{-t|A|} W_A(x). \end{aligned} \quad \square$$

*Proof of Proposition 2.10.* Let  $f, g : [N] \rightarrow X$ , denote by  $\bar{f} = \mathbb{E} f$ , and  $\bar{g} = \mathbb{E} g$ ,  $\tilde{f} = f - \bar{f}$ ,  $\tilde{g} = g - \bar{g}$ . Then  $\sum_{i=1}^N \tilde{f}(i) = \sum_{i=1}^N \tilde{g}(i) = 0$ , and so by the definition of  $\lambda$ ,  $\|A\tilde{f}\|_{L_p(X)} \leq \lambda \|\tilde{f}\|_{L_p(X)}$ , and  $\|A\tilde{g}\|_{L_p(X)} \leq \lambda \|\tilde{g}\|_{L_p(X)}$ . Hence

$$\begin{aligned} \left\| \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix} (\tilde{f} \oplus \tilde{g}) \right\|_{L_p(X)}^p &= \frac{\|A\tilde{g}\|_{L_p(X)}^p + \|A\tilde{f}\|_{L_p(X)}^p}{2} \\ &\leq \lambda^p \frac{\|\tilde{f}\|_{L_p(X)}^p + \|\tilde{g}\|_{L_p(X)}^p}{2} = \lambda^p \|\tilde{f} \oplus \tilde{g}\|_{L_p(X)}^p. \end{aligned}$$

It follows that,

$$(1 - \lambda) \|\tilde{f} \oplus \tilde{g}\|_{L_p(X)} \leq \|\tilde{f} \oplus \tilde{g}\|_{L_p(X)} - \left\| \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix} (\tilde{f} \oplus \tilde{g}) \right\|_{L_p(X)} \leq \left\| \begin{pmatrix} I & -A \\ -A & I \end{pmatrix} (\tilde{f} \oplus \tilde{g}) \right\|_{L_p(X)}.$$

So,

$$\begin{aligned}
\frac{1}{N^2} \sum_{i,j \in [N]} \|f(i) - g(j)\|_X^p &\leq 2^{p-1} \|\bar{f} - \bar{g}\|_X^p + \frac{2^{p-1}}{N^2} \sum_{i,j \in [N]} \|\tilde{f}(i) - \tilde{g}(j)\|_X^p \\
&\leq 2^{p-1} \|\bar{f} - \bar{g}\|_X^p + 4^{p-1} (\|\tilde{f}\|_{L_p(X)}^p + \|\tilde{g}\|_{L_p(X)}^p) \\
&= 2^{p-1} \|\bar{f} - \bar{g}\|_X^p + 2^{2p-1} \|\tilde{f} \oplus \tilde{g}\|_{L_p(X)}^p \\
&\leq 2^{p-1} \|\bar{f} - \bar{g}\|_X^p + 2^{2p-1} (1 - \lambda)^{-p} \left\| \begin{pmatrix} I & -A \\ -A & I \end{pmatrix} (\tilde{f} \oplus \tilde{g}) \right\|_{L_p(X)}^p.
\end{aligned}$$

Expanding the last term,

$$\begin{aligned}
&\left\| \begin{pmatrix} I & -A \\ -A & I \end{pmatrix} (\tilde{f} \oplus \tilde{g}) \right\|_{L_p(X)}^p \\
&= \frac{1}{2N} \sum_{i \in [N]} \left\| \sum_{j \in [N]} A_{ij} (\tilde{f}(i) - \tilde{g}(j)) \right\|_X^p + \frac{1}{2N} \sum_{i \in [N]} \left\| \sum_{j \in [N]} A_{ij} (\tilde{g}(i) - \tilde{f}(j)) \right\|_X^p \\
&\leq \frac{1}{N} \sum_{i,j \in [N]} A_{ij} \|\tilde{f}(i) - \tilde{g}(j)\|_X^p \\
&\leq 2^{p-1} \|\bar{f} - \bar{g}\|_X^p + \frac{2^{p-1}}{N} \sum_{i,j \in [N]} A_{ij} \|f(i) - g(j)\|_X^p.
\end{aligned}$$

We next observe that for any permutation  $\pi : [N] \rightarrow [N]$  (and  $p \geq 1$ ),  $\|\bar{f} - \bar{g}\|_X^p \leq \frac{1}{N} \sum_i \|f(i) - g(\pi(i))\|_X^p$ . Since  $A$  is symmetric stochastic, it is in the convex hull of the permutation matrices, so we can write  $A = \sum_\ell \alpha_\ell \Pi_\ell$ , where  $\{\Pi_\ell\}_\ell$  are permutation matrices, and  $\alpha_\ell \geq 0$ ,  $\sum_\ell \alpha_\ell = 1$ . Hence,

$$\|\bar{f} - \bar{g}\|_X^p \leq \sum_\ell \alpha_\ell \frac{1}{N} \sum_{i,j} \Pi_{ij}^\ell \|f(i) - g(j)\|_X^p = \frac{1}{N} \sum_{i,j} A_{ij} \|f(i) - g(j)\|_X^p.$$

We conclude that,

$$\frac{1}{N^2} \sum_{i,j \in [N]} \|f(i) - g(j)\|_X^p \leq \frac{8^p (1 - \lambda)^{-p}}{N} \sum_{i,j \in [N]} A_{ij} \|f(i) - g(j)\|_X^p. \quad \square$$

3.4.1. *Proof of Lemma 2.8.* The Rademacher projection of  $f$  is defined by

$$R_1 f = \sum_{j=1}^n \hat{f}(A) W_{\{j\}}.$$

The  $K$ -convexity constant of  $X$ , denoted  $K(X)$ , is the smallest constant  $K$  such that for every  $n$  and every  $f : \mathbb{F}_2^n \rightarrow X$ ,

$$\|(R_1 f)\|_{L_2(X)}^2 \leq K^2 \|f\|_{L_2(X)}^2.$$

We also define the  $k$ th level Rademacher projection by

$$R_k f = \sum_{\substack{A \subseteq [n] \\ |A|=k}} \hat{f}(A) W_A.$$

We assume throughout that  $X$  is a complex Banach space. The case of Banach spaces over  $\mathbb{R}$  follows formally from this case by complexification. Extend the heat semi-group  $T_z : L_p(X) \rightarrow L_p(X)$  to  $z \in \mathbb{C}$ , by setting:

$$T_z f = \sum_{A \subseteq [n]} e^{-z|A|} \hat{f}(A) W_A.$$

Another way to write it is  $T_z = \sum_{k=0}^n e^{-zk} R_k$ . We shall use the following deep theorem of Pisier[33]:

**Theorem 4** ([33]). *For  $K$ -convex  $X$ , and  $p \in (1, \infty)$ , there exists  $\phi, M >$  such that if  $|\arg z| < \phi$  then  $\|T_z\|_{L_p(X) \rightarrow L_p(X)} \leq M$ .*

See [21] for bounds on  $\phi, M$  in terms of  $K(X)$ . We will require the following standard corollary of Theorem 4:

**Corollary 3.4.**  $\|R_k\|_{L_p(X) \rightarrow L_p(X)} \leq M e^{ak}$ .

*Proof.* Define  $a = \frac{\pi}{\tan \phi}$ , so that all the points in the segment  $\{(a, iy) : y \in (-\pi, \pi)\}$  have argument at most  $\phi$ . Calculate,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} T_{a+it} dt &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} \sum_{r=0}^n e^{-(a+it)r} R_r dt \\ &= \frac{1}{2\pi} \sum_{r=0}^n e^{-ra} R_r \int_{-\pi}^{\pi} e^{i(k-r)t} dt = e^{-ka} R_k, \end{aligned}$$

and therefore, by convexity,

$$\|R_k\|_{L_p(X) \rightarrow L_p(X)} \leq \frac{e^{ka}}{2\pi} \int_{-\pi}^{\pi} |e^{ikt}| \cdot \|T_{a+it}\|_{L_p(X) \rightarrow L_p(X)} dt \leq M e^{ka}. \quad \square$$

**Corollary 3.5.** *Assume that  $\Re z \geq 2a$ . Then*

$$\begin{aligned} \|T_z\|_{L_p^{\geq m}(X) \rightarrow L_p^{\geq m}(X)} &= \left\| \sum_{k \geq m} e^{-zk} R_k \right\|_{L_p^{\geq m}(X) \rightarrow L_p^{\geq m}(X)} \\ &\leq \sum_{k \geq m} e^{-2ak} \cdot M e^{ak} = \frac{M}{1 - e^{-a}} e^{-ma}. \quad \square \end{aligned}$$

The ensuing argument is a quantitative version of the proof of the main theorem of Pisier in [34]. Let  $r = 2\sqrt{a^2 + \pi^2}$ , and define  $V = \{z \in \mathbb{C} : |z| \leq r \wedge |\arg z| \leq \phi\}$  (the set  $V \subseteq \mathbb{C}$  is depicted in Figure 2). Denote  $V_0 = \{x \pm ix \tan \phi : x \in [0, 2a)\}$  and  $V_1 = \{r e^{i\theta} : |\theta| \leq \phi\}$ , so that we have the disjoint union  $\partial V = V_0 \cup V_1$ .

Fix  $t \in (0, 2a)$ . We shall use the Harmonic measure corresponding to  $V$ , i.e., the Borel probability measure  $\mu = \mu_t$  on  $\partial V$  such that for every bounded analytic function on  $V$ ,

$$f(t) = \int_{\partial V} f(z) d\mu(z). \quad (58)$$

We refer to [7] for more information on this topic and the ensuing discussion. It suffices to say here that for a Borel set  $U \subseteq \partial V$ ,  $\mu_t(U)$  is the probability that Brownian motion

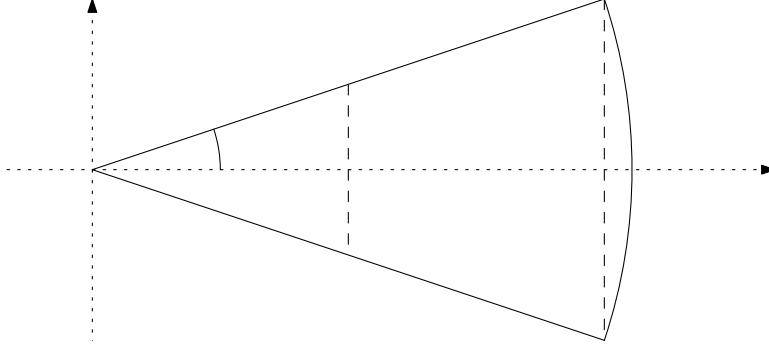


FIGURE 2. The sector  $V$ .

starting at  $t$  leaves  $V$  at  $U$ . Equivalently, by conformal invariance,  $\mu_t$  is the push-forward of the normalized Lebesgue measure on the unit circle  $S^1$  under the Riemann mapping from the unit disk to  $V$  which takes the origin to  $t$ .

Let  $\theta = \mu(V_1)$ , and denote  $\mu = (1 - \theta)\mu_0 + \theta\mu_1$ , where  $\mu_i$  is a probability measure on  $V_i$ ,  $i \in \{0, 1\}$ . We will require the following standard bound on  $\theta$ :

**Lemma 3.6.** *We have  $\theta \asymp \left(\frac{t}{r}\right)^{\frac{\pi}{2\phi}}$ .*

*Proof.* The proof is a simple exercise in conformal invariance. Let  $\mathbb{D}$  denote the unit disk centered at the origin, and let  $\mathbb{D}_+$  denote the intersection of  $\mathbb{D}$  with the right half plane  $\{z \in \mathbb{C} : \Re z \geq 0\}$ . The mapping  $z \mapsto \left(\frac{z}{r}\right)^{\frac{\pi}{2\phi}}$  is a conformal equivalence between  $V$  and  $\mathbb{D}_+$ . The mapping  $z \mapsto -i\frac{z+i}{z-i}$  is a conformal equivalence between  $\mathbb{D}_+$  and the positive quadrant  $\{x+iy : x, y \geq 0\}$ , and the mapping  $z \mapsto z^2$  is a conformal equivalence between the positive quadrant and the upper half-plane. Finally, the mapping  $z \mapsto \frac{z-i}{z+i}$  is a conformal equivalence between the upper half-plane and  $\mathbb{D}$ . By composing these mappings, we obtain the following conformal equivalence between  $V$  and  $\mathbb{D}$ :

$$F(z) = \frac{-\left(\left(\frac{z}{r}\right)^{\frac{\pi}{2\phi}} + i\right)^2 - i\left(\left(\frac{z}{r}\right)^{\frac{\pi}{2\phi}} - i\right)^2}{-\left(\left(\frac{z}{r}\right)^{\frac{\pi}{2\phi}} + i\right)^2 + i\left(\left(\frac{z}{r}\right)^{\frac{\pi}{2\phi}} - i\right)^2}.$$

Thus, by composing with the conformal automorphism of  $\mathbb{D}$ ,  $z \mapsto \frac{z-F(t)}{1-\overline{F(t)}z}$ , we obtain the mapping  $G(z) = \frac{F(z)-F(t)}{1-\overline{F(t)}F(z)}$ , which is a conformal equivalence between  $V$  and  $\mathbb{D}$  with  $G(t) = 0$ .

By conformal invariance,  $\theta$  is the length of the arc  $G(V_1) \subseteq \partial\mathbb{D} = S^1$ , divided by  $2\pi$ . The assertion of Lemma 3.6 now follows from a direct computation.  $\square$

**Lemma 3.7.** *For every  $\varepsilon \in (0, 1)$  there exists a bounded analytic function  $\Psi : V \rightarrow \mathbb{C}$ , such that  $\Psi(t) = 1$ , for every  $z \in V_0$ ,  $|\Psi(z)| \leq \varepsilon$ , and for every  $z \in V_1$ ,  $|\Psi(z)| \leq \varepsilon^{-\theta^{-1}+1}$ .*

*Proof.* The proof is the same as the proof of Claim 2 in [34]. Consider the strip  $S = \{z \in \mathbb{C} : \Re(z) \in [0, 1]\}$  and for  $j \in \{0, 1\}$  let  $S_j = \{z \in \mathbb{C} : \Re(z) = j\}$ . The argument in Claim 1 in [34] shows that there exists a conformal equivalence  $h : V \rightarrow S$  such that  $h(t) = \theta$ ,  $h(V_0) = S_0$ ,  $h(V_1) = S_1$ . Now define  $\Psi(z) = \varepsilon^{1-\theta^{-1}h(z)}$ .  $\square$

We next calculate

$$T_t = \Psi(t)T_t = \int_{\partial V} \Psi(z)T_z d\mu(z) = (1 - \theta) \int_{V_0} \Psi(z)T_z d\mu_0(z) + \theta \int_{V_1} \Psi(z)T_z d\mu_1(z).$$

Hence, using Lemma 3.7 with  $\varepsilon = e^{-am\theta}$ , we get the bound:

$$\|T_t\|_{L_p^{\geq m}(X) \rightarrow L_p^{\geq m}(X)} \leq (1 - \theta)\varepsilon M + \theta\varepsilon^{1-\theta^{-1}} C e^{-am} \leq (M + C)e^{-am\theta}. \quad (59)$$

Since by Lemma 3.6 we have  $\theta \asymp_{\phi} t^{\frac{\pi}{2\phi}}$  (where the implied constant depends on  $\phi$ ), the bound (59), which holds for  $t \in (0, 2a)$ , combined with Corollary 3.5, concludes the proof of Lemma 2.8.  $\square$

## REFERENCES

- [1] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172, 1992.
- [2] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Invent. Math.*, 115(3):463–482, 1994.
- [3] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric ramsey-type phenomena. *Annals of Mathematics*, 162(2):643–709, 2005.
- [4] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.
- [5] B. Brinkman, A. Karagiozova, and J. R. Lee. Vertex cuts, random walks, and dimension reduction in series-parallel graphs. In *STOC’07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 621–630. ACM, New York, 2007.
- [6] T. Figiel. On the moduli of convexity and smoothness. *Studia Math.*, 56:121–155, 1976.
- [7] J. B. Garnett and D. E. Marshall. *Harmonic measure*, volume 2 of *New Mathematical Monographs*. Cambridge University Press, Cambridge, 2008. Reprint of the 2005 original.
- [8] M. Gromov. Filling Riemannian manifolds. *J. Differential Geom.*, 18(1):1–147, 1983.
- [9] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.
- [10] M. Gromov. Random walk in random groups. *Geom. Funct. Anal.*, 13(1):73–146, 2003.
- [11] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the AMS*, 43(4):439–561, 2006.
- [12] R. C. James. A nonreflexive Banach space that is uniformly nonoctahedral. *Israel J. Math.*, 18:145–155, 1974.
- [13] R. C. James. Nonreflexive spaces of type 2. *Israel J. Math.*, 30(1-2):1–13, 1978.
- [14] R. C. James and J. Lindenstrauss. The octahedral problem for Banach spaces. In *Proceedings of the Seminar on Random Series, Convex Sets and Geometry of Banach Spaces (Mat. Inst., Aarhus Univ., Aarhus, 1974; dedicated to the memory of E. Asplund)*, pages 100–120. Various Publ. Ser., No. 24, Aarhus Univ., Aarhus, 1975. Mat. Inst.
- [15] G. Kasparov and G. Yu. The coarse geometric Novikov conjecture and uniform convexity. *Adv. Math.*, 206(1):1–56, 2006.
- [16] S. Khot and A. Naor. Nonembeddability theorems via Fourier analysis. *Mathematische Annalen*, 334(4):821–852, 2006.
- [17] V. Lafforgue. Un renforcement de la propriété (T). *Duke Math. J.*, 143(3):559–602, 2008.
- [18] V. Lafforgue. Propriété (T) renforcée Banachique et transformation de Fourier rapide. Preprint, 2009.
- [19] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.
- [20] J. Matoušek. On embedding expanders into  $\ell_p$  spaces. *Israel J. Math.*, 102:189–197, 1997.
- [21] B. Maurey. Type, cotype and  $K$ -convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1299–1332. North-Holland, Amsterdam, 2003.
- [22] M. Mendel and A. Naor. Euclidean quotients of finite metric spaces. *Adv. Math.*, 189(2):451–494, 2004.

- [23] M. Mendel and A. Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008. Preliminary version in SODA '06.
- [24] P.-A. Meyer. Transformations de Riesz pour les lois gaussiennes. In *Seminar on probability, XVIII*, volume 1059 of *Lecture Notes in Math.*, pages 179–193. Springer, Berlin, 1984.
- [25] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.
- [26] A. Naor, Y. Peres, O. Schramm, and S. Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197, 2006.
- [27] A. Naor and Y. Rabani. Spectral inequalities on curved spaces. Preprint, 2005.
- [28] A. Naor and L. Silberman. Poincaré inequalities, embeddings, and wild groups. Preprint, 2004.
- [29] S. Ohta. Markov type of Alexandrov spaces of nonnegative curvature. To appear in *Mathematika*, 2008.
- [30] N. Ozawa. A note on non-amenability of  $\mathcal{B}(l_p)$  for  $p = 1, 2$ . *Internat. J. Math.*, 15(6):557–565, 2004.
- [31] G. Pisier. Martingales with values in uniformly convex spaces. *Israel J. Math.*, 20(3-4):326–350, 1975.
- [32] G. Pisier. Some applications of the complex interpolation method to Banach lattices. *J. Analyse Math.*, 35:264–281, 1979.
- [33] G. Pisier. Holomorphic semigroups and the geometry of Banach spaces. *Ann. of Math. (2)*, 115(2):375–392, 1982.
- [34] G. Pisier. A remark on hypercontractive semigroups and operator ideals. Preprint, 2007.
- [35] G. Pisier. Complex interpolation between Hilbert, Banach and operator spaces. Preprint, 2008.
- [36] G. Pisier and Q. H. Xu. Random series in the real interpolation spaces between the spaces  $v_p$ . In *Geometrical aspects of functional analysis (1985/86)*, volume 1267 of *Lecture Notes in Math.*, pages 185–209. Springer, Berlin, 1987.
- [37] O. Reingold, L. Trevisan, and S. P. Vadhan. Pseudorandom walks on regular digraphs and the RL vs. L problem. In J. M. Kleinberg, editor, *STOC*, pages 457–466. ACM, 2006.
- [38] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Ann. of Math.*, 155(1):157–187, 2002.
- [39] J. Roe. *Lectures on coarse geometry*, volume 31 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2003.
- [40] E. Rozenman and S. Vadhan. Derandomized squaring of graphs. In *Approximation, randomization and combinatorial optimization*, volume 3624 of *Lecture Notes in Comput. Sci.*, pages 436–447. Springer, Berlin, 2005. Available from: <http://eccc.hpi-web.de/eccc-reports/2005/TR05-092>.

COMPUTER SCIENCE DIVISION, THE OPEN UNIVERSITY OF ISRAEL  
*E-mail address:* mendelma@gmail.com

COURANT INSTITUTE, NEW YORK UNIVERSITY  
*E-mail address:* naor@cims.nyu.edu