EXPANDERS WITH RESPECT TO HADAMARD SPACES AND RANDOM GRAPHS

MANOR MENDEL AND ASSAF NAOR

Abstract. It is shown that there exists a sequence of 3-regular graphs $\{G_n\}_{n=1}^{\infty}$ and a Hadamard space X such that $\{G_n\}_{n=1}^{\infty}$ forms an expander sequence with respect to X, yet random regular graphs are not expanders with respect to X. This answers a question of [NS11]. $\{G_n\}_{n=1}^{\infty}$ are also shown to be expanders with respect to random regular graphs, yielding a deterministic sublinear time constant factor approximation algorithm for computing the average squared distance in subsets of a random graph. The proof uses the Euclidean cone over a random graph, an auxiliary continuous geometric object that allows for the implementation of martingale methods.

Contents

1. Introduction

Throughout this paper all graphs are unweighted, non-oriented, and finite, and they are allowed to have parallel edges and self-loops, unless stated otherwise. Given a graph G, we denote its vertices by V_G and its edges by E_G. If G is connected then we denote the shortest-path metric that it induces on V_G by d_G.

Fix $d \in \mathbb{N}$ and let $\{G_n\}_{n=1}^{\infty}$ be a sequence of d-regular graphs such that $\lim_{n\to\infty} |V_{G_n}| = \infty$. Then $\{G_n\}_{n=1}^{\infty}$ is an expander sequence (see e.g. [HLW06]) if and only if for every sequence of Hilbert space valued functions $\{f_n : V_n \to \ell_2\}_{n=1}^{\infty}$ we have

$$\frac{1}{|V_{G_n}|^2} \sum_{(x,y)\in V_{G_n}\times V_{G_n}} \|f_n(x) - f_n(y)\|_2^2 \asymp \frac{1}{|E_{G_n}|} \sum_{\{u,v\}\in E_{G_n}} \|f_n(u) - f_n(v)\|_2^2. \quad (1)$$

Here and in what follows, when we write $A \asymp B$ we mean that there exist two universal constants $c, C \in (0,\infty)$ such that $cA \leqslant B \leqslant CB$. Also, in what follows the notations $A \lesssim B$ and $B \gtrsim A$ mean that $A \leqslant KB$ for some universal constant $K \in (0,\infty)$. If we need to allow K to depend on parameters, we indicate this by subscripts, thus e.g. $A \lesssim_{\alpha,\beta} B$ means that $A \leqslant K(\alpha,\beta)B$ for some $K(\alpha,\beta) \in (0,\infty)$ which is allowed to depend only on the parameters α and β.

The asymptotic identity (1) says that whenever one assigns a vector to each vertex of G_n, the average squared distance between these vectors can be estimated up to universal constant factors by averaging those squared distances that correspond to edges of G_n, an average of only $d|V_{G_n}|$ numbers rather than the full $|V_{G_n}|^2$ pairwise distances. This geometric characterization of expanders as "universal average Euclidean distance approximators" (following terminology of [BGS07]) is easy to prove (its proof will also be explained in the ensuing discussion).

It is natural to investigate the possible validity of (1) when the Hilbertian metric is replaced other metrics. (1) comprises of two asymptotic inequalities, one of which holds true in any metric space: if G is a d-regular graph and (X, d_X) is a metric space then for every $f : V_G \to X$ we have

$$\frac{1}{|E_G|} \sum_{\{u,v\}\in E_G} d_X\big(f(u), f(v)\big)^2 \leqslant \frac{4}{|V_G|^2} \sum_{(x,y)\in V_G\times V_G} d_X\big(f(x), f(y)\big)^2. \quad (2)$$

Indeed, by the triangle inequality and the convexity of $t \mapsto t^2$,

$$d_X(f(u), f(v))^2 \leqslant 2d_X(f(u), f(w))^2 + 2d_X(f(w), f(v))^2 \qquad (3)$$

for every $u, v, w \in V_G$. Since G is d-regular, the bound (2) follows by averaging (3) over the set $\{(u, v, w) \in V_G \times V_G \times V_G : \ \{u, v\} \in E_G\}$.

The nontrivial content of (1) is therefore the fact that the left hand side of (1) can be bounded by a multiple (independent of n and f) of the right hand side of (1). Thus, given a regular graph G and a metric space (X, d_X), let $\gamma(G, d_X^2)$ be the infimum over those $\gamma \in (0, \infty]$ such that for every $f : V_G \to X$,

$$\frac{1}{|V_G|^2} \sum_{(u,v)\in V_G\times V_G} d_X\big(f(u), f(v)\big)^2 \leqslant \frac{\gamma}{|E_G|} \sum_{\{u,v\}\in E_G} d_X\big(f(u), f(v)\big)^2. \quad (4)$$

If $X = \mathbb{R}$ with $d_{\mathbb{R}}(s, t) \stackrel{\mathrm{def}}{=} |s-t|$ then by expanding the squares in (4) one checks that

$$\gamma(G, d_{\mathbb{R}}^2) = \frac{1}{1 - \lambda_2(G)},$$

where $\lambda_2(G)$ denotes the second largest eigenvalue of the normalized adjacency matrix of the graph G. Despite the fact that there is no actual spectrum present in this geometric context, we think of $\gamma(G, d_X^2)$ as the reciprocal of the spectral gap of G with respect to (X, d_X). The value of $\gamma(G, d_X^2)$ is sensitive to the choice of metric space (X, d_X) and it can be very different from the reciprocal of the spectral gap of G. We refer to [MN14] for more information on nonlinear spectral gaps.

Given $d \in \mathbb{N}$, a sequence of d-regular graphs $\{G_n\}_{n=1}^\infty$ is said to be an expander sequence with respect to a metric space (X, d_X) if $\lim_{n\to\infty} |V_{G_n}| = \infty$ and $\sup_{n\in\mathbb{N}} \gamma(G_n, d_X^2) < \infty$. Note that by Cheeger's inequality [Che70, AM85] for every graph G we have

$$|X| \geqslant 2 \implies \gamma(G, d_X^2) \gtrsim \frac{1}{\sqrt{1 - \lambda_2(G)}}.$$

Hence, unless X is a singleton, if $\{G_n\}_{n=1}^\infty$ is an expander sequence with respect to (X, d_X) then $\sup_{n\in\mathbb{N}} \lambda_2(G_n) < 1$. This means that for every nontrivial metric space (X, d_X), being an expander sequence with respect to (X, d_X) is a stronger requirement than being an expander sequence in the classical sense.

Nonlinear spectral gaps first arose in the context of bi-Lipschitz embeddings; notable examples include the works of Enflo [Enf76], Gromov [Gro83], Bourgain, Milman and Wolfson [BMW86], Pisier [Pis86], Linial, London and Rabinovich [LLR95], and Matoušek [Mat97] (examples of more recent applications of nonlinear spectral gaps to bi-Lipschitz embeddings appear in [BLMN05, KN06]). Gromov [Gro03] studied nonlinear spectral gaps in the context of coarse embeddings and the Novikov conjecture, a direction that has been pursued in the works of Ozawa [Oza04], Kasparov and Yu [KY06], V. Lafforgue [Laf08, Laf09, Laf10], Pisier [Pis10], and ourselves [MN10, MN14]. Nonlinear spectral gaps also arise in fixed point theory for group actions; see the works of Wang [Wan98, Wan00], Gromov [Gro03], Izeki and Nayatani [IN05], Pansu [Pan09], Naor and Silberman [NS11], and Izeki, Kondo and Nayatani [IKN12]. We refer to the works of Gromov [Gro01] and Pichot [Pic08] for additional geometric applications of nonlinear spectral gaps. In Section 2 we discuss the relevance of nonlinear spectral gaps to approximation algorithms, based on ideas of Indyk [Ind99] and Barhum, Goldreich and Shraibman [BGS07].

Answering a question of Kasparov and Yu [KY06], V. Lafforgue proved [Laf08] that there exists a sequence of bounded degree graphs that are expanders with respect to *every* uniformly convex normed space; such graph sequences are called super-expanders. In [MN14] we found a different construction of super-expanders; here we show that our method can be applied to situations in which it seems difficult to use Lafforgue's (algebraic) approach.

It is a challenging question to characterize those metric spaces with respect to which there exist expander sequences. Random regular graphs are with high probability expanders in the classical sense (i.e., with respect to $\mathbb{R}$), but we are far from understanding those metric spaces with respect to which random regular graphs are expanders. One of the consequences of the results obtained here is that there exists a metric space (X, d_X) with respect to which there exists an expander sequence, yet almost surely random regular graphs are not expanders with respect to (X, d_X). No such example was previously known.

Let $\mathbb{G}_n$ be the set of all graphs on the vertex set $\{1, \ldots, n\}$. The subset of $\mathbb{G}_n$ consisting of all the connected graphs is denoted $\mathbb{G}_n^{\mathrm{con}}$. Given an integer $d \geqslant 3$, let $\mathcal{G}_{n,d}$ be the probability measure on $\mathbb{G}_n$ which is uniform over all those graphs in $\mathbb{G}_n$ that are d-regular and have no self-loops and no parallel edges. By [Wor81], $\lim_{n\to\infty} \mathcal{G}_{n,d}(\mathbb{G}_n^{\mathrm{con}}) = 1$.

A Hadamard space (also known as a complete $CAT(0)$ space) is a complete metric space (X, d_X) with the property that for every $x, y \in X$ there exists a point $w \in X$ such that for every $z \in X$ we have

$$d_X(z,w)^2 + \frac{1}{4}d_X(x,y)^2 \leqslant \frac{1}{2}d_X(z,x)^2 + \frac{1}{2}d_X(z,y)^2. \tag{5}$$

See the books [Jos97, BH99] and the survey [Stu03] for an extensive account of Hadamard spaces. One of the main questions left open in [NS11] (specifically, see page 1547 of [NS11]) is whether or not it is true that if (X, d_X) is a Hadamard space that admits at least one expander sequence then every classical expander sequence is also an expander sequence with respect to (X, d_X). Theorem 1.1 below, which is the first of our two main theorems, answers this question.

Theorem 1.1. *There exist a Hadamard space (X, d_X) and a sequence of 3-regular graphs $\{G_n\}_{n=1}^{\infty}$ with $\lim_{n\to\infty} |V_{G_n}| = \infty$ such that*

$$\sup_{n\in\mathbb{N}} \gamma(G_n, d_X^2) < \infty, \tag{6}$$

yet there exists $c \in (0,\infty)$ such that for every $d \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathcal{G}_{n,d}\left(\left\{H \in \mathbb{G}_n :\ \gamma(H, d_X^2) \geqslant c(\log_d n)^2\right\}\right) = 1. \tag{7}$$

Theorem 1.1 answers the above mentioned question from [NS11]. Indeed, by (6) the Hadamard space (X, d_X) admits some expander sequence. Random regular graphs are asymptotically almost surely classical expanders [Bol88], i.e., there exists $C \in (0,\infty)$ such that for every integer $d \geqslant 3$,

$$\lim_{n\to\infty} \mathcal{G}_{n,d}\left(\left\{H \in \mathbb{G}_n :\ \gamma(H, d_{\mathbb{R}}^2) \leqslant C\right\}\right) = 1. \tag{8}$$

Consequently, it follows from (7) and (8) that not all classical expander sequences are expanders with respect to (X, d_X).

Theorem 1.1 yields the first known example of a metric space (X, d_X) with respect to which random regular graphs are asymptotically almost surely not expanders, yet there does exist a special graph sequence $\{G_n\}_{n=1}^{\infty}$ that is an expander sequence with respect to (X, d_X). Observe that $\{G_n\}_{n=1}^{\infty}$ is a fortiori a classical expander sequence.

The graphs $\{G_n\}_{n=1}^{\infty}$ of Theorem 1.1 have desirable properties which no other expander sequence is known to satisfy. Specifically, $\{G_n\}_{n=1}^{\infty}$ are expanders with respect to random regular graphs. This is made precise in the following theorem.

Theorem 1.2. *There exists a universal constant $\Gamma \in (0,\infty)$ and a sequence of 3-regular graphs $\{G_n\}_{n=1}^{\infty}$ with*

$$\lim_{n\to\infty} |V_{G_n}| = \infty \qquad \text{and} \qquad \sup_{n\in\mathbb{N}} \frac{|V_{G_{n+1}}|}{|V_{G_n}|} < \infty, \tag{9}$$

such that for every integer $d \geqslant 3$ we have

$$\lim_{m\to\infty} \mathcal{G}_{m,d}\left(\left\{H \in \mathbb{G}_m^{\mathrm{con}} :\ \sup_{n\in\mathbb{N}} \gamma\big(G_n, d_H^2\big) < \Gamma\right\}\right) = 1. \tag{10}$$

In (10) we restrict to $H \in \mathbb{G}_m^{\mathrm{con}}$ because the metric d_H is defined only when H is connected. Explicit bounds on the rates of convergence in (7) and (10) are given Section 4.1.

Take $G \in \{G_n\}_{n=1}^{\infty}$ and write $V_G = \{1, \ldots, k\}$ for some $k \in \mathbb{N}$. Theorem 1.2 asserts that almost surely as $m \to \infty$, if H is a uniformly random m-vertex d-regular graph then for every $v_1, \ldots, v_k \in V_H$ the average of $d_H(v_i, v_j)^2$ over all $i, j \in \{1, \ldots, k\}$ is at most a constant multiple of the average of $d_H(v_i, v_j)^2$ over those $i, j \in \{1, \ldots, k\}$ that are joined by an edge of G. Note that the latter average is over $3k/2 \asymp k$ numbers while the former average is over $\binom{k}{2} \asymp k^2$ numbers.

Thus G is an especially constructed fixed graph that serves as a "sparse template" for computing the average squared distance between any k vertices in a random regular graph. This yields *sublinear (deterministic) time* approximate computation of average distances in random graphs: our input size is $\Omega(k^2)$, namely all the pairwise distances, while using G we estimate $\frac{1}{k^2}\sum_{i=1}^{k}\sum_{j=1}^{k} d_H(x_i, x_j)^2$ by making only $O(k)$ distance queries. See Section 2 for more on this topic.

Once the graph G is given to us, the above statement about the average squared shortest-path distance of any k vertices of the random graph H involves elementary combinatorics and probability. Nevertheless, our proof of this statement uses methods from analysis and geometry that are interesting in their own right.

Specifically, in [MN14] we introduced an iterative approach to the construction of super-expanders, building on the zigzag iteration of Reingold, Vadhan and Wigderson [RVW02]. This approach uses estimates on martingales in uniformly convex Banach spaces. In [MN13] we extended the estimates that were needed for the construction of super-expanders (namely nonlinear spectral calculus inequalities; see Remark 3.3 below) to Hadamard spaces using an appropriate notion of nonlinear martingale. In order to apply these methods in the present setting, we consider the one-dimensional simplicial complex obtained by including all the edges of H as unit intervals. We then work with the *Euclidean cone* over this one-dimensional simplicial complex, which is an auxiliary two-dimensional (random) continuous object on which martingale methods can be applied. The definition of the Euclidean cone over a metric space will be recalled in Section 3.6 below. We prove that if H is sampled from $\mathcal{G}_{n,d}$ then with high probability its Euclidean cone is a (non-disjoint) union of two sets A_1, A_2 such that A_1 admits a bi-Lipschitz embedding into L_1 and A_2 admits a bi-Lipschitz embedding into a Hadamard space. We then treat A_1 directly using Matoušek's extrapolation lemma for Poincaré inequalities [Mat97], and we treat A_2 using the methods of [MN14, MN13].

The implementation of the above strategy is not straightforward, relying on a variety of geometric and analytic tools; a more detailed overview of our proof of Theorem 1.2 appears in Section 3.10 and Section 4.2. At this juncture we only wish to stress that our proof of Theorem 1.2 introduces a way to reason about random graphs that is potentially useful in other contexts: we consider such a graph as being embedded in a larger auxiliary continuous geometric object that allows for the use of analytic methods, even though the statement being proved involves only the vertices of the original graph.

Previous work. Nonlinear spectral gaps have been studied in the literature from several points of view, leading to some notational inconsistencies. Here we use the notation that arises naturally from bi-Lipschitz embedding theory. Pichot [Pic08] denotes the reciprocal of $\gamma(G, d_X^2)$ by $\lambda_1(G, X)$. In reference to Gromov's original definition [Gro01, Gro03], Pansu [Pan09] and Kondo [Kon12] denote the same quantity by $\lambda^{\mathrm{Gro}}(G, X)$ and $\lambda_1^{\mathrm{Gro}}(G, X)$, respectively. When X is a Hadamard space, a closely related quantity, known today as Wang's invariant, was introduced by Wang [Wan98, Wan00]; Wang's invariant is always within a factor of 2 of the reciprocal of $\gamma(G, d_X^2)$. For Hadamard spaces, Izeki and Nayatani [IN05] introduced an invariant that can be used to control the ratio between Wang's invariant for a graph G and the classical spectral gap of G.

We were motivated to revisit the question of [NS11] that Theorem 1.1 answers by the recent work of Kondo [Kon12]. Kondo's goal in [Kon12] was to construct a Hadamard space for which the Izeki-Nayatani invariant is trivial. He succeeded to do so by using the Euclidean cone over certain expander graphs. In particular he obtained a Hadamard space that contains bi-Lipschitzly copies of some (classical) expanders. Gromov considered the same construction in [Gro01, Gro03] for a different but related purpose. The main point of Theorem 1.1 is to prove that the space X admits a sequence of expanders; we achieve this by modifying the Gromov-Kondo construction so that it will be compatible with the method to construct nonlinear expanders of [MN14]. The fact that our graphs are expanders with respect to random graphs requires additional work, relying on a structural result for Euclidean cones over random graphs that is presented in Section 3.9.

Roadmap. In Section 2 we describe an algorithmic implication of Theorem 1.2. Because the proofs of Theorem 1.1 and Theorem 1.2 use ingredients from several fields, Section 3 is devoted to a detailed explanation of the background and main tools that will be used in the proof of Theorem 1.1 and Theorem 1.2. The proofs themselves are given in Section 4, a section that is self-contained modulo some (quite substantial) ingredients that are presented in Section 3 and whose proof appears in later sections. The high-level structure of the argument is best discerned from reading Section 4, since it uses as a "black box" some conceptual ingredients whose proof is quite lengthy. Section 3.6 investigates the Lipschitz structure of Euclidean cones, proving in particular that the Euclidean cone over L_1 admits a bi-Lipschitz embedding into L_1. Section 6 is devoted to showing that sufficiently sparse graphs admit a bi-Lipschitz embedding into L_1. Section 7 deals with random regular graphs, proving in particular a crucial structure theorem (that holds true with high probability) for the Euclidean cone over a random graph. In Section 8 we present a partial result towards an open question that was posed by J. Kleinberg. For the formulation of Kleinberg's question itself see Section 2.1.

2. Sublinear average distance approximation algorithms

Suppose that (X, d_X) is a metric space and we have oracle access to pairwise distances in X. Given $x_1, \ldots, x_n \in X$ write

$$A \stackrel{\mathrm{def}}{=} \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_X(x_i, x_j)^2. \tag{11}$$

One can compute A exactly with n^2 distance queries. But, we wish to estimate A in *sublinear time*, i.e., with only $o(n^2)$ distance queries.

Indyk [Ind99] proved that it is possible to approximate A up to a factor of $1+\varepsilon$ by querying the distances between $O(n/\varepsilon^{7/2})$ pairs of points chosen uniformly at random. Barhum, Goldreich and Shraibman [BGS07] improved the required number of uniformly random pairs of points to $O(n/\varepsilon^2)$, which is asymptotically tight [BGS07].

The above simple randomized sampling algorithm shows that for every $x_1, \ldots, x_n \in X$ one can find $O(n)$ pairs of points from $\{x_1, \ldots, x_n\}$ whose average distance is within $O(1)$ of A, but these pairs depend on the initial point set $\{x_1, \ldots, x_n\} \subseteq X$. Following [BGS07], for $D \in [1, \infty)$ we say that a graph $G = (\{1, \ldots, n\}, E)$ is a *D-universal approximator* with respect to (X, d_X^2) if there exists (a scaling factor) $s \in (0, \infty)$ such that for every $x_1, \ldots, x_n \in X$ we have

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_X(x_i, x_j)^2 \leqslant \frac{s}{|E|}\sum_{\{i,j\}\in E} d_X(x_i, x_j)^2 \leqslant \frac{D}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_X(x_i, x_j)^2.$$

In [BGS07] it is shown that there exists $c \in (0, \infty)$ such that if $G = (\{1, \ldots, n\}, E)$ is a D-universal approximator with respect to

(X, d_X^2) for *every* metric space (X, d_X) then

$$|E| \gtrsim \frac{n^{1+c/\sqrt{D}}}{\sqrt{D}}.$$

It was also shown in [BGS07] that there exists $C \in (0, \infty)$ such that for every $D \in [1, \infty)$ and $n \in \mathbb{N}$ there exists a graph $G = (\{1, \ldots, n\}, E)$ that is a D-universal approximator with respect to (X, d_X^2) for *every* metric space (X, d_X), and such that

$$|E| \leqslant \sqrt{D} \cdot n^{1+C/\sqrt{D}}.$$

We note that [BGS07] deals with the analogous question for universal approximators when the quantity A in (11) is defined with the distances raised to power 1 rather than being squared. However, the arguments of [BGS07] easily extend mutatis mutandis to yield the above stated results (and, in fact, to analogous statements when A is defined in terms of distances raised to power p for any $p \geqslant 1$).

We thus have a satisfactory understanding of the size of universal approximators with respect to *all* metric spaces. But, for special metric spaces it is possible to obtain better tradeoffs. Indeed, in [BGS07] it is observed that an expander graph is a linear size $O(1)$-universal approximator with respect to Hilbert space; this is nothing more than an interpretation of (1), though by being more careful, and using Ramanujan graphs [LPS88, Mar88] of appropriate degree, it is shown in [BGS07] how to obtain a $1 + \varepsilon$ approximation for every $\varepsilon \in (0, 1)$.

Due to (2) and (4), for every graph G and every metric space (X, d_X), if we set $D = 4\gamma(G, d_X^2)$ then G is D-universal approximator with respect to (X, d_X^2). Hence, the super-expanders of [Laf09] and [MN14] are $O_X(1)$-universal approximators with respect to $(X, \|\cdot\|_X^2)$ for every uniformly convex Banach space $(X, \|\cdot\|_X)$ (by $O_X(1)$ we mean that the approximation factor depends only on X, in fact, it depends only on the modulus of uniform convexity of X).

Theorem 1.2 yields the only known construction of bounded degree $O(1)$-universal approximators with respect to random regular graphs; this is the content of Theorem 2.1 below. Graphs sampled from $\mathcal{G}_{n,d}$ occur in various application areas, e.g. in networking, where they serve as models for peer-to-peer networks [MS05]. Therefore, a data structure that can compute quickly the average squared distance of a given subset of a random regular graph is of theoretical interest. However, the potential practicality of our data structure is questionable because the approximation guarantee is a large universal constant; we made no attempt to improve this aspect of the construction.

Theorem 2.1. *There exists $D \in [1,\infty)$ and for every $n \in \mathbb{N}$ there exists a graph $U_n = (\{1,\ldots,n\}, E_n)$ with $|E_n| = O(n)$ such that for every two integers $m, d \geqslant 3$, if H is sampled from the restriction of $\mathcal{G}_{m,d}$ to $\mathbb{G}_m^{\mathrm{con}}$ then with probability that tends to 1 as $m \to \infty$ the graphs $\{U_n\}_{n=1}^{\infty}$ are D-universal approximators with respect to (V_H, d_H^2).*

Proof. Let $\{G_k\}_{k=1}^{\infty}$ be the graphs from Theorem 1.2. Fixing $n \in \mathbb{N}$, it follows from (9) that there exists $k \in \mathbb{N}$ such that $n \leqslant |V_{G_k}| \leqslant Mn$, where $M \in \mathbb{N}$ is a universal constant. Partition V_{G_k} into n disjoint sets $A_1, \ldots, A_n \subseteq V_{G_k}$ satisfying

$$\forall\, i \in \{1,\ldots,n\}, \qquad |A_i| \in \left\{ \left\lfloor \frac{|V_{G_k}|}{n} \right\rfloor, \left\lfloor \frac{|V_{G_k}|}{n} \right\rfloor + 1 \right\}. \tag{12}$$

Define the edge multi-set E_n of U_n by letting the number of edges joining $i \in \{1,\ldots,n\}$ and $j \in \{1,\ldots,n\}$ be equal to the total number of edges joining A_i and A_j in G_k. Then $|E_n| \leqslant |E_{G_k}| = \frac{3}{2}|V_{G_k}| \leqslant \frac{3}{2}Mn$.

Suppose that $H \in \mathbb{G}_m$ is connected and

$$\sup_{k\in\mathbb{N}} \gamma\left(G_k, d_H^2\right) < \Gamma, \tag{13}$$

where Γ is the constant from Theorem 1.2. It follows from (10) that if $H \in \mathbb{G}_m$ is sampled from $\mathcal{G}_{m,d}$ then the above asumptions hold true with probability that tends to 1 as $m \to \infty$.

Fix $x_1, \ldots, x_n \in V_H$ and define $f : V_{G_k} \to V_H$ by setting $f(u) = x_i$ for $u \in A_i$. By the definition of $\gamma(G_k, d_H^2)$ combined with (13) and (2),

$$\begin{aligned} \frac{1}{|V_{G_k}|^2} \sum_{(u,v)\in V_{G_k}\times V_{G_k}} d_H(f(u),f(v))^2 &\leqslant \frac{\Gamma}{|E_{G_k}|} \sum_{\{u,v\}\in E_{G_k}} d_H(f(u),f(v))^2 \\ &\leqslant \frac{4\Gamma}{|V_{G_k}|^2} \sum_{(u,v)\in V_{G_k}\times V_{G_k}} d_H(f(u),f(v))^2. \end{aligned} \tag{14}$$

Since

$$\sum_{(u,v)\in V_{G_k}\times V_{G_k}} d_H(f(u),f(v))^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} |A_i|\cdot|A_j|\cdot d_H(x_i,x_j)^2,$$

it follows from (12) that

$$\begin{aligned} \frac{n^2\lfloor |V_{G_k}|/n\rfloor^2}{|V_{G_k}|^2} &\leqslant \frac{\frac{1}{|V_{G_k}|^2}\sum_{(u,v)\in V_{G_k}\times V_{G_k}} d_H(f(u),f(v))^2}{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_H(x_i,x_j)^2} \\ &\leqslant \frac{n^2\left(\lfloor |V_{G_k}|/n\rfloor + 1\right)^2}{|V_{G_k}|^2}. \end{aligned} \tag{15}$$

Also, by the definition of E_n and f we have

$$\sum_{\{u,v\}\in E_{G_k}} d_H(f(u),f(v))^2 = \sum_{\{i,j\}\in E_n} d_H(x_i,x_j)^2. \tag{16}$$

By substituting (15) and (16) into (14) we conclude that

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_H(x_i,x_j)^2 &\leqslant \frac{\Gamma|V_{G_k}|^2|E_n|}{n^2\lfloor |V_{G_k}|/n\rfloor^2|E_{G_k}|}\cdot\frac{1}{|E_n|}\sum_{\{i,j\}\in E_n} d_H(x_i,x_j)^2\\ &\leqslant \frac{4\Gamma\left(1+\frac{1}{\lfloor |V_{G_k}|/n\rfloor}\right)^2}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_H(x_i,x_j)^2\\ &\leqslant \frac{16\Gamma}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_H(x_i,x_j)^2.\end{aligned}$$

This means that U_n is a D-universal approximator with respect to (V_H, d_H^2), where $D = 16\Gamma$. □

Remark 2.2. A straightforward inspection of our proof of Theorem 1.2 reveals that for every $n \in \mathbb{N}$ the universal approximator U_n of Theorem 2.1 can be constructed in deterministic $O(n)$ time.

Remark 2.3. Let (X, d_X) be a metric space and $p \in [1,\infty)$. As noted above, one can study universal approximators with respect to (X, d_X^p), i.e., given $x_1,\dots,x_n \in X$ the goal is approximate computation of the quantity $\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_X(x_i,x_j)^p$ with $o(n^2)$ distance queries. The references [Ind99, BGS07] deal with $p = 1$, but as we mentioned earlier, they extend painlessly to general $p \geqslant 1$. Here we have only dealt with the case $p = 2$, but we speculate that our argument can be modified to yield bounded degree universal approximators with respect to (H, d_H^p) for any $p \in (1,\infty)$, where H is a random d-regular graph. We did not, however, attempt to carry out our proof when $p \neq 2$. We also speculate that the case $p = 1$ requires more substantial new ideas, as the type of martingale arguments that we use typically fail at the endpoint $p = 1$.

2.1. **A question of J. Kleinberg.** In connection with Theorem 1.2 and the algorithmic context described in Section 2, Jon Kleinberg asked us whether or not two stochastically independent random 3-regular graphs are asymptotically almost surely expanders with respect to each other. Formally, we have the following open question.

Question 2.4 (Jon Kleinberg). Does there exist a universal constant $K \in (0,\infty)$ such that

$$\lim_{\min\{m,n\}\to\infty} \mathcal{G}_{m,3}\times\mathcal{G}_{n,3}\left[\left\{(G,H)\in\mathbb{G}_m\times\mathbb{G}_n^{\mathrm{con}}:\ \gamma\left(G,d_H^2\right)\right\}\leqslant K\right]=1.$$

While a positive solution of Question 2.4 does not formally imply that there exist graphs $\{G_n\}_{n=1}^\infty$ as in Theorem 1.2, such a statement would be a step towards a probabilistic construction of such graphs. At present we do not have methods to argue about the nonlinear spectral gap of random graphs. In particular, it is a major open question whether or not random 3-regular graphs are super-expanders with positive probability. Thus, all the known constructions in the context of nonlinear spectral gaps are deterministic, with the only two methods that are currently available being Lafforgue's algebraic approach [Laf08] and our iterative approach [MN14]. It seems, however, that the problem of obtaining a probabilistic proof of the existence of expanders with respect to random graphs is more tractable, and for this reason we believe that Question 2.4 is a promising research direction. In Section 8 we describe a partial result towards a positive solution of Question 2.4 that we obtained through discussions with Uriel Feige; we thank him for allowing us to include his insights here.

3. Preliminaries

In this section we set notation and terminology that will be used throughout the ensuing discussion. We also present the tools and main steps that will be used in the proofs of Theorem 1.1 and Theorem 1.2.

3.1. **Graph theoretical definitions.** We start by defining some basic concepts in graph theory. Most of what we describe here is standard or self-evident terminology, and in fact some of it was used in the introduction without being defined explicitly.

Recall that graphs can have parallel edges and self-loops unless stated otherwise. Thus, when discussing a graph G it will always be understood that E_G is a *multi-subset* of unordered pairs of vertices, i.e., for every $u,v\in V_G$ the unordered pair $\{u,v\}$ is allowed to appear in E_G multiple times. For $(u,v)\in V_G\times V_G$, denote by $E_G(u,v)$ the number of times that $\{u,v\}$ appears in E_G. The graph G is therefore determined by the integer matrix $(E_G(u,v))_{(u,v)\in V_G\times V_G}$. A graph is called *simple* if it contains no self-loops and no parallel edges, which is equivalent to requiring that $E_G(u,u)=0$ for every $u\in V_G$ and $E_G(u,v)\in\{0,1\}$ for every $u,v\in V_G$. Given $S,T\subseteq V_G$, we denote by $E_G(S,T)$ the multi-set of all edges in E_G that join a vertex in S and a vertex in T (hence for

$(u,v)\in V_G\times V_G$ we have $E_G(u,v)=|E_G(\{u\},\{v\})|$). When $S=T$ we use the simpler notation $E_G(S)=E_G(S,S)$.

The degree of a vertex $u\in V_G$ is $\deg_G(u)=\sum_{v\in V_G}E_G(u,v)$. Under this convention each self-loop contributes 1 to the degree of a vertex. For $d\in\mathbb{N}$, a graph G is d-regular if $\deg_G(u)=d$ for every $u\in V_G$.

For a graph G and a symmetric function $K:V_G\times V_G\to\mathbb{R}$ satisfying $K(u,u)=0$ for every $u\in V_G$, when we write the sum $\sum_{\{u,v\}\in E_G}K(u,v)$ we mean that each edge is counted once, i.e,

$$\begin{aligned}\sum_{\{u,v\}\in E_G}K(u,v)&\stackrel{\mathrm{def}}{=}\frac12\sum_{\substack{(u,v)\in V_G\times V_G\\ \{u,v\}\in E_G}}K(u,v)\\&=\frac12\sum_{(u,v)\in V_G\times V_G}E_G(u,v)K(u,v).\end{aligned}$$

If G is a connected graph then the diameter of the metric space (V_G,d_G) is denoted $\mathrm{diam}(G)$. The one-dimensional simplicial complex induced by a graph G is denoted $\Sigma(G)$. Thus $\Sigma(G)$ is obtained from G by including the edges of G as unit intervals. The geodesic distance on $\Sigma(G)$ is denoted $d_{\Sigma(G)}$. Then

$$\mathrm{diam}(G)\leqslant\mathrm{diam}(\Sigma(G))\leqslant\mathrm{diam}(G)+1. \tag{17}$$

Given a graph G, a subset $C\subseteq V_G$ is called a cycle of G if it is possible to write $C=\{x_1,\ldots,x_k\}$, where the vertices $x_1,\ldots,x_k\in V$ are distinct and

$$\{x_1,x_2\},\{x_2,x_3\},\ldots,\{x_{k-1},x_k\},\{x_k,x_1\}\in E_G. \tag{18}$$

Thus a self-loop is a cycle of length 1 and a parallel edge induces a cycle of length 2. By definition, $|E_G(C)|\geqslant|C|$ for every cycle $C\subseteq V_G$. A cycle $C\subseteq V_G$ is said to be an induced cycle if $|E_G(C)|=|C|$, i.e., $E_G(C)$ consists only of the edges listed in (18). Note that the smallest cycle in G is necessarily induced. The girth of a graph G, denoted $\mathrm{girth}(G)$, is the size of the smallest cycle in G. If a connected graph G does not contain a cycle, i.e., G is a tree, then we shall use the convention $\mathrm{girth}(G)=2\,\mathrm{diam}(G)$.

The normalized adjacency matrix of a d-regular graph G, denoted A_G, is the $|V_G|$ by $|V_G|$ symmetric stochastic matrix whose entry at $(u,v)\in V_G\times V_G$ is equal to $E_G(u,v)/d$. The decreasing rearrangement of the eigenvalues of A_G is denoted

$$1=\lambda_1(G)\geqslant\lambda_2(G)\geqslant\ldots\geqslant\lambda_{|V_G|}(G).$$

3.2. **Bi-Lipschitz embeddings.** A metric space (U, d_U) is said to admit a bi-Lipschitz embedding with distortion at most $D \in [1, \infty)$ into a metric space (V, d_V) if there exists (a scaling factor) $s \in (0, \infty)$ and $f : U \to V$ that satisfies

$$\forall\, x, y \in U, \qquad s d_U(x, y) \leqslant d_V\big(f(x), f(y)\big) \leqslant D s d_U(x, y). \tag{19}$$

The infimum over those $D \in (0, \infty)$ for which (U, d_U) admits a bi-Lipschitz embedding with distortion at most D into (V, d_V) is denoted $c_{(V,d_V)}(U, d_U)$, or simply $c_V(U)$ if the respective metrics are clear from the context (if no such $D \in (0, \infty)$ exists then we set $c_V(U) = \infty$). For $p \in [1, \infty]$ we use the simpler notation $c_p(U) = c_{L_p}(U)$. The quantity $c_2(U)$ is known as the Euclidean distortion of U and the quantity $c_1(U)$ is known as the L_1 distortion of U.

If G is a connected simple graph with $|V_G| = n$ then $c_p(\Sigma(G)) \lesssim n$ for every $p \in [1, \infty]$. Indeed, write $V_G = \{1, \ldots, n\}$ and let $e_1, \ldots, e_n \in \mathbb{R}^n$ be the coordinate basis of $\mathbb{R}^n$. If $\{i, j\} \in E_G$ and $x \in \Sigma(G)$ is a point on the unit interval that corresponds to the edge joining i and j, then map x to $(1 - t)e_i + te_j$, where t is the distance between x and i in $\Sigma(G)$. This trivial embedding can be improved: one has $c_p(\Sigma(G)) \lesssim \log(n + 1)$. The estimate $c_p(V_G, d_G) \lesssim \log(n + 1)$ is the classical Bourgain embedding theorem [Bou85], but it is simple to argue that the same bound holds true for embeddings of the one-dimensional simplicial complex of G rather than just its vertices. Using the better bound $c_p(\Sigma(G)) \lesssim \log(n+1)$ in what follows can improve the implicit constants in Theorem 1.1 and Theorem 1.2, but since we ignore constant factors here it will suffice to use the trivial bound $c_p(\Sigma(G)) \lesssim n$.

For future use we also recall that an argument from [LLR95] (see also [MN14, Sec. 1.1]) shows that for every connected d-regular graph H and every metric space (Z, d_Z),

$$c_{(Z,d_Z)}\left(V_H, d_H\right) \gtrsim \frac{\log_d n}{\sqrt{\gamma(H, d_Z^2)}}. \tag{20}$$

3.3. **Nonlinear absolute spectral gaps.** Fix $d, n \in \mathbb{N}$ and suppose that G is an n-vertex d-regular graph and that (X, d_X) is a metric space. Define $\gamma_+(G, d_X^2)$ to be the infimum over those $\gamma_+ \in (0, \infty]$ such that for every two mappings $f, g : V_G \to X$ we have

$$\frac{1}{n^2} \sum_{(u,v)\in V_G\times V_G} d_X\big(f(u), g(v)\big)^2 \leqslant \frac{\gamma_+}{nd} \sum_{(u,v)\in V_G\times V_G} E_G(u, v) \cdot d_X\big(f(u), g(v)\big)^2. \tag{21}$$

Comparison of (21) to (4) reveals that $\gamma(G, d_X^2) \leqslant \gamma_+(G, d_X^2)$.

With this notation one has

$$\gamma_+\big(G, d_{\mathbb{R}}^2\big) = \frac{1}{1 - \max\big\{\lambda_2(G), -\lambda_{|V_G|}(G\big\}}.$$

For this reason we think of $\gamma_+(G, d_X^2)$ as measuring the reciprocal of the nonlinear absolute spectral gap of G with respect to (X, d_X).

Despite the fact that Theorem 1.1 and Theorem 1.2 deal with the quantity $\gamma(\cdot,\cdot)$, for their proofs it will be very convenient to work with the quantity $\gamma_+(\cdot,\cdot)$. See [MN14] for more on nonlinear absolute spectral gaps, specifically Section 2.2 of [MN14] for information on the relation between $\gamma(\cdot,\cdot)$ and $\gamma_+(\cdot,\cdot)$.

As in [MN14], it is convenient to also work with nonlinear spectral gaps of symmetric stochastic matrices. Thus, letting $M = (m_{ij})$ be an n by n symmetric stochastic matrix and (X, d_X) be a metric space, define $\gamma(M, d_X^2)$ to be the infimum over those $\gamma \in (0, \infty]$ such that for every $x_1, \ldots, x_n \in X$ we have

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_X(x_i, x_j)^2 \leqslant \frac{\gamma}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij} d_X(x_i, x_j)^2.$$

Also, define $\gamma_+(M, d_X^2)$ to be the infimum over those $\gamma_+ \in (0, \infty]$ such that for every $x_1, \ldots, x_n, y_1, \ldots, y_n \in X$ we have

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} d_X(x_i, y_j)^2 \leqslant \frac{\gamma}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij} d_X(x_i, y_j)^2.$$

Under these definitions, one checks that for every regular graph G we have $\gamma(G, d_X^2) = \gamma(A_G, d_X^2)$ and $\gamma_+(G, d_X^2) = \gamma_+(A_G, d_X^2)$.

3.4. **Graph products, edge completion, and Cesàro averages.** Fix $d_1, d_2 \in \mathbb{N}$. Let G_1 be a d_1-regular graph and let G_2 be a d_2-regular graph. Suppose that $|V_{G_2}| = d_1$. Then one can construct a new graph, called the *zigzag product* of G_1 and G_2 and denoted $G_1 \text{ⓩ} G_2$. This construction is due to Reingold, Vadhan and Wigderson [RVW02]. We will not need to recall the definition of $G_1 \text{ⓩ} G_2$ here: all that we will use below is that $G_1 \text{ⓩ} G_2$ is d_2^2-regular, $|V_{G_1 \text{ⓩ} G_2}| = |V_{G_1}| \cdot |V_{G_2}|$, and for every metric space (X, d_X) we have

$$\gamma_+\big(G_1 \text{ⓩ} G_2, d_X^2\big) \leqslant \gamma_+\big(G_1, d_X^2\big) \cdot \gamma_+\big(G_2, d_X^2\big)^2. \tag{22}$$

The inequality (22) is due to [MN14].

Under the same assumptions on the graphs G_1, G_2, i.e., that G_1 is d_1-regular, G_2 is d_2-regular, and $|V_{G_2}| = d_1$, one can also construct a new graph, called the *replacement product* of G_1 and G_2 and denoted

$G_1 \circledR G_2$. This construction is due to Gromov [Gro83] (see also [RVW02] and [MN14, Sec 8.3]). Again, we will not need to recall the definition of $G_1 \circledR G_2$ here: all that we will use below is that the graph $G_1 \circledR G_2$ is (d_2+1)-regular, $|V_{G_1 \circledR G_2}| = |V_{G_1}| \cdot |V_{G_2}|$, and for every metric space (X, d_X) we have

$$\gamma_+\big(G_1 \circledR G_2, d_X^2\big) \leqslant 3(d_2+1) \cdot \gamma_+\big(G_1, d_X^2\big) \cdot \gamma_+\big(G_2, d_X^2\big)^2. \tag{23}$$

The inequality (23) is due to [MN14].

Fix $d \in \mathbb{N}$ and suppose that G is a d-regular graph. For every integer $D \geqslant d$ one can define a new graph called the *D-edge completion* of G, and denoted $\mathscr{C}_D(G)$. See [MN14, Def. 2.8] for the definition of $\mathscr{C}_D(G)$. All that we will use below is that $\mathscr{C}_D(G)$ is D-regular, $V_{\mathscr{C}_D(G)} = V_G$, and for every metric space (X, d_X) we have

$$\gamma_+\big(\mathscr{C}_D(G), d_X^2\big) \leqslant 2\gamma_+\big(G, d_X^2\big). \tag{24}$$

The proof of (24) is contained in [MN14, Lem. 2.9].

We will also work below with *Cesàro averages* of graphs. Given $d, m \in \mathbb{N}$ and a d-regular graph G, its mth Cesàro average $\mathscr{A}_m(G)$ is a new graph defined by $V_{\mathscr{A}_m(G)} = V_G$ and

$$\forall\, (u,v) \in V_G \times V_G, \qquad E_{\mathscr{A}_m(G)}(u,v) \stackrel{\mathrm{def}}{=} \sum_{t=0}^{m-1} d^{m-1-t} \left(A_G^t\right)_{u,v},$$

where we recall that A_G is a the normalized adjacency matrix of G. One checks that $\mathscr{A}_m(G)$ is md^{m-1}-regular and that the adjacency matrix of $\mathscr{A}_m(G)$ is the corresponding Cesàro average of A_G, i.e.,

$$A_{\mathscr{A}_m(G)} = \frac{1}{m}\sum_{t=0}^{m-1} A_G^t.$$

The following lemma will be used (twice) in Section 4.2.

Lemma 3.1. *Fix two integers $d, n \geqslant 3$ and let G be a d-regular graph with $|V_G| = 2n$. Then there exists a 3-regular graph G^* satisfying $|V_{G^*}| = 36dn = 18d|V_G|$ such that for every metric space (X, d_X),*

$$\gamma_+\big(G^*, d_X^2\big) \lesssim d^4 \gamma\big(G, d_X^2\big). \tag{25}$$

Proof. By Lemma 2.6 in [MN14], there exists a $4d$-regular graph G' with $|V_{G'}| = n$ such that $\gamma_+(G', d_X^2) \leqslant 8\gamma(G, d_X^2)$ for every metric space (X, d_X). Let C_{4d}° be the cycle of length $4d$ in which each vertex has exactly one self-loop (thus C_{4d}° is a 3-regular graph). Since G' is a $4d$-regular graph, we may form the zigzag product $G'' \circledz C_{4d}^\circ$, which is a 9-regular graph with $4dn$ vertices. By [MN14, Lem. 2.1] we have $\gamma_+(C_{4d}^{\circ}, d_X^2) \leqslant 192d^2$. It therefore follows from (22) that

$$\gamma_+\big(G'', d_X^2\big) \leqslant 8\gamma(G, d_X^2) \cdot (192d^2)^2 \lesssim d^4\gamma(G, d_X^2). \tag{26}$$

Now, let C_9 be the cycle of length 9 (thus C_9 is a 2-regular graph with 9 vertices). Define G^* to be the replacement product $G^* = G'' \textcircled{r} C_9$, which is a 3-regular graph with $36dn$ vertices. By [MN14, Lem. 2.1] we have $\gamma_+(C_9, d_X^2) \leqslant 648$, so (25) follows from (26) and (23). $\square$

3.4.1. *A zigzag iteration.* Theorem 3.2 below is a (much simpler) variant of the iterative procedure by which we constructed super-expanders in [MN14], building on the zigzag iteration of Reingold, Vadhan and Wigderson [RVW02].

Theorem 3.2. *Fix $K \in [1, \infty)$ and two integers $n, d \geqslant 3$ with $n \geqslant d^3$. Suppose that (X, d_X) is a metric space with the property that for every regular graph G we have*

$$\gamma_+\big(\mathscr{A}_m(G), d_X^2\big) \leqslant K \max\left\{1, \frac{\gamma_+(G, d_X^2)}{m}\right\}, \tag{27}$$

where

$$m \stackrel{\mathrm{def}}{=} \left\lfloor \frac{\log n}{3\log d} \right\rfloor. \tag{28}$$

Suppose further that there exists a d-regular graph H with $|V_H| = n$ such that

$$\gamma_+\big(H, d_X^2\big) \leqslant \sqrt{\frac{m}{2K}}. \tag{29}$$

Then there exists a sequence of 3-regular graphs $\{G_j\}_{j=1}^{\infty}$ satisfying

$$\forall\, j \in \mathbb{N}, \qquad \big|V_{G_j}\big| = 9d^2 n^j, \tag{30}$$

and

$$\forall\, j \in \mathbb{N}, \qquad \gamma_+\big(G_j, d_X^2\big) \lesssim d^8 K \gamma_+\big(H, d_X^2\big)^2. \tag{31}$$

Proof. We first claim that for every $j \in \mathbb{N}$ there exists a d^2-regular graph W_j such that $|V_{W_j}| = n^j$ and

$$\gamma_+\big(W_j, d_X^2\big) \leqslant 2K\gamma_+\big(H, d_X^2\big)^2. \tag{32}$$

The proof of this statement is by induction on j. Define W_1 to be the d^2-edge completion of H, i.e,

$$W_1 \stackrel{\mathrm{def}}{=} \mathscr{C}_{d^2}(H).$$

Supposing that W_j has been defined, the Cesàro average $\mathscr{A}_m(W_j)$ is an $md^{2(m-1)}$-regular graph with $|V_{\mathscr{A}_m(W_j)}| = |V_{W_j}| = n^j$. Recalling (28), we have $md^{2(m-1)} \leqslant n$. We can therefore form the edge completion

$\mathscr{C}_n(\mathscr{A}_m(W_j))$, which is an n-regular graph, and consequently the following zigzag product is well defined.

$$W_{j+1} \stackrel{\mathrm{def}}{=} \mathscr{C}_n(\mathscr{A}_m(W_j)) \textcircled{z} H. \tag{33}$$

The graph W_{j+1} is d^2-regular and $|V_{W_{j+1}}| = n|V_{\mathscr{C}_n(\mathscr{A}_m(W_j))}| = n^{j+1}$. Moreover,

$$\begin{aligned}
\gamma_+\big(W_{j+1}, d_X^2\big) &\stackrel{(22)\wedge(33)}{\leqslant} 2\gamma_+\big(\mathscr{C}_n(\mathscr{A}_m(W_j)), d_X^2\big)\cdot\gamma_+\big(H, d_X^2\big)^2 \\
&\stackrel{(24)}{\leqslant} 2\gamma_+\big(\mathscr{A}_m(W_j), d_X^2\big)\cdot\gamma_+\big(H, d_X^2\big)^2 \\
&\stackrel{(27)}{\leqslant} 2K\max\left\{1, \frac{\gamma_+(W_j, d_X^2)}{m}\right\}\cdot\gamma_+\big(H, d_X^2\big)^2 \\
&\stackrel{(32)}{\leqslant} 2K\max\left\{1, \frac{2K\gamma_+(H, d_X^2)^2}{m}\right\}\cdot\gamma_+\big(H, d_X^2\big)^2 \\
&\stackrel{(29)}{=} 2K\gamma_+\big(H, d_X^2\big)^2.
\end{aligned}$$

This completes the inductive construction of $\{W_j\}_{j=1}^\infty$.

Next, let $C_{d^2}^\circ$ be the cycle of length d^2 in which each vertex has exactly one self-loop. We can form the zigzag product $W_j \textcircled{z} C_{d^2}^\circ$, which is a 9-regular graph with $d^2 n^j$ vertices. By [MN14, Lem. 2.1] we have $\gamma_+\big(C_{d^2}^\circ, d_X^2\big) \leqslant 12d^4$, so we deduce from (22) and (32) that

$$\gamma_+\big(W_j \textcircled{z} C_{d^2}^\circ, d_X^2\big) \leqslant 2K\gamma_+\big(H, d_X^2\big)^2\cdot(12d^4)^2 \lesssim d^8 K\gamma_+\big(H, d_X^2\big)^2. \tag{34}$$

Recalling that C_9 denotes the simple cycle of length 9, define G_j to be the replacement product

$$G_j \stackrel{\mathrm{def}}{=} \big(W_j \textcircled{z} C_{d^2}^\circ\big) \textcircled{r} C_9.$$

Then G_j is a 3-regular graph and $|V_{G_j}| = 9d^2 n^j$. By [MN14, Lem. 2.1] we have $\gamma_+(C_9, d_X^2) \leqslant 648$. It therefore follows from (23) and (34) that the desired estimate (31) holds true. $\square$

Remark 3.3. The condition (27) is called a *nonlinear spectral calculus inequality*; see [MN14] for an explanation of this terminology. Looking ahead, we will rely here on the fact that such a nonlinear spectral calculus inequality is available for Hadamard spaces, as proved in [MN13].

3.5. $CAT(0)$ **and** $CAT(1)$ **spaces.** We need to briefly recall basic definitions related to curvature upper bounds in the sense of Aleksandrov; see [BH99] for much more on this topic.

A metric space (X, d_X) is a $CAT(1)$ space if it satisfies the following conditions. First, for every $x, y \in X$ with $d_X(x, y) < \pi$ there exists a geodesic joining x to y, i.e., a curve $\phi : [0,1] \to X$ that satisfies $d_X(\phi(t), x) = td_X(x, y)$ and $d_X(\phi(t), y) = (1-t)d_X(x, y)$ for every $t \in [0,1]$. Suppose that $x, y, z \in X$ satisfy

$$d_X(x,y) + d_X(y,z) + d_X(z,x) < 2\pi,$$

and that there exist geodesics $\phi_{x,y}, \phi_{y,z}, \phi_{z,x} : [0,1] \to X$ joining x to y, y to z, and z to x, respectively. Let S^2 be the unit Euclidean sphere in $\mathbb{R}^3$, and let d_{S^2} denote the geodesic metric on S^2 (thus the diameter of S^2 equals π under this metric). As explained in [BH99], there exist $a, b, c \in S^2$ such that $d_{S^2}(a,b) = d_X(x,y)$, $d_{S^2}(b,c) = d_X(y,z)$ and $d_{S^2}(c,a) = d_X(z,x)$. Let $\varphi_{a,b}, \varphi_{b,c}, \varphi_{c,a} : [0,1] \to S^2$ be geodesics joining a to b, b to c, and c to a, respectively. Then the remaining requirement in the definition of a $CAT(1)$ space is that for every $s, t \in [0,1]$ we have $d_X(\phi_{x,y}(s), \phi_{y,z}(t)) \leqslant d_{S^2}(\varphi_{a,b}(s), \varphi_{b,c}(t))$.

Following Gromov [Gro01, Gro03] and Kondo [Kon12], we shall now describe a simple way to obtain $CAT(1)$ spaces from graphs. Suppose that $\mathcal{F}$ is a family of connected graphs that satisfies

$$R_{\mathcal{F}} \stackrel{\mathrm{def}}{=} \sup_{G \in \mathcal{F}} \frac{\mathrm{diam}(G)}{\mathrm{girth}(G)} < \infty. \tag{35}$$

Define a metric space $(U_{\mathcal{F}}, d_{\mathcal{F}})$ as follows. $U_{\mathcal{F}}$ is the formal disjoint union of the one-dimensional simplicial complexes $\{\Sigma(G)\}_{G \in \mathcal{F}}$, i.e.,

$$U_{\mathcal{F}} \stackrel{\mathrm{def}}{=} \bigsqcup_{G \in \mathcal{F}} \Sigma(G). \tag{36}$$

The metric $d_{\mathcal{F}} : U_{\mathcal{F}} \times U_{\mathcal{F}} \to [0,\infty)$ is defined by

$$d_{\mathcal{F}}(x,y) \stackrel{\mathrm{def}}{=} \begin{cases} \frac{2\pi d_{\Sigma(G)}(x,y)}{\mathrm{girth}(G)} & \text{if } \exists G \in \mathcal{F} \text{ such that } x, y \in \Sigma(G), \\ 2\pi\left(R_{\mathcal{F}} + 1\right) & \text{otherwise.} \end{cases} \tag{37}$$

$d_{\mathcal{F}}$ is indeed a metric because for every $G \in \mathcal{F}$ and every $x, y \in \Sigma(G)$,

$$\frac{2\pi d_{\Sigma(G)}(x,y)}{\mathrm{girth}(G)} \leqslant \frac{2\pi\,\mathrm{diam}(\Sigma(G))}{\mathrm{girth}(G)} \stackrel{(17)}{\leqslant} \frac{2\pi\left(\mathrm{diam}(G)+1\right)}{\mathrm{girth}(G)} \stackrel{(35)}{\leqslant} \frac{2\pi\left(R_{\mathcal{F}}\,\mathrm{girth}(G)\right)+1)}{\mathrm{girth}(G)} \leqslant 2\pi\left(R_{\mathcal{F}}+1\right). \tag{38}$$

We claim that $(U_{\mathcal{F}}, d_{\mathcal{F}})$ is a $CAT(1)$ space. Indeed, since for distinct $G, H \in \mathcal{F}$ we have $d_{\mathcal{F}}(\Sigma(G), \Sigma(H)) > 2\pi$ and for every $G \in \mathcal{F}$ the metric space $(\Sigma(G), 2\pi d_{\Sigma(G)}/\mathrm{girth}(G))$ is geodesic, every $x, y \in U_{\mathcal{F}}$ with $d_{\mathcal{F}}(x,y) < \pi$ can be joined by a geodesic. Moreover, if $x, y, z \in U_{\mathcal{F}}$

satisfy $d_{\mathcal{F}}(x,y)+d_{\mathcal{F}}(y,z)+d_{\mathcal{F}}(z,x) < 2\pi$ then necessarily $x,y,z \in \Sigma(G)$ for some $G \in \mathcal{F}$, in which case we have

$$d_{\Sigma(G)}(x,y) + d_{\Sigma(G)}(y,z) + d_{\Sigma(G)}(z,x) < \mathrm{girth}(G). \tag{39}$$

It follows that

$$\max\left\{d_{\Sigma(G)}(x,y), d_{\Sigma(G)}(y,z), d_{\Sigma(G)}(z,x)\right\} < \frac{\mathrm{girth}(G)}{2}. \tag{40}$$

Indeed, if (40) fails then without loss of generality we may assume that $d_{\Sigma(G)}(z,x) \geqslant \mathrm{girth}(G)/2$. By the triangle inequality we would then have $d_{\Sigma(G)}(x,y) + d_{\Sigma(G)}(y,z) \geqslant d_{\Sigma(G)}(z,x) \geqslant \mathrm{girth}(G)/2$, so that (39) would be violated. By (40) the geodesic triangle whose vertices are x,y,z is contained (isometrically) in a metric tree, and since metric trees are easily seen to be $CAT(1)$ spaces (see [BH99, Sec. II.1.15]), this completes the verification that $(U_{\mathcal{F}}, d_{\mathcal{F}})$ is a $CAT(1)$ space.

A metric space (X,d_X) is said to be a $CAT(0)$ space if every two points $x,y \in X$ can be joined by a geodesic, and for every $x,y,z \in X$, every geodesic $\phi : [0,1] \to X$ with $\phi(0) = y$ and $\phi(1) = z$, and every $t \in [0,1]$ we have

$$d_X(x,\phi(t))^2 \leqslant (1-t)d_X(x,y)^2 + td_X(x,z)^2 - t(1-t)d_X(y,z)^2. \tag{41}$$

A Hadamard space is a complete $CAT(0)$ space, in which case it suffices to require the validity of (41) for $t = \frac{1}{2}$.

3.5.1. *Nonlinear spectral calculus for $CAT(0)$ spaces.* In [MN13] we proved that $CAT(0)$ metric spaces satisfy a nonlinear spectral calculus inequality such as (27), a tool that will be used crucially in what follows. This is a special case of a more general result that is proved in [MN13], but we will not formulate it here in order to avoid introducing terminology that is not needed for the present context (we will use the zigzag iteration of Theorem 3.2 *only* for $CAT(0)$ spaces).

Theorem 3.4 ([MN13]). *Let (X,d_X) be a $CAT(0)$ space. Then for every $m,n \in \mathbb{N}$ and every n by n symmetric and stochastic matrix M,*

$$\gamma_+\left(\frac{1}{m}\sum_{t=0}^{m-1} M^t, d_X^2\right) \lesssim \max\left\{1, \frac{\gamma_+(M,d_X^2)}{m}\right\}, \tag{42}$$

3.6. **The Euclidean cone.** Let (X,d_X) be a metric space. The Euclidean cone over X, denoted $\mathbf{Cone}(X,d_X)$ or simply $\mathbf{Cone}(X)$ when the metric on X is clear from the context, is defined to be the completion of $(0,\infty)\times X$ under the metric

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) \stackrel{\mathrm{def}}{=} \sqrt{s^2+t^2-2st\cos\left(\min\{\pi, d_X(x,y)\}\right)}. \tag{43}$$

This definition is due to [Ber83]; see also the exposition in [ABN86] and [BH99] (in particular, the fact that (43) satisfies the triangle inequality is proved in Proposition 5.9 of [BH99, Chapter I.5]). When X itself is complete, the completion of $(0,\infty)\times X$ under the metric defined in (43) amounts to extending the definition (43) to $[0,\infty)\times X$ and identifying all of the points in $\{0\}\times X$ (i.e., adding one additional point to $(0,\infty)\times X$ as the "cusp" of the cone). The results that are discussed below pass immediately from a space to its completion, so we will mostly ignore this one-point completion in the ensuing arguments.

Observe that (43) can be rewritten as follows

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) = \sqrt{(s-t)^2+2st\left(1-\cos\left(\min\{\pi,d_X(x,y)\}\right)\right)}. \quad (44)$$

For the sake of later comparisons between cones, we record here the following very simple fact.

Fact 3.5. *Fix $D\in[1,\infty)$ and two metric space (X,d_X) and (Y,d_Y). Suppose that $f:X\to Y$ satisfies*

$$\forall\, x,y\in X,\qquad d_X(x,y)\leqslant d_Y(f(x),f(y))\leqslant Dd_X(x,y). \quad (45)$$

Then for every distinct $(s,x),(t,y)\in(0,\infty)\times X$ we have

$$1\leqslant \frac{d_{\mathbf{Cone}(Y)}\big((s,f(x)),(t,f(y))\big)}{d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)}\leqslant D. \quad (46)$$

Proof. Since $\psi(a)\stackrel{\mathrm{def}}{=}\sqrt{1-\cos(\min\{\pi,a\})}=\sqrt{2}\sin(\min\{\pi/2,a/2\})$ is concave and increasing on $[0,\infty)$, and $\psi(0)=0$, for every $a,b\in[0,\infty)$ with $a\leqslant b$ we have $\psi(b)\leqslant b\psi(a)/a$. Hence by (45) we have

$$\begin{aligned}1-\cos\left(\min\{\pi,d_X(x,y)\}\right)&\leqslant 1-\cos\left(\min\{\pi,d_Y(f(x),f(y))\}\right)\\&\leqslant D^2\left(1-\cos\left(\min\{\pi,d_X(x,y)\}\right)\right),\end{aligned}$$

which implies (46) due to (44). □

Corollary 3.6. *For every metric space (X,d_X) and every Banach space $(Y,\|\cdot\|_Y)$ we have $c_{\mathbf{Cone}(Y)}(\mathbf{Cone}(X))\leqslant c_Y(X)$.*

Proof. For $D>c_Y(X)$ take $f:X\to Y$ that satisfies (45). Define $F:(0,\infty)\times X\to(0,\infty)\times Y$ by $F(s,x)=(s,f(x))$, and use Fact 3.5. The role of Y being a Banach space was only to ensure that we can take the scaling factor s in the definition (19) to be equal to 1. □

Proposition 3.7 below, whose proof is given in Section 5.1, asserts that $c_1(\mathbf{Cone}(L_1))<\infty$, a fact that plays an important role in our proofs of Theorem 1.1 and Theorem 1.2.

Proposition 3.7. $\mathbf{Cone}(L_1)$ *admits a bi-Lipschitz embedding into* L_1.

By combining Proposition 3.7 with Corollary 3.6 we obtain the following useful corollary.

Corollary 3.8. *Every metric space* (X, d_X) *satisfies*

$$c_1(\mathbf{Cone}(X)) \lesssim c_1(X).$$

3.7. $CAT(0)$ **cones.** An important theorem of Berestovskiĭ [Ber83] (see also [ABN86] and Theorem 3.14 of [BH99, Ch. II.3]) asserts that if X is a $CAT(1)$ metric space then $\mathbf{Cone}(X)$ is a $CAT(0)$ metric space. Consequently, we have the following lemma.

Lemma 3.9. *Let* $\mathcal{F}$ *be a family of connected graphs satisfying* (35), *and let* $(U_{\mathcal{F}}, d_{\mathcal{F}})$ *be the metric space defined by* (36) *and* (37). *Then* $\mathbf{Cone}(U_{\mathcal{F}}, d_{\mathcal{F}})$ *is a Hadamard space. Moreover, for every* $G \in \mathcal{F}$ *the metric space* $(\Sigma(G), d_{\Sigma(G)})$ *embeds into* $\mathbf{Cone}(U_{\mathcal{F}}, d_{\mathcal{F}})$ *with* $O(1)$ *distortion, namely,*

$$c_{\mathbf{Cone}(U_{\mathcal{F}},d_{\mathcal{F}})}\left(\Sigma(G), d_{\Sigma(G)}\right) \leqslant \frac{\pi(R_{\mathcal{F}}+1)}{2}. \tag{47}$$

Proof. In Section 3.5 we verified that $(U_{\mathcal{F}}, d_{\mathcal{F}})$ is a $CAT(1)$ space, so the fact that $\mathbf{Cone}(U_{\mathcal{F}}, d_{\mathcal{F}})$ is a Hadamard space is a consequence of Berestovskiĭ's theorem. Fixing $G \in \mathcal{F}$, define $f : \Sigma(G) \to (0,\infty) \times U_{\mathcal{F}}$ by $f(x) = (1/\sqrt{2}, x)$. Recalling (37) and (43), for every $x, y \in \Sigma(G)$,

$$d_{\mathbf{Cone}(U_{\mathcal{F}},d_{\mathcal{F}})}(f(x), f(y)) = \sqrt{1 - \cos\left(\min\left\{\pi, \frac{2\pi d_{\Sigma(G)}(x,y)}{\mathrm{girth}(G)}\right\}\right)}.$$

Due to the elementary inequalities

$$\forall\, \theta \in [0,\pi], \qquad \frac{2\theta^2}{\pi^2} \leqslant 1 - \cos\theta \leqslant \frac{\theta^2}{2}, \tag{48}$$

it follows that

$$d_{\mathbf{Cone}(U_{\mathcal{F}},d_{\mathcal{F}})}(f(x), f(y)) \leqslant \frac{\pi d_{\Sigma(G)}(x,y)}{\sqrt{2}\,\mathrm{girth}(G)}, \tag{49}$$

and

$$\begin{aligned} d_{\mathbf{Cone}(U_{\mathcal{F}},d_{\mathcal{F}})}(f(x), f(y)) &\geqslant \frac{\sqrt{2}}{\pi}\cdot \min\left\{\pi, \frac{2\pi d_{\Sigma(G)}(x,y)}{\mathrm{girth}(G)}\right\} \\ &\geqslant \frac{\sqrt{2} d_{\Sigma(G)}(x,y)}{R_{\mathcal{F}}\,\mathrm{girth}(G)+1}, \end{aligned} \tag{50}$$

where the last step of (50) relies on the penultimate inequality in (38). The desired distortion estimate (47) follows from (49) and (50). $\square$

3.8. **Auxiliary families of graphs.** The family of graphs to which we will apply Lemma 3.9 is given in the following definition.

Definition 3.10. Fix $K \in (1,\infty)$ and denote by $\mathcal{F}_K$ all those connected graphs G with the following properties.

(1) $\mathrm{diam}(G) \leqslant K\,\mathrm{girth}(G)$.
(2) For every $S \subseteq \Sigma(G)$ with $|S| \leqslant \sqrt{|V_G|}$ we have

$$c_1\left(\mathbf{Cone}\left(S, \frac{2\pi}{\mathrm{girth}(G)} \cdot d_{\Sigma(G)}\right)\right) < K. \tag{51}$$

We also define

$$\mathscr{X}_K \stackrel{\mathrm{def}}{=} \mathbf{Cone}\left(U_{\mathcal{F}_K}, d_{\mathcal{F}_K}\right). \tag{52}$$

By Lemma 3.9 the metric space $\mathscr{X}_K$ is a Hadamard space and for every $G \in \mathcal{F}_K$ we have $c_{\mathscr{X}_K}(\Sigma(G), d_{\Sigma(G)}) \lesssim K$. Eventually K will be chosen to be a large enough universal constant.

Definition 3.11. Fix two integers $n, d \geqslant 3$ and $K \in (1,\infty)$. Denote by $\mathscr{L}_K^{n,d}$ the set of all those connected n-vertex d-regular graphs H for which there exists a subset of edges $I \subseteq E_H$ with $|I| \leqslant \sqrt{n}$ such that if we let L be the graph $(V_H, E_H \smallsetminus I)$ (i.e. we remove the edges in I from H) then L has the following properties.

(a) $L \in \mathcal{F}_K$.
(b) $\mathrm{diam}(L) \leqslant K \log_d n$.
(c) For every $S \subseteq V_L$ with $|S| \leqslant n/2$ we have

$$|E_L(S, V_L \smallsetminus S)| \geqslant \frac{|S|}{K}.$$

The following lemma asserts that a random n-vertex d-regular graph is asymptotically almost surely in $\mathscr{L}_K^{n,d}$. Due to item (a) of Definition 3.11, it follows in particular that for a large enough universal constant K, the graph family $\mathcal{F}_K$ contains many graphs.

Lemma 3.12. *There exists a universal constant $K \in (1,\infty)$ and for every integer $d \geqslant 3$ there exists $C(d) \in (0,\infty)$ such that for all $n \in \mathbb{N}$,*

$$\mathcal{G}_{n,d}\left(\mathscr{L}_K^{n,d}\right) \geqslant 1 - \frac{C(d)}{\sqrt[3]{n}}. \tag{53}$$

From now on we will assume that K is the universal constant from Lemma 3.12. We shall now proceed to establish some useful corollaries of Lemma 3.12; the proof of Lemma 3.12 itself will be given in Section 7.1.

Corollary 3.13. *There exists a universal constant $c \in (0,\infty)$ such that for every two integers $d,n \geqslant 3$ we have*

$$\mathcal{G}_{n,d}\left(\left\{H \in \mathbb{G}_n :\ \gamma\left(H, d_{\mathscr{X}_K}^2\right) \geqslant c(\log_d n)^2\right\}\right) \geqslant 1 - \frac{C(d)}{\sqrt[3]{n}},$$

where $K, C(d)$ are the constant from Lemma 3.12 and $\mathscr{X}_K$ is as in (52).

Proof. By Lemma 3.12, it suffices to show that there exists $c \in (0,\infty)$ such that if $H \in \mathscr{L}_K^{n,d}$ then $\gamma\left(H, d_{\mathscr{X}_K}^2\right) \geqslant c(\log_d n)^2$. So, supposing that $H \in \mathscr{L}_K^{n,d}$, let $I \subseteq E_H$ and $L = (V_H, E_H \smallsetminus I)$ be as in Definition 3.11.

By part (a) of Definition 3.11 we have $L \in \mathcal{F}_K$, so that by Lemma 3.9 we have $c_{\mathscr{X}_K}(V_H, d_L) \leqslant \pi K$. This means that there exists a mapping $f : V_H \to \mathscr{X}_K$ and a scaling factor $s \in (0,\infty)$ such that

$$\forall\, x,y \in V_H, \qquad s d_L(x,y) \leqslant d_{\mathscr{X}_K}\left(f(x), f(y)\right) \leqslant \pi K s d_L(x,y).$$

By the definition of $\gamma\left(H, d_{\mathscr{X}_K}^2\right)$, we therefore have

$$\frac{1}{n^2}\sum_{(x,y)\in V_H\times V_H} d_L(x,y)^2 \lesssim \frac{\gamma\left(H, d_{\mathscr{X}_K}^2\right)}{dn}\sum_{\{u,v\}\in E_H} d_L(x,y)^2. \tag{54}$$

Because L was obtained from H by deleting edges, we trivially have $d_L(x,y) \geqslant d_H(x,y)$ for every $x,y \in V_H$. Since H is a d-regular graph, for a universal constant fraction of the pairs $(x,y) \in V_H \times V_H$ we have $d_H(x,y) \gtrsim \log_d n$ (for a proof of this standard and simple fact, see e.g. page 193 of [Mat97]). Hence,

$$\frac{1}{n^2}\sum_{(x,y)\in V_H\times V_H} d_L(x,y)^2 \geqslant \frac{1}{n^2}\sum_{(x,y)\in V_H\times V_H} d_H(x,y)^2 \gtrsim (\log_d n)^2. \tag{55}$$

Since $E_L = E_H \smallsetminus I$, if $\{x,y\} \in E_H \smallsetminus I$ then $d_L(x,y) = 1$, and by part (b) of Definition 3.11 we have $d_L(x,y) \leqslant K \log_d n$ if $\{x,y\} \in I$. Hence,

$$\begin{aligned}\frac{1}{dn}\sum_{\{u,v\}\in E_H} d_L(x,y)^2 &\leqslant \frac{(|E_H| - |I|) + K^2(\log_d n)^2|I|}{dn}\\ &\lesssim \frac{dn + (\log_d n)^2\sqrt{n}}{dn} \lesssim 1, \end{aligned}\tag{56}$$

where we used the fact that $|I| \leqslant \sqrt{n}$, as stipulated in Definition 3.11. The desired lower bound $\gamma\left(H, d_{\mathscr{X}_K}^2\right) \gtrsim (\log_d n)^2$ now follows by contrasting (54) with (55) and (56). $\square$

Corollary 3.14. *Let K and $C(3)$ be the constants from Lemma 3.12 (with $d=3$). Then for every integer $n>C(3)^3$ there exists an n-vertex 3-regular graph Γ_n such that*
$$\sup_{n>C(3)^3} \lambda_2(\Gamma_n)<1, \tag{57}$$
and
$$\sup_{n>C(3)^3} c_{\mathscr{X}_K}\left(V_{\Gamma_n}, d_{\Gamma_n}\right)<\infty, \tag{58}$$
where $\mathscr{X}_K$ is as in (52).

Proof. Since the right hand side of (53) (with $d=3$) is positive, by Lemma 3.12 there exists an n-vertex 3-regular graph H such that $H\in\mathscr{L}_K^{n,d}$. Let $L=(V_H,E_H\smallsetminus I)$ be as in Definition 3.11. Since L was obtained from H by deleting edges, we can add self-loops to those vertices of V_H whose degree in L is less than 3 so that the resulting graph, which will be denoted Γ_n, is 3-regular. By part (c) of Definition (3.11), for every $S\subseteq V_{\Gamma_n}$ with $|S|\leqslant n/2$ we have $|E_{\Gamma_n}(S,V_{\Gamma_n}\smallsetminus S)|\gtrsim|S|$. By Cheeger's inequality [Che70, AM85], this implies (57). Also, $d_{\Gamma_n}=d_L$, so (58) follows from Lemma 3.9. □

3.9. **On the structure of cones over random graphs.** We have already described the tools that will suffice for the proof of Theorem 1.1, but in order to also prove Theorem 1.2 we need to have a better understanding of the structure of the Euclidean cone over a random graph.

Proposition 3.15. *For every integer $d\geqslant 3$ there exists $C'(d)\in(0,\infty)$ such that for every $n\in\mathbb{N}$, if H is distributed according to $\mathcal{G}_{n,d}$ then with probability at least $1-C'(d)/\sqrt[3]{n}$ the graph H is connected and there exist $\sigma\in(0,\infty)$ and two subsets $A_1,A_2\subseteq\Sigma(H)$ such that the following assertions hold true.*

(I) $A_1\cup A_2=\Sigma(H)$.
(II) $d_{\Sigma(H)}\left(A_1\smallsetminus A_2, A_2\smallsetminus A_1\right)\gtrsim\frac{1}{\sigma}$.
(III) $c_1\left(\mathbf{Cone}\left(A_1,\sigma d_{\Sigma(H)}\right)\right)\lesssim 1$.
(IV) $c_{\mathscr{X}_K}\left(\mathbf{Cone}\left(A_2,\sigma d_{\Sigma(H)}\right)\right)\lesssim 1$.
(V) $c_{\mathbf{Cone}\left(\Sigma(H),\sigma d_{\Sigma(H)}\right)}\left(\Sigma(H),d_{\Sigma(H)}\right)\lesssim 1$.

Here K is the constant from Lemma 3.12 and $\mathscr{X}_K$ is as in (52).

By naturally identifying the cusps of $\mathbf{Cone}(A_1)$ and $\mathbf{Cone}(A_2)$ with the cusp of $\mathbf{Cone}\left(\Sigma(H)\right)$, we have
$$\mathbf{Cone}\left(\Sigma(H)\right)=\mathbf{Cone}(A_1)\cup\mathbf{Cone}(A_2). \tag{59}$$
Proposition 3.15 therefore says that $\mathbf{Cone}\left(\Sigma(H),\sigma d_{\Sigma(H)}\right)$ can be written as a (non-disjoint) union of two subsets, the first of which embeds

bi-Lipschitzly into L_1 and the second of which embeds bi-Lipschitzly into the Hadamard space $\mathscr{X}_K$. The relevance of the remaining assertions of Proposition 3.15 will become clear in Section 3.10 below.

3.10. **Unions of cones.** Proposition 3.15 will be used to prove Theorem 1.2 via the following strategy. We will construct a sequence of 3-regular graphs $\{G_n\}_{n=1}^\infty$ satisfying (9) and such that

$$\sup_{n\in\mathbb{N}} \gamma_+\big(G_n, d_{\mathscr{X}_K}^2\big) < \infty.$$

By assertion (IV) of Proposition 3.15 we therefore have

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(A_2,\sigma d_{\Sigma(H)}\right)}\Big) < \infty. \tag{60}$$

Since $\{G_n\}_{n=1}^\infty$ are necessarily also classical expanders, we will argue using Matoušek's extrapolation [Mat97] that

$$\sup_{n\in\mathbb{N}} \gamma_+\big(G_n, d_{L_1}^2\big) < \infty,$$

where $d_{L_1}(f,g) \stackrel{\mathrm{def}}{=} \|f-g\|_1$ is the standard metric on L_1. Assertion (III) of Proposition 3.15 therefore implies that

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(A_1,\sigma d_{\Sigma(H)}\right)}\Big) < \infty. \tag{61}$$

We would like to combine (60) and (61) to deduce that $\{G_n\}_{n=1}^\infty$ are expanders with respect to $\mathbf{Cone}\left(\Sigma(H), \sigma d_{\Sigma(H)}\right)$. For this purpose, we prove the following lemma in Section 5.3.

Lemma 3.16. *Fix $\beta \in (0,\pi]$ and $n \in \mathbb{N}$. Let (X, d_X) be a metric space and suppose that $A, B \subseteq X$ satisfy $A \cup B = X$ and*

$$d_X(A \smallsetminus B, B \smallsetminus A) \geqslant \beta. \tag{62}$$

Then every n by n symmetric stochastic matrix $M = (m_{ij})$ satisfies

$$\gamma_+\big(M, d^2_{\mathbf{Cone}(X)}\big) \lesssim \frac{\gamma_+\big(M, d^2_{\mathbf{Cone}(A)}\big) + \gamma_+\big(M, d^2_{\mathbf{Cone}(B)}\big)}{\beta^4}. \tag{63}$$

By virtue of assertion (II) of Proposition 3.15, we may use Lemma 3.16 to deduce from (60) and (61) that

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(\Sigma(H),\sigma d_{\Sigma(H)}\right)}\Big) < \infty.$$

By assertion (V) of Proposition 3.15 it follows that

$$\sup_{n\in\mathbb{N}} \gamma_+\big(G_n, d^2_{\Sigma(H)}\big) < \infty,$$

implying Theorem 1.2. It remains, of course, to explain how to construct the graphs $\{G_n\}_{n=1}^{\infty}$; this is done in Section 4 below.

Remark 3.17. In light of (59), a positive answer to the following natural open question would yield an alternative (more general) way to carry out the above argument. Suppose that (X, d_X) is a metric space and $A, B \subseteq X$ satisfy $A \cup B = X$. Write d_A and d_B for the restriction of d_X to A and B, respectively. Let $\{G_n\}_{n=1}^{\infty}$ be a sequence of 3-regular graphs such that $\sup_{n\in\mathbb{N}} \gamma_+(G_n, d_A^2) < \infty$ and $\sup_{n\in\mathbb{N}} \gamma_+(G_n, d_B^2) < \infty$. Does this imply that $\sup_{n\in\mathbb{N}} \gamma_+(G_n, d_X^2) < \infty$? The same question could be also asked with $\gamma_+(\cdot,\cdot)$ replaced throughout by $\gamma(\cdot,\cdot)$. We speculate that the answer to the above questions is negative, in which case one would ask for conditions on A and B which would imply a positive answer. Lemma 3.16 is a step in this direction.

4. Proof of Theorem 1.1 and Theorem 1.2

Here we prove Theorem 1.1 and Theorem 1.2 while assuming only the validity of Lemma 3.12, Proposition 3.15 and Lemma 3.16, which will be proven in Section 7.1, Section 7.2 and Section 5.3, respectively. Proposition 3.7, whose proof appears in Section 5.1, serves as a step towards the proof of Proposition 3.15.

4.1. A stronger theorem.

We state below a theorem that directly implies both Theorem 1.1 and Theorem 1.2.

Theorem 4.1. *There exist a Hadamard space (X, d_X) and two sequences of connected 3-regular graphs $\{G_n\}_{n=1}^{\infty}$ and $\{\Gamma_n\}_{n=1}^{\infty}$ satisfying $\lim_{n\to\infty} |V_{G_n}| = \lim_{n\to\infty} |V_{\Gamma_n}| = \infty$ and $\sup_{n\in\mathbb{N}} |V_{G_{n+1}}|/|V_{G_n}| < \infty$, such that the following properties hold true.*

(1) $\sup_{n\in\mathbb{N}} \gamma_+(G_n, d_X^2) < \infty$.
(2) $\sup_{n\in\mathbb{N}} \lambda_2(\Gamma_n) < 1$.
(3) $\sup_{n\in\mathbb{N}} c_X(V_{\Gamma_n}, d_{\Gamma_n}) < \infty$.
(4) There exists $c, C \in (0,\infty)$ and for every integer $d \geqslant 1$ there exists $\kappa(d) \in (0,\infty)$ such that for every $n \in \mathbb{N}$ we have

$$\mathcal{G}_{n,d}\left(\left\{H \in \mathbb{G}_n :\ \gamma\big(H, d_X^2\big) \geqslant c(\log_d n)^2\right\}\right) \geqslant 1 - \frac{\kappa(d)}{\sqrt[3]{n}},$$

and

$$\mathcal{G}_{n,d}\left(\left\{H \in \mathbb{G}_n^{\mathrm{con}} :\ \sup_{k\in\mathbb{N}} \gamma_+\big(G_k, d_{\Sigma(H)}^2\big) < C\right\}\right) \geqslant 1 - \frac{\kappa(d)}{\sqrt[3]{n}}.$$

Theorem 4.1 is stronger than Theorem 1.1 and Theorem 1.2. Indeed, the estimate (6) in Theorem 1.1 is weaker than assertion *(1)* of

Theorem 4.1 because $\gamma_+(G_n, d_X^2) \geqslant \gamma(G_n, d_X^2)$. The estimate (7) in Theorem 1.1 and the estimate (10) in Theorem 1.2 follow from assertion *(4)* of Theorem 4.1.

The role of the graphs $\{\Gamma_n\}_{n=1}^\infty$ of Theorem 4.1 is to supply the following additional information that is not contained in Theorem 1.1 and Theorem 1.2. By assertions *(2)* and *(3)* of Theorem 4.1, we know that X contains bi-Lipschitz copies of some 3-regular (classical) expander sequence. Using (20) we deduce from assertion *(3)* of Theorem 4.1 $\gamma(\Gamma_n, d_X^2) \gtrsim (\log n)^2$. As was discussed in the paragraph that immediately follows the statement of Theorem 1.1, we already knew that as a consequence of Theorem 1.1 there exists a sequence of 3-regular graphs $\{H_n\}_{n=1}^\infty$ for which $\gamma(H_n, d_X^2) \gtrsim (\log n)^2$. Theorem 4.1 says that moreover, we can even arrange it for X to contain bi-Lipschitz copies of such an expander sequence.

4.2. **Proof of Theorem 4.1.** In the proof of Theorem 4.1 we will use the following easy lemma, which is in the spirit of Lemma 3.16 but simpler to prove.

Lemma 4.2. *Let (X, d_X) be a metric space and suppose that $\{A_i\}_{i\in I}$ are subsets of X such that*

$$X = \bigcup_{i\in I} A_i, \quad \text{and} \quad \forall\, i,j \in I,\ (i \neq j) \implies d_X(A_i, A_j) \geqslant \pi. \tag{64}$$

Then for every $n \in \mathbb{N}$, any n by n symmetric stochastic $M = (m_{ij})$ satisfies

$$\gamma_+\big(M, d^2_{\mathbf{Cone}(X)}\big) \lesssim 2 \sup_{i\in I} \gamma_+\big(M, d^2_{\mathbf{Cone}(A_i)}\big)\,, \tag{65}$$

and

$$\gamma\left(M, d^2_{\mathbf{Cone}(X)}\right) \lesssim 2 \sup_{i\in I} \gamma\left(M, d^2_{\mathbf{Cone}(A_i)}\right)\,, \tag{66}$$

Proof. Let $o \in \mathbf{Cone}(X)$ denote the equivalence class obtained by identifying all the points in $\{0\} \times X$, thus $d_{\mathbf{Cone}(X)}\big((s,x), o\big) = s$ for every $(s,x) \in (0,\infty) \times X$. With a slight abuse of notation we think of o as the cusp $\{\mathbf{Cone}(A_i)\}_{i\in I}$ as well (formally the cusps of $\{\mathbf{Cone}(A)\}_{i\in I}$ should be labeled differently, as the equivalence classes $\{\{0\} \times A_i\}_{i\in I}$, but dropping this notation will not cause confusion in what follows).

For every $i \in I$ define $\mathfrak{a}_i : \mathbf{Cone}(X) \to \mathbf{Cone}(A_i)$ by

$$\mathfrak{a}_i(s,x) \stackrel{\mathrm{def}}{=} \begin{cases} (s,x) & \text{if } x \in A_i \\ o & \text{otherwise.} \end{cases}$$

Due to (64) every $(s,x),(t,y)\in \mathbf{Cone}(X)$ satisfy

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)=\sum_{i\in I} d_{\mathbf{Cone}(A_i)}\big(\mathfrak{a}_i(s,x),\mathfrak{a}_i(t,y)\big), \tag{67}$$

where the sum in the right-hand side of (67) contains at most two nonzero terms. Consequently,

$$1\leqslant \frac{d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2}{\sum_{i\in I} d_{\mathbf{Cone}(A_i)}\big(\mathfrak{a}_i(s,x),\mathfrak{a}_i(t,y)\big)^2}\leqslant 2.$$

Every $(s_1,x_1),\ldots,(s_n,x_n),(t_1,y_1),\ldots,(t_n,y_n)\in \mathbf{Cone}(X)$ therefore satisfy

$$\begin{aligned}
&\frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n d_{\mathbf{Cone}(X)}\big((s_j,x_j),(t_k,y_k)\big)^2\\
&\leqslant 2\sum_{i\in I}\frac{1}{n^2}\sum_{j=1}^n\sum_{k=1}^n d_{\mathbf{Cone}(A_i)}\big(\mathfrak{a}_i(s_j,x_j),\mathfrak{a}_i(t_k,y_k)\big)^2\\
&\leqslant 2\sum_{i\in I}\frac{\gamma_+\big(M,d^2_{\mathbf{Cone}(A_i)}\big)}{n}\sum_{j=1}^n\sum_{k=1}^n m_{jk}d_{\mathbf{Cone}(A_i)}\big(\mathfrak{a}_i(s_j,x_j),\mathfrak{a}_i(t_k,y_k)\big)^2\\
&\leqslant \frac{2\sup_{i\in I}\gamma_+\big(M,d^2_{\mathbf{Cone}(A_i)}\big)}{n}\sum_{j=1}^n\sum_{k=1}^n m_{jk}d_{\mathbf{Cone}(X)}\big((s_j,x_j),(t_k,y_k)\big)^2.
\end{aligned}$$

This proves (65). The proof of (66) is analogous. $\square$

The following lemma will be used (twice) in the proof of Theorem 4.1. Its proof relies on Matoušek's extrapolation lemma for Poincaré inequalities [Mat97], combined with an idea from [NS11].

Lemma 4.3. *Fix $d,n\in\mathbb{N}$ and let G be a d-regular n-vertex graph with $\lambda_2(G)<1$. Then*

$$\gamma\big(G,d^2_{L_1}\big)\lesssim \frac{1}{\big(1-\lambda_2(G)\big)^2}.$$

Proof. Since $\gamma(G,d^2_{\mathbb{R}})=1/(1-\lambda_2(G))$, it follows from Matoušek's extrapolation lemma for Poincaré inequalities (see [Mat97, Prop. 3] for Matoušek's original version and [BLMN05, Lem. 5.5] for the form that we use here) that for every $f : V_G \to L_4$ we have

$$\frac{1}{|V_G|^2} \sum_{(u,v)\in V_G\times V_G} \|f(u)-f(v)\|_4^4 \lesssim \frac{1}{(1-\lambda_2(G))^2|E_G|} \sum_{\{u,v\}\in E_G} \|f(u)-f(v)\|_4^4. \tag{68}$$

The metric space $(L_1, \sqrt{d_{L_1}})$ admits an isometric embedding into L_2 (see e.g. [DL97]), and L_2 admits an isometric embedding into L_4 (see e.g. [Woj91]), so $(L_1, \sqrt{d_{L_1}})$ admits an isometric embedding into L_4. It therefore follows from (68) that

$$\frac{1}{|V_G|^2} \sum_{(u,v)\in V_G\times V_G} \|f(u)-f(v)\|_1^2 \lesssim \frac{1}{(1-\lambda_2(G))^2|E_G|} \sum_{\{u,v\}\in E_G} \|f(u)-f(v)\|_1^2. \qquad \square$$

Proof of Theorem 4.1. Let K be the constant from Lemma 3.12 and write $X = \mathscr{X}_K$. The existence of the sequence of 3-regular graphs $\{\Gamma_n\}_{n=1}^\infty$, together with assertions *(2)* and *(3)* of Theorem 4.1, is precisely Corollary 3.14. The first displayed equation in assertion *(4)* of Theorem 4.1 follows from Corollary 3.13.

Our goal is to construct $\{G_n\}_{n=1}^\infty$ using Theorem 3.2. To this end, since the assumption (27) holds true by virtue of the fact that (X, d_X) is a Hadamard space and Theorem 3.4, we need to prove the existence of a "base graph" that satisfies the assumption (29).

Below it will be convenient to give a name to the implicit universal constant in (42). Thus let $\alpha \in (1,\infty)$ be such that for every $CAT(0)$ metric space (Y, d_Y), every $m, n \in \mathbb{N}$ and every n by n symmetric and stochastic matrix M, we have

$$\gamma_+\left(\frac{1}{m}\sum_{t=0}^{m-1} M^t, d_Y^2\right) \leqslant \alpha \max\left\{1, \frac{\gamma_+(M, d_Y^2)}{m}\right\}. \tag{69}$$

Fix from now on any sequence $\{W_k\}_{k=1}^\infty$ of 3-regular graphs of even cardinality that forms an expander sequence, i.e., $|V_{W_k}|$ is even for every $k \in \mathbb{N}$, $\lim_{k\to\infty} |V_{W_k}| = \infty$ and $\sup_{k\in\mathbb{N}} \lambda_2(W_k) < 1$. By Lemma 4.3 we have $\sup_{k\in\mathbb{N}} \gamma(W_k, d_{L_1}^2) < \infty$. By Lemma 3.1 there exists a sequence of 3-regular graphs $\{W_k^*\}_{k=1}^\infty$ with

$$\forall\, k \in \mathbb{N}, \qquad \left|V_{W_k^*}\right| = 54\,|V_{W_k}|, \tag{70}$$

such that

$$\beta \stackrel{\mathrm{def}}{=} \sup_{k\in\mathbb{N}} \gamma_+\big(W_k^*, d_{L_1}^2\big) < \infty. \tag{71}$$

Choose the minimum $k \in \mathbb{N}$ such that

$$|V_{W_k^*}| \geqslant 3^{48\alpha\beta^2K^4}. \tag{72}$$

Note that since α, β, K are all universal constants, so are k and $|V_{W_k^*}|$.

Define a subfamily $\mathcal{F}_K^*$ of the graph family $\mathcal{F}_K$ by

$$\mathcal{F}_K^* \stackrel{\mathrm{def}}{=} \left\{G \in \mathcal{F}_K :\ |V_G| \geqslant 4|V_{W_k^*}|^2\right\}. \tag{73}$$

Fix $G \in \mathcal{F}_K^*$ and two mappings

$$\varphi, \psi : V_{W_k^*} \to \mathbf{Cone}\left(\Sigma(G), \frac{2\pi}{\mathrm{girth}(G)} \cdot d_{\Sigma(G)}\right). \tag{74}$$

Observe that

$$\left|\varphi\left(V_{W_k^*}\right) \cup \psi\left(V_{W_k^*}\right)\right| \leqslant 2\left|V_{W_k^*}\right| \stackrel{(73)}{\leqslant} \sqrt{|V_G|}.$$

Consequently, there exists $S \subseteq \Sigma(G)$ with $|S| \leqslant \sqrt{|V_G|}$ such that

$$\varphi\left(V_{W_k^*}\right) \cup \psi\left(V_{W_k^*}\right) \subseteq \mathbf{Cone}\left(S, \frac{2\pi}{\mathrm{girth}(G)} \cdot d_{\Sigma(G)}\right).$$

By assertion (2) of Definition 3.10, it follows that

$$c_1\left(\varphi\left(V_{W_k^*}\right) \cup \psi\left(V_{W_k^*}\right), d_{\mathbf{Cone}\left(\Sigma(G), \frac{2\pi}{\mathrm{girth}(G)}\cdot d_{\Sigma(G)}\right)}\right) \leqslant K. \tag{75}$$

Since (75) holds true for every φ, ψ as in (74), it follows from the definition of β in (71) that

$$\gamma_+\left(W_k^*, d^2_{\mathbf{Cone}\left(\Sigma(G), \frac{2\pi}{\mathrm{girth}(G)}\cdot d_{\Sigma(G)}\right)}\right) \leqslant \beta K^2. \tag{76}$$

Recalling the definitions (36) and (37), $U_{\mathcal{F}_K^*}$ is the disjoint union of $\{\Sigma(G)\}_{G\in\mathcal{F}_K^*}$, where the distances (in terms of $d_{\mathcal{F}_K^*}$, which is the same as the restriction to $U_{\mathcal{F}_K^*}$ of $d_{\mathcal{F}_K}$) between different sets in this disjoint union being at least 2π. We are therefore allowed to use Lemma 4.2, thus concluding from (76) that

$$\gamma_+\left(W_k^*, d^2_{\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)}\right) \leqslant 2\beta K^2 \stackrel{(72)}{\leqslant} \sqrt{\frac{1}{2\alpha}\left\lfloor\frac{\log\left|V_{W_k^*}\right|}{3\log 3}\right\rfloor}. \tag{77}$$

By the definition of α in (69), due to (77) (which corresponds to condition (29)) we can apply Theorem 3.2 to the $CAT(0)$ metric space $\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)$ and the "base graph" $H = W_k^*$, thus obtaining a sequence of 3-regular graphs $\{F_n\}_{n=1}^{\infty}$ with

$$\forall\, n \in \mathbb{N}, \qquad |V_{F_n}| = 81 \left|V_{W_k^*}\right|^n \overset{(70)}{=} 81 \cdot (54\, |V_{W_k}|)^n\,, \tag{78}$$

and

$$\sup_{n\in\mathbb{N}} \gamma_+\left(F_n, d^2_{\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)}\right) < \infty. \tag{79}$$

Finally, define $G_n \stackrel{\mathrm{def}}{=} F_n^*$, where F_n^* is obtained from F_n by Lemma 3.1 (we are allowed to apply Lemma 3.1 here since by (78) we know that $|V_{F_n}|$ is even). Then $|V_{G_n}| = 54\,|V_{F_n}|$ for every $n \in \mathbb{N}$. Moreover, by Lemma 3.1 for every $n \in \mathbb{N}$ we have

$$\begin{aligned}\sup_{n\in\mathbb{N}} \gamma_+\left(G_n, d^2_{\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)}\right) &\lesssim \sup_{n\in\mathbb{N}} \gamma\left(F_n, d^2_{\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)}\right)\\ &\leqslant \sup_{n\in\mathbb{N}} \gamma_+\left(F_n, d^2_{\mathbf{Cone}\left(U_{\mathcal{F}_K^*}, d_{\mathcal{F}_K^*}\right)}\right) < \infty. \end{aligned}\tag{80}$$

Due to (79) the graphs $\{F_n\}_{n=1}^\infty$ are necessarily also classical expanders, i.e., $\sup_{n\in\mathbb{N}} \lambda_2(F_n) < 1$. By another application of Lemma 4.3 we therefore conclude that $\sup_{n\in\mathbb{N}} \gamma(F_n, d^2_{L_1}) < \infty$, and by Lemma 3.1 this implies that

$$\sup_{n\in\mathbb{N}} \gamma_+(G_n, d^2_{L_1}) \lesssim \sup_{n\in\mathbb{N}} \gamma(F_n, d^2_{L_1}) < \infty. \tag{81}$$

If $H \in \mathcal{F}_K \smallsetminus \mathcal{F}_K^*$ then $|V_H| < 4|V_{W_k^*}|^2$. Hence, as explained in Section 3.2 we have

$$\sup_{H\in\mathcal{F}_K\smallsetminus\mathcal{F}_K^*} c_1(\Sigma(H)) < \infty.$$

By Corollary 3.8 we therefore have

$$\sup_{H\in\mathcal{F}_K\smallsetminus\mathcal{F}_K^*} c_1\left(\mathbf{Cone}\left(\Sigma(H), \frac{2\pi}{\mathrm{girth}(H)} \cdot d_{\Sigma(H)}\right)\right) < \infty. \tag{82}$$

In conjunction with (81), it follows from (82) that

$$\sup_{H\in\mathcal{F}_K\smallsetminus\mathcal{F}_K^*} \sup_{n\in\mathbb{N}} \gamma_+\left(G_n, d^2_{\mathbf{Cone}\left(\Sigma(H), \frac{2\pi}{\mathrm{girth}(H)}\cdot d_{\Sigma(H)}\right)}\right) < \infty. \tag{83}$$

Recalling the definitions (36) and (37), $U_{\mathcal{F}_K}$ is the disjoint union of $U_{\mathcal{F}_K^*}$ and $\{\Sigma(H)\}_{H\in\mathcal{F}_K\smallsetminus\mathcal{F}_K^*}$ with the distances (in terms of $d_{\mathcal{F}_K}$) between different sets in this disjoint union being at least 2π. We can therefore apply Lemma 4.2 to (80) and (83), yielding assertion *(1)* of Theorem 4.1.

It remains to justify the second displayed equation in assertion *(4)* of Theorem 4.1. This is done following the strategy that was sketched in Section 3.10. Let H be a connected 3-regular n-vertex graph that satisfies the conclusion of Proposition 3.15. By Proposition 3.15 we know that the $\mathcal{G}_{n,d}$-probability that this occurs is at least $1-C'(d)/\sqrt[3]{n}$. Continuing here with the notation of Proposition 3.15, it follows from assertion *(1)* of Theorem 4.1 (which we already proved), combined with assertion (IV) of Proposition 3.15 that

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(A_2,\sigma d_{\Sigma(H)}\right)}\Big) \lesssim 1. \tag{84}$$

Due to (81), assertion (III) of Proposition 3.15 implies that

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(A_1,\sigma d_{\Sigma(H)}\right)}\Big) \lesssim 1. \tag{85}$$

By assertions (I) and (II) of Proposition 3.15, we may use Lemma 3.16 to deduce from (84) and (85) that

$$\sup_{n\in\mathbb{N}} \gamma_+\Big(G_n, d^2_{\mathbf{Cone}\left(\Sigma(H),\sigma d_{\Sigma(H)}\right)}\Big) \lesssim 1.$$

By assertion (V) of Proposition 3.15 it follows that

$$\sup_{n\in\mathbb{N}} \gamma_+\big(G_n, d^2_{\Sigma(H)}\big) \lesssim 1. \qquad \square$$

5. On the Lipschitz structure of Euclidean cones

Here we study various aspects of the Lipschitz structure of Euclidean cones. In particular, we will prove Proposition 3.7 and Lemma 3.16. We will also establish several additional geometric results of independent interest, yielding as a side product progress on a question that we raised in [MN04]. Despite the fact that in this paper the value of universal constants is mostly insignificant, when proving results such as Proposition 3.7 we will attempt to state reasonably good (though still suboptimal) explicit distortion bounds, because such geometric results are interesting in their own right and might be useful elsewhere.

Before proceeding we record for future use the following simple facts.

Lemma 5.1. *Let (X, d_X) be a metric space. For every $s,t \in (0,\infty)$ and $x, y \in X$ we have*

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) \geqslant \max\{s,t\}\cdot \sin\left(\min\left\{\frac{\pi}{2}, d_X(x,y)\right\}\right). \tag{86}$$

Proof. If $d_X(x,y) \geqslant \pi/2$ then $\cos\left(\min\{\pi, d_X(x,y)\}\right) \leqslant 0$ and it follows from (43) that $d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) \geqslant \sqrt{s^2+t^2} \geqslant \max\{s,t\}$. We may therefore assume that $d_X(x,y) < \pi/2$ and $s \geqslant t$. The minimum of the function $t \mapsto s^2+t^2-2st\cos\left(d_X(x,y)\right)$ is attained at $t = s\cos\left(d_X(x,y)\right)$, implying (86). $\square$

Lemma 5.2. *Let (X,d_X) be a bounded metric space. Suppose that $f: X\to(0,\infty)$ is Lipschitz and define $F:\mathbf{Cone}(X)\to\mathbf{Cone}(X)$ by*

$$F(s,x)\stackrel{\mathrm{def}}{=}\big(f(x)s,x\big).$$

Then F is Lipschitz with

$$\|F\|_{\mathrm{Lip}}\leqslant\sqrt{\operatorname{diam}(X)^2\cdot\|f\|_{\mathrm{Lip}}^2+2\|f\|_\infty^2}.$$

Proof. Fix $x,y\in X$ and set $\theta\stackrel{\mathrm{def}}{=}\min\{\pi,d_X(x,y)\}$. For $s,t\in(0,\infty)$ with $s\leqslant t$ we have

$$\begin{aligned}|f(x)s-f(y)t|&\leqslant|f(x)-f(y)|s+|f(y)|\cdot|s-t|\\&\leqslant\|f\|_{\mathrm{Lip}}d_X(x,y)s+\|f\|_\infty|s-t|\\&\leqslant\frac{\operatorname{diam}(X)\|f\|_{\mathrm{Lip}}}{\pi}\theta\sqrt{st}+\|f\|_\infty|s-t|\\&\leqslant\frac{\operatorname{diam}(X)\|f\|_{\mathrm{Lip}}}{\sqrt{2}}\sqrt{st(1-\cos\theta)}+\|f\|_\infty|s-t|,\end{aligned}\tag{87}$$

where in (87) we used (48). By squaring (87) we see that

$$\big(f(x)s-f(y)t\big)^2\leqslant2\|f\|_\infty^2(s-t)^2+\operatorname{diam}(X)^2\|f\|_{\mathrm{Lip}}^2st(1-\cos\theta),\tag{88}$$

and therefore,

$$\begin{aligned}&d_{\mathbf{Cone}(X)}\big(F(s,x),F(t,y)\big)^2\\&\stackrel{(44)}{=}\big(f(x)s-f(y)t\big)^2+2f(x)f(y)st(1-\cos\theta)\\&\stackrel{(88)}{\leqslant}2\|f\|_\infty^2(s-t)^2+\big(\operatorname{diam}(X)^2\|f\|_{\mathrm{Lip}}^2+2\|f\|_\infty^2\big)st(1-\cos\theta)\\&\stackrel{(44)}{\leqslant}\big(\operatorname{diam}(X)^2\|f\|_{\mathrm{Lip}}^2+2\|f\|_\infty^2\big)d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2.\end{aligned}$$ □

Lemma 5.3. *Let (X,d_X) be a metric space and suppose that (s,x) and (t,y) are distinct points in $\mathbf{Cone}(X)$. Then*

$$\frac13\leqslant\frac{d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)}{\max\Big\{|s-t|,\max\{s,t\}\cdot\sqrt{2\left(1-\cos\left(\min\{\pi,d_X(x,y)\}\right)\right)}\Big\}}\leqslant\sqrt{2}.$$

Proof. Denoting $\theta\stackrel{\mathrm{def}}{=}\min\{\pi,d_X(x,y)\}$, it follows from (44) that

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)=\sqrt{(s-t)^2+2st(1-\cos\theta)}.\tag{89}$$

The desired upper bound on $d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)$ is therefore immediate from (89), using $\sqrt{st}\leqslant\max\{s,t\}$. Since $\sqrt{st}\geqslant\max\{s,t\}-|s-t|$, it follows from (89) that

$$\begin{aligned}
&d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)\\
&\geqslant \max\left\{|s-t|, \max\{s,t\}\cdot\sqrt{2(1-\cos\theta)} - |s-t|\sqrt{2(1-\cos\theta)}\right\}\\
&\geqslant \frac{\max\left\{|s-t|, \max\{s,t\}\cdot\sqrt{2(1-\cos\theta)}\right\}}{1+\sqrt{2(1-\cos\theta)}},
\end{aligned}$$

yielding the desired lower bound on $d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)$. □

5.1. **The Euclidean cone over L_1.** Our goal here is to prove Proposition 3.7. In preparation for this we first show that if one equips L_1 with the metric $\rho(x,y)=\min\{1,\|x-y\|_1\}$, then the resulting metric space admits a bi-Lipschitz embedding into L_1. In [MN04] we showed that the metric space $(L_2,\min\{1,\|x-y\|_2\})$ admits a bi-Lipschitz embedding into L_2 but left open the question whether $c_p(L_p,\min\{1,\|x-y\|_p\})<\infty$ for $p\in[1,2)$; Lemma 5.4 below shows that this is indeed the case when $p=1$, but the question remains open for $p\in(1,2)$. As explained in [MN04, Rem. 5.12], for every $p\in(2,\infty]$ the metric space $(L_p,\min\{1,\|x-y\|_p\})$ does not admit a bi-Lipschitz embedding into L_q for any $q\in[1,\infty)$.

Lemma 5.4 (truncated L_1 embeds into L_1)**.** *For every $M\in(0,\infty)$ there exists a mapping $T_M:L_1\to L_1$ such that*

- *$\|T_M(x)\|_1=M$ for every $x\in L_1$,*
- *For every distinct $x,y\in L_1$,*

$$1-\frac{1}{e}\leqslant \frac{\|T_M(x)-T_M(y)\|_1}{\min\{M,\|x-y\|_1\}}\leqslant 1. \tag{90}$$

Proof. Fix an integer $n\geqslant 2$. For every $A\subseteq\{1,\ldots,n\}$ the Walsh function $W_A:\{0,1\}^n\to\{-1,1\}$ is defined as usual by

$$\forall\, z=(z_1,\ldots,z_n)\in\{0,1\}^n,\qquad W_A(z)\stackrel{\mathrm{def}}{=}\prod_{j\in A}(-1)^{z_j}.$$

For $\lambda\in[1/2,1]$ observe that every $x,y\in\{0,1\}^n$ satisfy

$$\begin{aligned}
1-(2\lambda-1)^{\|x-y\|_1} &= 1-\prod_{j=1}^n\big(\lambda+(-1)^{x_j+y_j}(1-\lambda)\big)\\
&=\sum_{A\subseteq\{1,\ldots,n\}}\lambda^{n-|A|}(1-\lambda)^{|A|}\left(1-W_A(x+y)\right)\\
&=\sum_{A\subseteq\{1,\ldots,n\}}\lambda^{n-|A|}(1-\lambda)^{|A|}\left|W_A(x)-W_A(y)\right|.
\end{aligned}\tag{91}$$

Let $\{h_A\}_{A\subseteq\{1,\ldots,n\}} \subseteq L_1$ be disjointly supported functions with unit L_1 norm, and define $\Phi_M^n : \{0,1\}^n \to L_1$ by setting for every $z\in\{0,1\}^n$,

$$\Phi_M^n(z) \stackrel{\mathrm{def}}{=} M \sum_{A\subseteq\{1,\ldots,n\}} \left(\frac{1+e^{-1/M}}{2}\right)^{n-|A|}\left(\frac{1-e^{-1/M}}{2}\right)^{|A|} W_A(z)h_A.$$

Then $\|\Phi_M^n(z)\|_1 = M$ for every $z\in\{0,1\}^n$, and it follows from (91) that for every $x,y\in\{0,1\}^n$ we have

$$\frac{\|\Phi_M^n(x)-\Phi_M^n(y)\|_1}{\min\{M,\|x-y\|_1\}} = \frac{M\left(1-e^{-\|x-y\|_1/M}\right)}{\min\{M,\|x-y\|_1\}} \in \left[1-\frac{1}{e},1\right]. \qquad (92)$$

This proves the existence of the desired embedding for the metric space $(\{0,1\}^n,\min\{M,\|\cdot\|_1\})$; we pass to all of L_1 via the following standard approximation argument. If $S\subseteq L_1$ is finite and $\varepsilon\in(0,1)$ then by approximating by step functions we can choose $d,K,Q\in\mathbb{N}$ and $k_1,\ldots k_d : S\to\{1,\ldots,K\}$ such that for every $x,y\in S$,

$$\frac{1}{Q}\sum_{j=1}^{d}|k_i(x)-k_i(y)| \leqslant \|x-y\|_1 \leqslant \frac{1+\varepsilon}{Q}\sum_{j=1}^{d}|k_i(x)-k_i(y)|. \qquad (93)$$

Write $n=dK$ and define $\psi : S\to\{0,1\}^n$ by

$$\forall\, x\in S, \qquad \psi(x) \stackrel{\mathrm{def}}{=} \sum_{i=1}^{d}\sum_{j=(i-1)K+1}^{(i-1)K+k_i(x)} e_j,$$

where $e_1,\ldots,e_n$ is the standard basis of $\mathbb{R}^n$. Then by (93),

$$\forall\, x,y\in S, \qquad \frac{Q}{1+\varepsilon}\|x-y\|_1 \leqslant \|\psi(x)-\psi(y)\|_1 \leqslant Q\|x-y\|_1. \quad (94)$$

Define $F_\varepsilon : S\to L_1$ by

$$F_\varepsilon \stackrel{\mathrm{def}}{=} \frac{1}{Q}\Phi_{QM}^n\circ\psi.$$

Then $\|F_\varepsilon(x)\|_1 = M$ for every $x\in S$, and by (92) and (94) we have

$$\frac{1}{1+\varepsilon}\left(1-\frac{1}{e}\right) \leqslant \frac{\|F_\varepsilon(x)-F_\varepsilon(y)\|_1}{\min\{M,\|x-y\|_1\}} \leqslant 1$$

for all distinct $x,y\in S$. Since this holds for every finite $S\subseteq L_1$ and every $\varepsilon\in(0,1)$, the existence of the mapping T_M from the statement of Lemma 5.4 now follows from a standard ultrapower argument (see [Kap09, Ch. 9] for background on ultrapowers of metric spaces), using the fact that any ultrapower of L_1 is an $L_1(\mu)$ space [Hei80]. □

Remark 5.5. As explained in [MN04, Rem. 5.4], it follows formally from Lemma 5.4 that there exists a universal constant $C \in (0,\infty)$ such that if $\omega : [0,\infty) \to [0,\infty)$ is a concave nondecreasing function with $\omega(0) = 0$ and $\omega(t) > 0$ for $t > 0$ then the metric space $(L_1, \omega(\|x-y\|_1))$ embeds into L_1 with distortion at most C. The analogous assertion with L_1 replaced by L_p for $p \in (1,2)$ remains open.

Proof of Proposition 3.7. Let $T_\pi : L_1 \to L_1$ be the mapping constructed in Lemma 5.4 (with $M = \pi$). Define $F : \mathbf{Cone}(L_1) \to L_1$ by

$$\forall (s,x) \in (0,\infty) \times L_1, \qquad F(s,x) \stackrel{\mathrm{def}}{=} sT_\pi(x). \tag{95}$$

Fix $(s,x),(t,y) \in (0,\infty) \times L_1$ and assume without loss of generality that $t \geqslant s$. Recalling (44), we have

$$d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big) = \sqrt{(s-t)^2 + 2st(1-\cos\theta)}, \tag{96}$$

where

$$\theta \stackrel{\mathrm{def}}{=} \min\{\pi, \|x-y\|_1\}. \tag{97}$$

Moreover,

$$\|T_\pi(x)\|_1 = \|T_\pi(y)\|_1 = \pi \tag{98}$$

and (90) becomes

$$\left(1 - \frac{1}{e}\right)\theta \leqslant \|T_\pi(x) - T_\pi(y)\|_1 \leqslant \theta. \tag{99}$$

Using also (48), we therefore have

$$\begin{align}
\|F(s,x) - F(t,y)\|_1 &= \|sT_\pi(x) - tT_\pi(y)\|_1 \tag{100}\\
&\leqslant s\|T_\pi(x) - T_\pi(y)\|_1 + (t-s)\|T_\pi(y)\|_1 \tag{101}\\
&\leqslant s\theta + \pi(t-s) \tag{102}\\
&\leqslant \frac{\pi}{2}\sqrt{2st(1-\cos\theta)} + \pi(t-s) \tag{103}\\
&\leqslant \frac{\pi\sqrt{5}}{2}\sqrt{2st(1-\cos\theta) + (s-t)^2} \tag{104}\\
&= \frac{\pi\sqrt{5}}{2} d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big), \tag{105}
\end{align}$$

where in (100) we used (95), in (101) we used the triangle inequality in L_1, in (102) we used (98) and the rightmost inequality in (99), in (103) we used the leftmost inequality in (48) and that $s \leqslant \sqrt{st}$, in (104) we used the Cauchy-Schwarz inequality, and in (105) we used (96).

Next, we claim that

$$\|F(s,x)-F(t,y)\|_1 \geqslant \frac{\pi(e-1)}{\sqrt{4\pi^2e^2+(e-1)^2}} d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big). \tag{106}$$

Once proved, (106) in conjunction with (105) would imply that

$$c_1\big(\mathbf{Cone}(L_1)\big) \leqslant \frac{\sqrt{20\pi^2e^2+5(e-1)^2}}{2(e-1)} \leqslant 11.17, \tag{107}$$

thus completing the proof of Proposition 3.7.

We prove (106) by distinguishing between two cases as follows.

Case 1: $t-s \geqslant \frac{e-1}{2\pi e} t\theta$. Since $t \geqslant \sqrt{st}$ this assumption combined with the rightmost inequality appearing in (48) implies that

$$t-s \geqslant \frac{e-1}{2\pi e}\sqrt{2st(1-\cos\theta)}. \tag{108}$$

Consequently,

$$d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big) \overset{(96)\wedge(108)}{\leqslant} (t-s)\frac{\sqrt{4\pi^2e^2+(e-1)^2}}{e-1}, \tag{109}$$

implying the desired inequality (106) as follows.

$$\begin{aligned}
\|F(s,x)-F(t,y)\|_1 &\overset{(95)}{=} \|sT_\pi(x)-tT_\pi(y)\|_1\\
&\geqslant t\|T_\pi(y)\|_1 - s\|T_\pi(x)\|_1\\
&\overset{(98)}{=} \pi(t-s)\\
&\overset{(109)}{\geqslant} \frac{\pi(e-1)}{\sqrt{4\pi^2e^2+(e-1)^2}} d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big).
\end{aligned}$$

Case 2: we have

$$t-s < \frac{e-1}{2\pi e} t\theta. \tag{110}$$

It follows that

$$\begin{aligned}
\|F(s,x)-F(t,y)\|_1 &\overset{(95)}{=} \|sT_\pi(x)-tT_\pi(y)\|_1\\
&\geqslant t\|T_\pi(y)-T_\pi(x)\|_1 - (t-s)\|T_\pi(x)\|_1\\
&\overset{(99)\wedge(98)}{\geqslant} \left(1-\frac{1}{e}\right)t\theta - \pi(t-s)\\
&\overset{(110)}{>} \frac{e-1}{2e}t\theta.
\end{aligned} \tag{111}$$

At the same time, since $st \leqslant t^2$, it follows from (96), combined with (110) and the rightmost inequality appearing in (48), that

$$d_{\mathbf{Cone}(L_1)}\big((s,x),(t,y)\big) \leqslant t\theta\frac{\sqrt{4\pi^2e^2+(e-1)^2}}{2\pi e}. \tag{112}$$

In combination with (111), it follows from (112) that the desired inequality (106) holds true in Case 2 as well, thus completing the proof of Proposition 3.7. □

5.2. **Snowflakes of cones.** For $\alpha \in (0,1]$, the α-snowflake of a metric space (X,d) is defined to be the metric space (X,d^α). It is well known that the $\frac12$-snowflake of L_1 embeds isometrically into L_2 (see e.g. [DL97] or [Nao10, Sec. 3]). Consequently, Proposition 3.7 implies that the $\frac12$-snowflake of $\mathbf{Cone}(L_1)$ admits a bi-Lipschitz distortion into L_2. Proposition 5.6 below yields an alternative proof of this fact. Note that Proposition 5.6 does not imply Proposition 3.7 since by [KV05] there exist metric spaces which do not admit a bi-Lipschitz embedding into L_1 yet their $\frac12$-snowflake admits an isometric embedding into L_2.

In the proof of Theorem 4.1 (most significantly, in the proofs of Lemma 3.12 and Proposition 3.15) we use Proposition 3.7, but one can use mutatis mutandis Proposition 5.6 instead. We include Proposition 5.6 here due to its intrinsic interest, but it will not be used in this paper other than as an indication of an alternative approach to the parts of the proof of Theorem 4.1 eluded to above (note, however, that the approach below yields a better bound; see Remark 5.7). Readers who familiarized themselves with the contents of Section 5.1 can skip the present section if they only wish to understand the proof of Theorem 4.1.

Proposition 5.6. *For every separable metric space (X,d_X) and every $\alpha\in(0,1]$ we have*

$$c_2\left(\mathbf{Cone}(X), d^\alpha_{\mathbf{Cone}(X)}\right) \lesssim c_2\left(X, d_X^\alpha\right). \tag{113}$$

Proof. By a classical theorem of Schoenberg [Sch38] (see also [WW75]) there exists a mapping $h_\alpha:\mathbb{R}\to L_2$ with $h_\alpha(0)=0$ such that

$$\forall\, s,t\in\mathbb{R},\qquad \|h_\alpha(s)-h_\alpha(t)\|_2 = |s-t|^\alpha. \tag{114}$$

Also, as shown in [MN04, Lem. 5.2], there exists $\tau_\alpha: L_2\to L_2$ such that every distinct $x,y\in L_2$ satisfy

$$\|\tau_\alpha(x)\|_2 = \|\tau_\alpha(y)\|_2 = \frac{\pi^\alpha}{\sqrt{2}}, \tag{115}$$

and

$$\sqrt{1-\frac{1}{e}} \leqslant \frac{\|\tau_\alpha(x)-\tau_\alpha(y)\|_2}{\min\{\pi^\alpha, \|x-y\|_2\}} \leqslant 1. \tag{116}$$

Writing $D \stackrel{\mathrm{def}}{=} c_2\left(X, d_X^\alpha\right)$, there exists $f : X \to L_2$ such that

$$\forall\, x,y \in X \qquad \frac{d_X(x,y)^\alpha}{D} \leqslant \|f(x)-f(y)\|_2 \leqslant d_X(x,y)^\alpha. \tag{117}$$

We can now define an embedding $\phi : \mathbf{Cone}(X) \to L_2 \otimes L_2$ by

$$\forall (s,x) \in (0,\infty) \times X, \qquad \phi(s,x) \stackrel{\mathrm{def}}{=} h_\alpha(s) \otimes \tau_\alpha(f(x)). \tag{118}$$

Here the tensor product $L_2 \otimes L_2$ is equipped with the Hilbert space tensor product (Hilbert-Schmidt) norm. Thus $L_2 \otimes L_2$ is isometric to L_2 and $\|x \otimes y\|_{L_2 \otimes L_2} = \|x\|_2 \cdot \|y\|_2$ for every $x, y \in L_2$. Consequently, for every $a, b, x, y \in L_2$ we have

$$\begin{aligned}\|a\otimes x - b \otimes y\|_{L_2\otimes L_2}^2 &= \|a\|_2^2 \cdot \|x\|_2^2 + \|b\|_2^2 \cdot \|y\|_2^2 \\ &\quad - \frac{1}{2}\left(\|a\|_2^2 + \|b\|_2^2 - \|a-b\|_2^2\right)\left(\|x\|_2^2 + \|y\|_2^2 - \|x-y\|_2^2\right). \end{aligned}\tag{119}$$

Fix $x, y \in X$ and $s, t \in (0,\infty)$ with $t \geqslant s$. Since $h_\alpha(0)=0$ it follows from (114) that $\|h_\alpha(s)\|_2 = s^\alpha$ and $\|h_\alpha(t)\|_2 = t^\alpha$. In conjunction with (115), this shows that due to (114), (118) and (119),

$$\begin{aligned}&\|\phi(s,x)-\phi(t,y)\|_2^2 \\ &= \frac{\pi^{2\alpha}}{2}\left(s^{2\alpha}+t^{2\alpha}\right) \\ &\qquad - \frac{1}{2}\left(s^{2\alpha}+t^{2\alpha}-(t-s)^{2\alpha}\right)\left(\pi^{2\alpha} - \|\tau_\alpha(f(x))-\tau_\alpha(f(y))\|_2^2\right) \\ &= \frac{\pi^{2\alpha}}{2}(t-s)^{2\alpha} + \frac{s^{2\alpha}+t^{2\alpha}-(t-s)^{2\alpha}}{2}\,\|\tau_\alpha(f(x))-\tau_\alpha(f(y))\|_2^2 .\end{aligned}\tag{120}$$

Let $\theta \in [0,\pi]$ be given by

$$\theta \stackrel{\mathrm{def}}{=} \min\{\pi, d_X(x,y)\}. \tag{121}$$

Recalling (44), we therefore have

$$d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) = \sqrt{(s-t)^2 + 2st(1-\cos\theta)}, \tag{122}$$

and

$$\|\tau_\alpha(f(x))-\tau_\alpha(f(y))\|_2 \overset{(116)\wedge(117)}{\leqslant} \theta^\alpha \overset{(48)}{\leqslant} \frac{\pi^\alpha}{2^{\alpha/2}}(1-\cos\theta)^{\alpha/2}. \tag{123}$$

To bound (120) from above we also note the following elementary inequality, which holds for every $s \in [0,\infty)$, $t \in [s,\infty)$ and $\alpha \in [0,1]$.

$$s^{2\alpha}+t^{2\alpha}-(t-s)^{2\alpha} \leqslant 2(st)^\alpha. \tag{124}$$

To verify the validity of (124), normalization by $t^{2\alpha}$ shows that it suffices to check that $\psi(\alpha) = 2u^\alpha - u^{2\alpha} - 1 + (1-u)^{2\alpha} \geqslant 0$ for every fixed $u \in (0,1)$. Indeed, $\psi'(\alpha) = 2u^\alpha(1-u^\alpha)\log u + 2(1-u)^{2\alpha}\log(1-u) \leqslant 0$, so $\psi(\alpha) \geqslant \psi(1) = 0$.

By substituting (123) and (124) into (120) we have

$$\begin{aligned}\|\phi(s,x)-\phi(t,y)\|_2 &\leqslant \frac{\pi^\alpha}{2^{\alpha+1/2}}\sqrt{2^{2\alpha}|s-t|^{2\alpha}+(2st(1-\cos\theta))^\alpha}\\ &\leqslant \frac{\pi^\alpha\left(2^{2\alpha/(1-\alpha)}+1\right)^{1-\alpha}}{2^{\alpha+1/2}} d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^\alpha\\ &\lesssim d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^\alpha, \qquad (125)\end{aligned}$$

where the penultimate step of (125) uses Hölder's inequality and (122).

To bound (120) from below, observe first that

$$\begin{aligned}\|\tau_\alpha(f(x))-\tau_\alpha(f(y))\|_2 &\overset{(116)\wedge(117)}{\geqslant} \sqrt{1-\frac{1}{e}}\cdot\min\left\{\pi^\alpha,\frac{d_X(x,y)^\alpha}{D}\right\}\\ &\overset{(121)}{\geqslant} \frac{\sqrt{e-1}}{D\sqrt{e}}\theta^\alpha \overset{(48)}{\geqslant} \frac{2^{\alpha/2}\sqrt{e-1}}{D\sqrt{e}}(1-\cos\theta)^{\alpha/2}. \qquad (126)\end{aligned}$$

Substituting (126) and the trivial estimate $s^{2\alpha}+t^{2\alpha}-(t-s)^{2\alpha} \geqslant s^{2\alpha}$ into (120) shows that

$$\|\phi(s,x)-\phi(t,y)\|_2 \geqslant \frac{\sqrt{\frac{\pi^{2\alpha}eD^2}{e-1}|s-t|^{2\alpha}+(2s^2(1-\cos\theta))^\alpha}}{D\left(\frac{2e}{e-1}\right)^{\alpha/2}}. \qquad (127)$$

But, recalling (122), since $st \leqslant 2s^2+(s-t)^2$ we have

$$\begin{aligned}d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^\alpha &\leqslant \big(3(s-t)^2+2s^2(1-\cos\theta)\big)^{\alpha/2}\\ \leqslant \sqrt{3^\alpha|s-t|^{2\alpha}+(2s^2(1-\cos\theta))^\alpha} &\overset{(127)}{\lesssim} D\|\phi(s,x)-\phi(t,y)\|_2. \qquad (128)\end{aligned}$$

Due to (125) and (128) the proof of Proposition 5.6 is complete. □

Remark 5.7. When $\alpha = \frac{1}{2}$ and $D = 1$, a more careful analysis of the proof of Proposition 5.6 yields the estimate

$$c_2\left(\mathbf{Cone}(L_1), d_{\mathbf{Cone}(L_1)}^{1/2}\right) \leqslant 2.574.$$

We omit the (tedious) proof of this estimate, noting that it is better than the bound that follows from an application of Proposition 3.7 (specifically, recall (107)) and the fact that the $\frac{1}{2}$-snowflake of L_1 embeds isometrically into L_2.

5.3. **Poincaré inequalities with respect to unions of cones.** Our goal here is to prove Lemma 3.16. Before doing so, we record the following simple observation.

Lemma 5.8. *Fix $\lambda,\kappa\in(0,\infty)$ and $n\in\mathbb{N}$. Let (A,d_A), (B,d_B) and (X,d_X) be metric spaces. Suppose that there exist λ-Lipschitz mappings $\mathfrak{a}:X\to A$ and $\mathfrak{b}:X\to B$ such that for every $x,y\in X$,*

$$\sqrt{d_A(\mathfrak{a}(x),\mathfrak{a}(y))^2+d_B(\mathfrak{b}(x),\mathfrak{b}(y))^2}\geqslant\frac{d_X(x,y)}{\kappa}. \tag{129}$$

Then every n by n symmetric stochastic matrix $M=(m_{ij})$ satisfies

$$\gamma_+\big(M,d_X^2\big)\leqslant(\kappa\lambda)^2\left(\gamma_+\big(M,d_A^2\big)+\gamma_+\big(M,d_B^2\big)\right). \tag{130}$$

Proof. For every $x_1,\ldots,x_n,y_1,\ldots,y_n\in X$ we have

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_X(x_i,y_j)^2&\\ &\stackrel{(129)}{\leqslant}\frac{\kappa^2}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(d_A(\mathfrak{a}(x_i),\mathfrak{a}(y_j))^2+d_B(\mathfrak{b}(x_i),\mathfrak{b}(y_j))^2\right).\end{aligned} \tag{131}$$

By the definition of $\gamma_+(M,d_A^2)$ and $\gamma_+(M,d_B^2)$, combined with the fact that $\mathfrak{a}$ and $\mathfrak{b}$ are λ-Lipschitz,

$$\begin{aligned}&\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(d_A(\mathfrak{a}(x_i),\mathfrak{a}(y_j))^2+d_B(\mathfrak{b}(x_i),\mathfrak{b}(y_j))^2\right)\\ &\leqslant\frac{\gamma_+(M,d_A^2)}{n}\sum_{i=1}^n\sum_{j=1}^n m_{ij}d_A(\mathfrak{a}(x_i),\mathfrak{a}(y_j))^2\\ &\qquad+\frac{\gamma_+(M,d_B^2)}{n}\sum_{i=1}^n\sum_{j=1}^n m_{ij}d_B(\mathfrak{b}(x_i),\mathfrak{b}(y_j))^2\\ &\leqslant\frac{\lambda^2\left(\gamma_+(M,d_A^2)+\gamma_+(M,d_B^2)\right)}{n}\sum_{i=1}^n\sum_{j=1}^n m_{ij}d_X(x_i,y_j)^2.\end{aligned}$$

In combination with (131), this completes the proof of Lemma 5.8. □

Remark 5.9. The identical argument shows that (130) also holds true with the quantities $\gamma_+(M,d_X^2),\gamma_+(M,d_A^2),\gamma_+(M,d_B^2)$ replaced by the quantities $\gamma\,(M,d_X^2),\gamma\,(M,d_A^2),\gamma\,(M,d_B^2)$, respectively.

Proof of Lemma 3.16. Since the Euclidean cone over (X,d_X) is isometric to the Euclidean cone over the metric space $(X,\min\{\pi,d_X\})$, it suffices to prove (63) under the additional assumption

$$\operatorname{diam}(X)\leqslant\pi. \tag{132}$$

Let o denote the cusp of $\mathbf{Cone}(X)$, i.e, the equivalence class obtained by identifying all of the points in $\{0\}\times X$. Thus $d_{\mathbf{Cone}(X)}\big((s,x),o\big)=s$ for every $(s,x)\in(0,\infty)\times X$. Define $\mathfrak{a}:\mathbf{Cone}(X)\to\mathbf{Cone}(A)$ and $\mathfrak{b}:\mathbf{Cone}(X)\to\mathbf{Cone}(B)$ by setting for every $(s,x)\in(0,\infty)\times X$,

$$\mathfrak{a}(s,x)\stackrel{\mathrm{def}}{=}\big(d_X(x,B\smallsetminus A)s,x\big)\quad\text{and}\quad \mathfrak{b}(s,x)\stackrel{\mathrm{def}}{=}\big(d_X(x,A\smallsetminus B)s,x\big).$$

We first show that $\mathfrak{a}$ and $\mathfrak{b}$ are $\sqrt{3}\pi$-Lipschitz. By symmetry it suffices check this for $\mathfrak{a}$, i.e., to show that for every $s,t\in(0,\infty)$ and $x,y\in X$,

$$d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x),\mathfrak{a}(t,y)\big)\leqslant\sqrt{3}\pi d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big). \tag{133}$$

(133) is trivial if $d_X(x,B\smallsetminus A)=d_X(y,B\smallsetminus A)=0$. If $d_X(x,B\smallsetminus A)>0$ and $d_X(y,B\smallsetminus A)=0$ then

$$\begin{aligned}
&d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x),\mathfrak{a}(t,y)\big)\\
&=d_{\mathbf{Cone}(A)}\Big(\big(d_X(x,B\smallsetminus A)s,x\big),o\Big)\\
&=d_X(x,B\smallsetminus A)s\\
&\leqslant d_X(x,y)s\\
&\leqslant\frac{d_X(x,y)}{\sin\big(\min\left\{\frac{\pi}{2},d_X(x,y)\right\}\big)}d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big) &(134)\\
&\leqslant\pi d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big), &(135)
\end{aligned}$$

where (134) uses Lemma 5.1 and (135) uses (132) and the fact that the function $u\mapsto u/\sin u$ is increasing on $(0,\pi/2]$. The validity of (133) when $d_X(x,B\smallsetminus A),d_X(y,B\smallsetminus A)>0$ follows from Lemma 5.2 with X replaced by $X\smallsetminus\overline{(B\smallsetminus A)}$ and $f(z)=d_X(z,B\smallsetminus A)$ (using (132) and consequently $\|f\|_\infty\leqslant\pi$, combined with $\|f\|_{\mathrm{Lip}}\leqslant 1$).

Having proved (133), it follows from Lemma 5.8 that Lemma 3.16 will be proven once we show that for every $s,t\in(0,\infty)$ and $x,y\in X$,

$$\begin{aligned}
d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x),\mathfrak{a}(t,y)\big)^2&+d_{\mathbf{Cone}(B)}\big(\mathfrak{b}(s,x),\mathfrak{b}(t,y)\big)^2\\
&\geqslant\frac{\beta^4}{72\pi^2}d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2. &(136)
\end{aligned}$$

We will actually show that

$$\begin{aligned}
&s\geqslant t\quad\text{and}\quad d_X\left(x,B\smallsetminus A\right)\geqslant\frac{\beta}{2}\\
&\implies d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x),\mathfrak{a}(t,y)\big)^2\geqslant\frac{\beta^4}{72\pi^2}d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2. &(137)
\end{aligned}$$

The validity of the implication (137) yields (136) since by replacing (s,x) by (t,y) if necessary we may assume without loss of generality that $s \geqslant t$, and (137) implies that we are done if $d_X(x, B \smallsetminus A) \geqslant \beta/2$. If $d_X(x, B \smallsetminus A) < \beta/2$ then note that by the triangle inequality,

$$d_X(x, B \smallsetminus A) + d_X(x, A \smallsetminus B) \geqslant d_X(B \smallsetminus A, A \smallsetminus B) \overset{(62)}{\geqslant} \beta.$$

Consequently, since we are assuming that $d_X(x, B \smallsetminus A) < \beta/2$, we have $d_X(x, A \smallsetminus B) > \beta/2$, so an application of (137) with A replaced by B and $\mathfrak{a}$ replaced by $\mathfrak{b}$ implies that

$$d_{\mathbf{Cone}(B)}\big(\mathfrak{b}(s,x), \mathfrak{b}(t,y)\big)^2 \geqslant \frac{\beta^4}{18\pi^2} d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2,$$

yielding (136) in the remaining case.

It therefore remains to prove the implication (137). So, suppose from now on that

$$s \geqslant t \qquad \text{and} \qquad d_X(x, B \smallsetminus A) \geqslant \frac{\beta}{2}. \tag{138}$$

It will be convenient to use below the following notation.

$$d_x \overset{\mathrm{def}}{=} d_X(x, B \smallsetminus A), \qquad d_y \overset{\mathrm{def}}{=} d_X(y, B \smallsetminus A), \qquad \theta \overset{\mathrm{def}}{=} d_X(x,y). \tag{139}$$

We proceed by considering the cases $d_x \geqslant d_y$ and $d_x < d_y$ separately.

Case 1: $d_x \geqslant d_y$. In this case, since we are assuming (138) we have $sd_x \geqslant td_y$. Consequently,

$$\begin{aligned} &d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x), \mathfrak{a}(t,y)\big) \\ &\geqslant \frac{1}{3}\max\left\{d_x s - d_y t, d_x s\sqrt{2(1-\cos\theta)}\right\} \qquad (140)\\ &\geqslant \frac{d_x}{3}\max\left\{s-t, s\sqrt{2(1-\cos\theta)}\right\} \\ &\geqslant \frac{\beta}{6\sqrt{2}} d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big), \qquad (141)\end{aligned}$$

where (140) uses Lemma 5.3 and (141) uses (138) and Lemma 5.3. Since $\beta \in [0,\pi]$, (141) implies the conclusion of (137).

Case 2: $d_x < d_y$. Fix $\xi \in (0,1)$ that will be determined later. By the convexity of the function $u \mapsto u^2$ we have

$$\forall\, u, v \in \mathbb{R}, \qquad (u-v)^2 \geqslant \xi u^2 - \frac{\xi}{1-\xi} v^2.$$

Consequently,

$$
\begin{aligned}
(d_x s - d_y t)^2 &= (d_x(s-t) + (d_x - d_y)t)^2 \\
&\geqslant \xi d_x^2 (s-t)^2 - \frac{\xi}{1-\xi}(d_y - d_x)^2 t^2 \\
&\geqslant \frac{\xi\beta^2}{4}(s-t)^2 - \frac{\xi}{1-\xi}\theta^2 t^2 \qquad (142)\\
&\geqslant \frac{\xi\beta^2}{4}(s-t)^2 - \frac{\pi^2\xi}{2(1-\xi)} st(1-\cos\theta), \qquad (143)
\end{aligned}
$$

where (142) uses the assumption $d_x \geqslant \beta/2$ and, by the triangle inequality, that $|d_y - d_x| \leqslant d_X(x,y) = \theta$. In (143) we used (48) and the assumption $s \geqslant t$. Now,

$$
\begin{aligned}
& d_{\mathbf{Cone}(A)}\big(\mathfrak{a}(s,x), \mathfrak{a}(t,y)\big)^2 \\
&= (d_x s - d_y t)^2 + 2 d_x d_y st(1-\cos\theta) \\
&\geqslant (d_x s - d_y t)^2 + \frac{\beta^2}{2} st(1-\cos\theta) \qquad (144)\\
&\geqslant \frac{\xi\beta^2}{4}(s-t)^2 + \frac{1}{4}\left(\beta^2 - \frac{\pi^2\xi}{1-\xi}\right)\cdot 2st(1-\cos\theta) \qquad (145)\\
&\geqslant \frac{1}{4}\min\left\{\xi\beta^2, \beta^2 - \frac{\pi^2\xi}{1-\xi}\right\}\cdot d_{\mathbf{Cone}(X)}\big((s,x),(t,y)\big)^2, \qquad (146)
\end{aligned}
$$

where in (144) we used the fact that in Case 2 we are assuming that $d_y > d_x \geqslant \beta/2$ and in (145) we used (143). Choosing $\xi = \beta^2/(3\pi^2)$ in (146) implies the conclusion of (137) in Case 2 as well. □

6. Embedding $(1+\delta)$-sparse graphs in L_1

In this section all graphs are simple. Following [ABLT06], given $\delta \in (0,1)$ we say that a graph G is $(1+\delta)$-sparse if

$$\forall\, S \subseteq V_G, \qquad |E_G(S)| \leqslant (1+\delta)|S|, \qquad (147)$$

where we recall that $E_G(S)$ denotes the set of those edges in E_G both of whose endpoints lie in S, i.e., $E_G(S) = \{\{u,v\} \in E_G:\ u,v \in S\}$.

For $t > 0$ we denote the set of all cycles of G of size less than t by

$$\mathfrak{C}_t(G) \stackrel{\mathrm{def}}{=} \{C \subseteq V:\ C \text{ is a cycle of } G \text{ and } |C| < t\}. \qquad (148)$$

Claim 6.1. *Fix $\delta \in (0,1)$ and a graph G that is $(1+\delta)$-sparse. Then every $C \in \mathfrak{C}_{1/\delta}(G)$ is an induced cycle, i.e., $|E_G(C)| = |C|$.*

Proof. If C is a cycle of G and $|E_G(C)| > |C|$ then it follows from (147) that $|C|+1 \leqslant |E_G(C)| \leqslant (1+\delta)|C|$, implying that $|C| \geqslant 1/\delta$. $\square$

The starting point of our investigations in the present section is the following result from [ALN+12].

Lemma 6.2 (Lemma 3.11 of [ALN+12])**.** *Fix $\delta \in (0,1)$ and a connected graph G that satisfies*

$$\forall\, S \subseteq V_G, \qquad |E_G(S)| \leqslant (1+\delta)(|S|-1). \tag{149}$$

Then

$$c_1(V_G, d_G) \lesssim 1 + \delta\, \mathrm{diam}(G). \tag{150}$$

While the condition (149) is very close to the $(1+\delta)$-sparseness condition (147), it in fact implies that G also has high girth. The relation between (149) and (147) is clarified in the following lemma.

Lemma 6.3. *Fix $\delta \in (0,1)$ and an integer $g \geqslant 3$. Suppose that a graph G satisfies* (149)*. Then G is $(1+\delta)$-sparse and the girth of G is at least $1+1/\delta$. Conversely, suppose that G is $(1+\delta)$-sparse and the girth of G is at least g. Then*

$$\forall\, S \subseteq V_G, \qquad |E_G(S)| \leqslant (1+\delta)\left(1+\frac{1}{g-1}\right)(|S|-1).$$

Proof. If G contains a cycle of size k then $k \leqslant (1+\delta)(k-1)$ due to (149). Thus $k \geqslant 1+1/\delta$, proving the first assertion of Lemma 6.3. Conversely, suppose that G is $(1+\delta)$-sparse and the girth of G is at least g. Take $S \subseteq V_G$. If $|S| < g$ then the graph $(S, E_G(S))$ cannot contain a cycle, and it is therefore a forest. Consequently $|E_G(S)| \leqslant |S|-1$. If $|S| \geqslant g$ then it follows from (147) that

$$|E_G(S)| \leqslant (1+\delta)|S| \leqslant (1+\delta)\left(|S|-1+\frac{|S|-1}{g-1}\right). \qquad \square$$

Lemma 6.2 therefore implicitly assumes that the graph in question has high girth, an assumption that will not be available in our context. In fact, in our setting δ will tend to 0 as $|V_G| \to \infty$, so the implicit high girth assumption of Lemma 6.2 is quite restrictive. We observe that this high girth assumption is essential for the proof of Lemma 6.2 in [ALN+12]. Indeed, denoting the set of all spanning trees of a connected graph G by $\mathfrak{T}(G)$, the main step of the proof of Lemma 6.2 in [ALN+12] is the following statement.

Claim 6.4 (Claim 3.14 of [ALN+12])**.** *Fix $\delta \in (0,1)$ and suppose that a connected graph G satisfies the assumptions of Lemma 6.2. Then* *there exists a probability measure* μ *on* $\mathfrak{T}(G)$ *such that*

$$\forall\, e \in E_G, \qquad \mu\big(\{T = (V_G, E_T) \in \mathfrak{T}(G):\ e \in E_T\}\big) \geqslant \frac{1}{1+\delta}. \tag{151}$$

The existence of a probability measure μ on $\mathfrak{T}(G)$ that satisfies (151) implies that the girth of G is at least $1+1/\delta$. Indeed, if $C \subseteq V_G$ is a cycle of G then since every $T \in \mathfrak{T}(G)$ satisfies $|E_T \cap E_G(C)| \leqslant |C|-1$ it would follow from (151) that

$$\begin{aligned}|C|-1 &\geqslant \int_{\mathfrak{T}(G)} |E_T \cap E_G(C)| d\mu(T)\\ &= \sum_{e\in E_G(C)} \mu\big(\{T \in \mathfrak{T}(G):\ e \in E_T\}\big) \geqslant \frac{|E_G(C)|}{1+\delta} \geqslant \frac{|C|}{1+\delta},\end{aligned}$$

implying that $|C| \geqslant 1+1/\delta$.

In what follows, given a connected graph G, a spanning tree T of its associated 1-dimensional simplicial complex $\Sigma(G)$ is defined as follows. Given a spanning tree $T_0 = (V_G, E_{T_0}) \in \mathfrak{T}(G)$, include in T all the edges of E_{T_0} as unit intervals. If $e = \{x, y\} \in E_G$ is an edge of G that is not an edge of T_0 then partition the unit interval in $\Sigma(G)$ that corresponds to e into two connected subsets I_x, I_y, where $x \in I_x$ and $y \in I_y$ (thus one of I_x, I_y is a closed interval and the other is a half open interval). T is now obtained from T_0 by gluing I_x to T_0 at the vertex x, and gluing I_y to T_0 at the vertex y. Notice that the completion of T (i.e., taking the closures of the half open intervals that were glued to T_0) is the 1-dimensional simplicial complex of a graph theoretical tree. The set of all spanning trees of $\Sigma(G)$ is denoted below by $\mathfrak{T}(\Sigma(G))$.

Here we prove the following theorem, which assumes only that the graph in question is connected and $(1+\delta)$-sparse.

Theorem 6.5. *Fix* $\delta \in (0,1)$ *and a* $(1+\delta)$*-sparse graph* G. *There exists a probability measure* μ *on* $\mathfrak{T}(\Sigma(G))$ *with respect to which every* $x, y \in \Sigma(G)$ *satisfy*

$$\int_{\mathfrak{T}(\Sigma(G))} \min\{d_T(x,y), \mathrm{diam}(G)\}\, d\mu(T) \lesssim (1+\delta\,\mathrm{diam}(G)) d_{\Sigma(G)}(x,y). \tag{152}$$

Before passing to the proof of Theorem 6.5 we state the following corollary that explains its link to Lemma 6.2.

Corollary 6.6. *Under the assumptions of Theorem 6.5 we have*

$$c_1(\Sigma(G)) \lesssim 1+\delta\,\mathrm{diam}(G).$$

Proof. By Lemma 5.4 there exists a mapping $\Phi : L_1 \to L_1$ such that

$$\forall x, y \in L_1, \qquad \|\Phi(x) - \Phi(y)\|_1 \asymp \min\left\{\|x-y\|_1, \mathrm{diam}(G)\right\}.$$

For every $T \in \mathfrak{T}(\Sigma(G))$ fix an isometric embedding $F_T : (T, d_T) \to L_1$. Let $L_1(\mu, L_1)$ denote the space of all mappings $f : \mathfrak{T}(\Sigma(G)) \to L_1$, equipped with the norm

$$\|f\|_{L_1(\mu,L_1)} \stackrel{\mathrm{def}}{=} \int_{\mathfrak{T}(\Sigma(G))} \|f(T)\|_1 d\mu(T).$$

Then $L_1(\mu, L_1)$ is isometric to L_1. Define $h : \Sigma(G) \to L_1(\mu, L_1)$ by

$$\forall\, T \in \mathfrak{T}(\Sigma(G)),\ \forall\, x \in \Sigma(G), \qquad h(x)(T) = \Phi(F_T(x)).$$

Then, for every $x, y \in \Sigma(G)$,

$$\|h(x) - h(y)\|_{L_1(\mu,L_1)} \asymp \int_{\mathfrak{T}(\Sigma(G))} \min\left\{d_T(x,y), \mathrm{diam}(G)\right\} d\mu(T). \quad (153)$$

Observe that $\min\left\{d_T(x,y), \mathrm{diam}(G)\right\} \gtrsim d_{\Sigma(G)}(x,y)$ for every spanning tree $T \in \mathfrak{T}(\Sigma(G))$ and $x, y \in \Sigma(G)$, because $d_T(x,y) \geqslant d_{\Sigma(G)}(x,y)$ and $d_{\Sigma(G)}(x,y) \leqslant \mathrm{diam}(G) + 1$ (recall (17)). Hence, it follows from (152) and (153) that

$$d_{\Sigma(G)}(x,y) \lesssim \|h(x) - h(y)\|_{L_1(\mu,L_1)} \lesssim (1 + \delta\,\mathrm{diam}(G)) d_{\Sigma(G)}(x,y). \quad \square$$

Remark 6.7. Lemma 6.2 is proved in [ALN+12] by first establishing the same statement as Theorem 6.5 under the stronger assumption (149), with the conclusion (152) holding for every $x, y \in V_G$. It is then shown in [ALN+12, Claim 3.13] that for every tree T and every threshold $a \in (0, \infty)$ the metric $\min\{d_T, a\}$ embeds with distortion $O(1)$ into a convex combination of dominating tree metrics. Thus for the purpose of proving Corollary 6.6 one can use Claim 3.13 of [ALN+12] rather than the stronger statement of Lemma 5.4, which deals with truncation of all of L_1 and not just trees.

Remark 6.8. The fact that Theorem 6.5 yields an embedding of $\Sigma(G)$ rather than (V_G, d_G) is needed in the ensuing arguments, but it does not add significant difficulties to the proof of Theorem 6.5.

The new contribution of Theorem 6.5 is that its conclusion holds under the assumption that G is $(1+\delta)$-sparse rather than under the assumption (149), i.e., without requiring that the girth of G is large. This is achieved via the following strategy: in [ALN+12], given an edge $\{x, y\} \in E_G$ and a spanning tree $T \in \mathfrak{T}(G)$ that does not contain $\{x, y\}$ as an edge, the quantity $\min\left\{d_T(x,y), \mathrm{diam}(G)\right\}$ is bounded from above by $\mathrm{diam}(G)$. But by Claim 6.1, if $\{x, y\}$ is an edge of a cycle $C \in \mathfrak{C}_{1/\delta}(G)$ and $|E_T \cap E_G(C)| \geqslant |C| - 1$ then $d_T(x,y) \leqslant |C| - 1$.

Our strategy is therefore to use this better bound on $d_T(x,y)$, and to show that it is possible to only deal with probability measures μ on $\mathfrak{T}(G)$ that are supported on those spanning trees $T \in \mathfrak{T}(G)$ satisfying $|E_T \cap E_G(C)| \geqslant |C| - 1$ for *every* small cycle $C \in \mathfrak{C}_{1/(3\delta)}(G)$.

We next prove some lemmas as steps towards the proof of Theorem 6.5, starting with the following combinatorial fact whose obvious and short proof is included for completeness.

Lemma 6.9. *Let G be a graph and $C_1, C_2 \subseteq V$ be distinct cycles of G such that $C_1 \cap C_2 \neq \emptyset$. Then $|E_G(C_1 \cup C_2)| \geqslant |C_1 \cup C_2| + 1$.*

Proof. By removing redundant edges and vertices we may assume without loss of generality that $V_G = C_1 \cup C_2$ and if $C_1 = \{x_1, \ldots, x_m\}$ and $C_2 = \{y_1, \ldots, y_n\}$ then $E_G = E_1 \cup E_2$, where

$$E_1 \stackrel{\mathrm{def}}{=} \big\{\{x_1, x_2\}, \{x_2, x_3\}, \ldots, \{x_{m-1}, x_m\}, \{x_m, x_1\}\big\},$$

and

$$E_2 \stackrel{\mathrm{def}}{=} \big\{\{y_1, y_2\}, \{y_2, y_3\}, \ldots, \{y_{n-1}, y_n\}, \{y_n, y_1\}\big\}.$$

If $E_1 \cap E_2 = \emptyset$ then

$$|E_G(C_1 \cup C_2)| = |E| = |C_1| + |C_2| \geqslant |C_1 \cup C_2| + 1,$$

where the final inequality uses $C_1 \cap C_2 \neq \emptyset$. So, suppose that there exist $i, j \in \{1, \ldots, m\}$ and $s, t \in \{1, \ldots, n\}$ with $|i - j| \in \{1, m-1\}$, $|s - t| \in \{1, n-1\}$ and $x_i = y_s$, $x_j = y_t$. If $|C_1|, |C_2| > 3$ then we may contract the edge $\{x_i, x_j\}$ while identifying its endpoints, thus obtaining a graph G' which is the union of two strictly smaller cycles, with $|V_{G'}| = |V| - 1$ and $|E_{G'}| = |E| - 1$. By continuing in this manner we see that it suffices to prove the desired result under the additional assumptions $|C_1| = 3$ and $E_1 \cap E_2 \neq \emptyset$. Thus either $|C_1 \cap C_2| = 2$ or $C_1 \subseteq C_2$. In the former case we have

$$|E_G(C_1 \cup C_2)| = |C_2| + 2 = |C_1 \cup C_2| + 1,$$

and in the latter case, since $C_1 \neq C_2$, we have

$$|E_G(C_1 \cup C_2)| \geqslant |C_2| + 1 = |C_1 \cup C_2| + 1. \qquad \square$$

Lemma 6.10. *Fix $\delta \in (0,1)$ and a connected $(1+\delta)$-sparse graph G. Suppose that $C_1, C_2 \subseteq V_G$ are distinct cycles of G. Then*

$$d_G(C_1, C_2) \geqslant 1 + \frac{1}{\delta} - |C_1| - |C_2|. \tag{154}$$

Proof. It suffices to prove (154) under the assumption

$$|C_1| + |C_2| < 1 + \frac{1}{\delta}. \tag{155}$$

We first note that (155) implies that $C_1 \cap C_2 = \emptyset$. Indeed, otherwise, using Lemma 6.9 and the fact that G is $(1+\delta)$-sparse, we have

$$|C_1 \cup C_2| + 1 \leqslant |E_G(C_1 \cup C_2)| \leqslant (1+\delta)|C_1 \cup C_2|.$$

Consequently $|C_1 \cup C_2| \geqslant 1/\delta$, which contradicts (155) since $C_1 \cap C_2 \neq \emptyset$.

Having proved that $C_1 \cap C_2 = \emptyset$, we have $t \stackrel{\mathrm{def}}{=} d_G(C_1, C_2) > 0$. Take $x \in C_1$ and $y \in C_2$ with $d_G(x,y) = t$ and choose $w_0, \ldots, w_t \in V_G$ such that $w_0 = x$, $w_t = y$ and $\{w_{i-1}, w_i\} \in E_G$ for every $i \in \{1, \ldots, t\}$. By the choice of x and y as the vertices at which $d_G(C_1, C_2)$ is attained, necessarily $w_1, \ldots, w_{t-1} \in V_G \smallsetminus (C_1 \cup C_2)$. Hence,

$$|C_1 \cup C_2 \cup \{w_1, \ldots, w_{t-1}\}| = |C_1| + |C_2| + t - 1,$$

and

$$|E_G\left(C_1 \cup C_2 \cup \{w_1, \ldots, w_{t-1}\}\right)| \geqslant |C_1| + |C_2| + t.$$

Since G is $(1+\delta)$-sparse, it follows that

$$|C_1| + |C_2| + t \leqslant (1+\delta)(|C_1| + |C_2| + t - 1),$$

which simplifies to give $d_G(C_1, C_2) = t \geqslant 1 + \frac{1}{\delta} - |C_1| - |C_2|$. $\square$

Fixing $\delta \in (0,1)$ and a connected $(1+\delta)$-sparse graph G, write

$$t \stackrel{\mathrm{def}}{=} \frac{1}{3\delta}, \tag{156}$$

and define $\Gamma \subseteq V_G$ by

$$\Gamma \stackrel{\mathrm{def}}{=} \bigcup_{C \in \mathfrak{C}_t(G)} C. \tag{157}$$

We shall work below with a "quotient graph"

$$G/\mathfrak{C}_t(G) = \left(V_{G/\mathfrak{C}_t(G)}, E_{G/\mathfrak{C}_t(G)}\right),$$

which is defined as follows.

$$V_{G/\mathfrak{C}_t(G)} \stackrel{\mathrm{def}}{=} (V_G \smallsetminus \Gamma) \cup \mathfrak{C}_t(G), \tag{158}$$

i.e., the vertex set of $G/\mathfrak{C}_t(G)$ consists of those vertices of G that do not belong to any cycle of G of length less than t, and we append to these vertices an additional vertex for every cycle $C \in \mathfrak{C}_t(G)$. Thus

$$\left|V_{G/\mathfrak{C}_t(G)}\right| = |V_G| - |\Gamma| + |\mathfrak{C}_t(G)|. \tag{159}$$

The edges of $G/\mathfrak{C}_t(G)$ are defined as follows. For every $x, y \in V_G \smallsetminus \Gamma$,

$$\{x,y\} \in E_{G/\mathfrak{C}_t(G)} \iff \{x,y\} \in E_G,$$

and for every $(x, C) \in (V_G \smallsetminus \Gamma) \times \mathfrak{C}_t(G)$,

$$\{x, C\} \in E_{G/\mathfrak{C}_t(G)} \iff d_G(x, C) = 1.$$

Under this definition $G/\mathfrak{C}_t(G)$ is simple and connected. Moreover, by Lemma 6.10 every distinct $C_1, C_2 \in \mathfrak{C}_t(G)$ are disjoint, and by Claim 6.1 we have $E_G(C) = |C|$ for every $C \in \mathfrak{C}_t(G)$. Consequently,

$$\left|E_{G/\mathfrak{C}_t(G)}\right| = |E_G| - \sum_{C\in\mathfrak{C}_t(G)} |C| = |E_G| - |\Gamma|. \tag{160}$$

Lemma 6.11. *For every distinct $C_1, C_2 \in \mathfrak{C}_t(G)$ we have*

$$d_{G/\mathfrak{C}_t(G)}(C_1, C_2) > t + 1.$$

Proof. Suppose that $d_{G/\mathfrak{C}_t(G)}(C_1, C_2)$ is minimal among all distinct cycles $C_1, C_2 \in \mathfrak{C}_t(G)$. Then by the definition of $G/\mathfrak{C}_t(G)$, the non-endpoint vertices on the shortest path joining C_1 and C_2 in $G/\mathfrak{C}_t(G)$ consist of vertices in $V_G \smallsetminus \Gamma$. Consequently,

$$d_{G/\mathfrak{C}_t(G)}(C_1, C_2) \geqslant d_G(C_1, C_2) \geqslant 1+\frac{1}{\delta}-|C_1|-|C_2| > 1+\frac{1}{\delta}-2t = t+1,$$

where we used Lemma 6.10, the fact that $|C_1|, |C_2| < t$, and (156). □

Lemma 6.12. *If $0 < \delta < \frac{1}{3}$ then*

$$\left|E_{G/\mathfrak{C}_t(G)}\right| \leqslant \frac{1+\delta}{1-3\delta}\left(\left|V_{G/\mathfrak{C}_t(G)}\right| - 1\right). \tag{161}$$

Proof. If $\mathfrak{C}_t(G) = \emptyset$ then by definition $G = G/\mathfrak{C}_t(G)$ and the girth of G is at least t. Recalling the choice of t in (156), we see that (161) holds true due to the fact that G is $(1+\delta)$-sparse and Lemma 6.3. We may therefore assume from now on that $\mathfrak{C}_t(G) \neq \emptyset$. We may also assume from now on that $G/\mathfrak{C}_t(G)$ contains a cycle, since otherwise it is a tree and therefore $|E_{G/\mathfrak{C}_t(G)}| \leqslant |V_{G/\mathfrak{C}_t(G)}| - 1$.

Suppose first that $|\mathfrak{C}_t(G)| = 1$ and write $\mathfrak{C}_t(G) = \{C\}$. Thus, recalling (157), we have $|\Gamma| = |C| < t$. Since $G/\mathfrak{C}_t(G)$ contains a cycle, it follows that $|V_G \smallsetminus \Gamma| \geqslant t/2$. Indeed, if there is a cycle in $G/\mathfrak{C}_t(G)$ that does not contain the vertex $C \in V_{G/\mathfrak{C}_t(G)}$ then since no cycle of G other than C has length less than t, this cycle contains at least t vertices of $V_G \smallsetminus \Gamma$. If on the other hand there exist distinct $x_1, \ldots, x_k \in V_G \smallsetminus \Gamma$ such that $\{C, x_1\}, \{x_1, x_2\}, \ldots, \{x_{k-1}, x_k\}, \{x_k, C\} \in E_{G/\mathfrak{C}_t(G)}$ then $d_G(x_1, C) = d_G(x_k, C) = 1$. By adding to $\{x_1, \ldots, x_k\}$ the vertices on the shortest path in C joining the nearest neighbors of x_1 and x_k in C, we obtain a cycle of length less than $k+t/2$ in G that differs from C. Since we are assuming that no cycle of G other than C has length less than t, it follows that $k + t/2 \geqslant t$, implying that $|V_G \smallsetminus \Gamma| \geqslant k \geqslant t/2$, as required. We have thus shown that

$$|\Gamma| < t \leqslant 2|V_G \smallsetminus \Gamma|, \tag{162}$$

and consequently, since G is $(1+\delta)$-sparse,

$$\begin{aligned}\left|E_{G/\mathfrak{C}_t(G)}\right| &\stackrel{(160)}{=} |E_G|-|\Gamma| \leqslant (1+\delta)|V_G|-|\Gamma| = (1+\delta)|V_G\setminus\Gamma|+\delta|\Gamma| \\ &\stackrel{(162)}{\leqslant} (1+3\delta)|V_G\setminus\Gamma| = (1+3\delta)\left(\left|V_{G/\mathfrak{C}_t(G)}\right|-1\right), \quad (163)\end{aligned}$$

where the last step of (163) uses (159) and the fact that we are treating the case $|\mathfrak{C}_t(G)|=1$. Since $1+3\delta\leqslant(1+\delta)/(1-3\delta)$, the desired bound (161) follows from (163).

It remains to prove (161) under the assumption $|\mathfrak{C}_t(G)|\geqslant 2$. In this case, for every $C\in\mathfrak{C}_t(G)$ fix an arbitrary $C'\in\mathfrak{C}_t(G)$ with $C\neq C'$. By Lemma 6.11 if $\{C=v_0,v_1,\ldots,v_k=C'\}$ is the shortest path in $G/\mathfrak{C}_t(G)$ joining C and C' then $k>t+1=1/(3\delta)+1\geqslant 2$ and $v_1,\ldots,v_{\lceil t\rceil}\in V\setminus\Gamma$. If we define $A(C)\subseteq V\setminus\Gamma$ by

$$A(C)\stackrel{\mathrm{def}}{=}\left\{v_1,\ldots,v_{\lfloor (t+1)/2\rfloor}\right\},$$

then $A(C_1)\cap A(C_2)=\emptyset$ for every distinct $C_1,C_2\in\mathfrak{C}_t(G)$, since otherwise $d_{G/\mathfrak{C}_t(G)}(C_1,C_2)\leqslant 2\lfloor (t+1)/2\rfloor\leqslant t+1$, contradicting Lemma 6.11. This shows that

$$\begin{aligned}|V_G\setminus\Gamma| &\geqslant \left|\bigcup_{C\in\mathfrak{C}_t(G)} A(C)\right| = \sum_{C\in\mathfrak{C}_t(G)}|A(C)| \\ &\geqslant |\mathfrak{C}_t(G)|\cdot\left\lfloor\frac{t+1}{2}\right\rfloor \geqslant \frac{t-1}{2}|\mathfrak{C}_t(G)|. \quad (164)\end{aligned}$$

Hence,

$$\begin{aligned}|\Gamma| &= \sum_{C\in\mathfrak{C}_t(G)}|C| < t|\mathfrak{C}_t(G)| \\ &\stackrel{(164)}{\leqslant} \frac{2t}{t-1}|V_G\setminus\Gamma| \stackrel{(156)}{=} \frac{2}{1-3\delta}|V_G\setminus\Gamma|. \quad (165)\end{aligned}$$

Now, arguing similarly to (163) we have

$$\begin{aligned}\left|E_{G/\mathfrak{C}_t(G)}\right| &\stackrel{(160)}{=} |E_G|-|\Gamma| \leqslant (1+\delta)|V_G|-|\Gamma| = (1+\delta)|V_G\setminus\Gamma|+\delta|\Gamma| \\ &\stackrel{(165)}{\leqslant} \left(1+\delta+\frac{2\delta}{1-3\delta}\right)|V_G\setminus\Gamma| \stackrel{(159)}{\leqslant} \frac{1-3\delta^2}{1-3\delta}\left(\left|V_{G/\mathfrak{C}_t(G)}\right|-2\right). \qquad \square\end{aligned}$$

Corollary 6.13. *If $0<\delta<\frac13$ then the quotient graph $G/\mathfrak{C}_t(G)$ satisfies (149) with δ replaced by $4\delta/(1-3\delta)$, i.e,*

$$\forall\, S\subseteq V_{G/\mathfrak{C}_t(G)}, \qquad \left|E_{G/\mathfrak{C}_t(G)}(S)\right|\leqslant\frac{1+\delta}{1-3\delta}(|S|-1).$$

Proof. Fix $S \subseteq V_{G/\mathfrak{C}_t(G)}$ and let $S_1, \ldots, S_m \subseteq S$ be the connected components of the graph $(S, E_{G/\mathfrak{C}_t(G)}(S))$. For every $i \in \{1, \ldots, m\}$ we lift each S_i to a subset $U_i \subseteq V_G$ that is given by

$$U_i \stackrel{\mathrm{def}}{=} \big(S_i \cap (V_G \smallsetminus \Gamma)\big) \bigcup \left(\bigcup_{C \in S_i \cap \mathfrak{C}_t(G)} C \right).$$

Since $(S_i, E_{G/\mathfrak{C}_t(G)}(S_i))$ is connected, the graph $H_i \stackrel{\mathrm{def}}{=} (U_i, E_G(U_i))$ is connected as well. By definition $H_i/\mathfrak{C}_t(H_i) = (S_i, E_{G/\mathfrak{C}_t(G)}(S_i))$, so, since H_i is connected and $(1+\delta)$-sparse, by Lemma 6.12 we have

$$\forall\, i \in \{1, \ldots, m\}, \qquad \big|E_{G/\mathfrak{C}_t(G)}(S_i)\big| \leqslant \frac{1+\delta}{1-3\delta}(|S_i| - 1). \tag{166}$$

Consequently,

$$\begin{aligned}\big|E_{G/\mathfrak{C}_t(G)}(S)\big| = \sum_{i=1}^{m} \big|E_{G/\mathfrak{C}_t(G)}(S_i)\big| &\overset{(166)}{\leqslant} \frac{1+\delta}{1-3\delta}\left(\sum_{i=1}^{m} |S_i| - m\right) \\ &= \frac{1+\delta}{1-3\delta}\,(|S| - m) \leqslant \frac{1+\delta}{1-3\delta}\,(|S| - 1)\,. \qquad \square\end{aligned}$$

Before proceeding it will be convenient to introduce the following notation. Firstly, recalling (157), the set of edges $E_G(\Gamma) \subseteq E_G$ consists of those edges in E_G that belong to a cycle of length less than t. Since, by Lemma 6.10, these cycles are pairwise disjoint, we can associate to every $e \in E_G(\Gamma)$ a unique cycle $C_e \in \mathfrak{C}_t(G)$ such that $e \in E_G(C_e)$. Secondly, let $\mathfrak{T}_{\mathrm{good}}(G) \subseteq \mathfrak{T}(G)$ be the set of those spanning trees T of G satisfying $|E_T \cap E_G(C)| = |C| - 1$ for every $C \in \mathfrak{C}_t(G)$, i.e.,

$$\mathfrak{T}_{\mathrm{good}}(G) \stackrel{\mathrm{def}}{=} \{T \in \mathfrak{T}(G) :\ \forall\, C \in \mathfrak{C}_t(G), \quad |E_T \cap E_G(C)| = |C| - 1\}\,.$$

Lemma 6.14. *Suppose that $0 < \delta < \frac{1}{3}$. There exists a probability measure μ on $\mathfrak{T}_{\mathrm{good}}(G)$ such that for every $e \in E_G \smallsetminus E_G(\Gamma)$ we have*

$$\mu\left(\{T \in \mathfrak{T}_{\mathrm{good}}(G) :\ e \in E_T\}\right) \geqslant \frac{1-3\delta}{1+\delta}, \tag{167}$$

and in addition for every $e \in E_G(\Gamma)$ we have

$$\mu\left(\{T \in \mathfrak{T}_{\mathrm{good}}(G) :\ e \in E_T\}\right) \geqslant \frac{|C_e| - 1}{|C_e|}. \tag{168}$$

Proof. Due to Corollary 6.13, we can apply Claim 6.4 to the quotient graph $G/\mathfrak{C}_t(G)$, thus obtaining a probability measure ν on $\mathfrak{T}(G/\mathfrak{C}_t(G))$ such that for every $e \in E_{G/\mathfrak{C}_t(G)}$ we have

$$\nu\big(\{T \in \mathfrak{T}(G/\mathfrak{C}_t(G)) :\ e \in E_T\}\big) \geqslant \frac{1-3\delta}{1+\delta}. \tag{169}$$

By the definition of $G/\mathfrak{C}_t(G)$, there is a bijection between the quotient edges $E_{G/\mathfrak{C}_t(G)}$ and $E_G \setminus E_G(\Gamma)$. Under this bijection, ν can be lifted to a probability measure σ on the subsets of $E_G \setminus E_G(\Gamma)$. Let τ be the probability measure on the subsets of $E_G(\Gamma)$ given by selecting a subset of size $|C|-1$ uniformly at random from each $C \in \mathfrak{C}_t(\Gamma)$, where these selections are performed independently for different cycles in $\mathfrak{C}_t(G)$. If $A \subseteq E_G \setminus E_G(\Gamma)$ and $B \subseteq E_G(\Gamma)$ are such that $\sigma(A), \tau(B) > 0$ then $A \cup B$ form the edges of a spanning tree of G, which by design belongs to $\mathfrak{T}_{\mathrm{good}}(G)$. Thus $\sigma \times \tau$ induces a probability measure μ on $\mathfrak{T}_{\mathrm{good}}(G)$. The desired estimate (167) holds true due to (169). The desired estimate (168) holds true because if $e \in E_G(\Gamma)$ and $T \in \mathfrak{T}_{\mathrm{good}}(G)$ is distributed according to μ then each subset of C_e of size $|C_e|-1$ is contained in E_T with probability $1/|C_e|$. □

Proof of Theorem 6.5. Let μ be the probability measure on $\mathfrak{T}_{\mathrm{good}}(G)$ from Lemma 6.14. We obtain from μ a probability distribution μ_Σ over spanning trees of the 1-dimensional simplicial complex $\Sigma(G)$ as follows. Given $T_0 \in \mathfrak{T}_{\mathrm{good}}(G)$ that is distributed according to μ, define $T \in \mathfrak{T}(\Sigma(G))$ as follows. Include all the edges of T_0 as unit intervals in T. Let $\{U_e\}_{E_G \setminus E_{T_0}}$ be i.i.d. random variables that are uniformly distributed on $[0,1]$, and attach to T a closed interval of length U_e at the vertex x and a half open interval of length $1-U_e$ at the vertex y (here we arbitrarily choose a labeling of the endpoints of every edge in E_G as x and y). The distribution of the resulting spanning tree $T \in \mathfrak{T}(\Sigma(G))$ is denoted μ_Σ.

It suffices to prove (152) when x and y lie on the same interval corresponding to an edge $e \in E_G$. Let $[x,y]$ denote the geodesic joining x and y in $\Sigma(G)$ (it is a sub-interval of the unit interval corresponding to e). Given $T \in \mathfrak{T}(\Sigma(G))$, write $[x,y] \subseteq T$ if and only if the geodesic $[x,y]$ is also a geodesic in T. We distinguish between two cases.

Case 1. We have $e \in E_G \setminus E_G(\Gamma)$. Then there exists a sub-interval $I_{x,y} \subseteq [0,1]$ of length $d_{\Sigma(G)}(x,y)$ such that

$$\begin{aligned}
&\mu_\Sigma\big(\{T \in \mathfrak{T}(\Sigma(G)) :\ [x,y] \not\subseteq T\}\big) \\
&\qquad = \mu\big(\{T \in \mathfrak{T}(G) :\ e \notin E_T\}\big) \cdot \Pr[U_e \in I_{x,y}] \\
&\qquad \overset{(167)}{\leqslant} \frac{4\delta}{1+\delta} d_{\Sigma(G)}(x,y). \qquad (170)
\end{aligned}$$

Consequently,

$$\begin{aligned}
&\int_{\mathfrak{T}(\Sigma(G))} \min\{d_T(x,y), \mathrm{diam}(G)\}\, d\mu_\Sigma(T)\\
&\quad\lesssim \quad \mu_\Sigma\big(\{T\in\mathfrak{T}(\Sigma(G)):\ [x,y]\subseteq T\}\big)\cdot d_{\Sigma(G)}(x,y)\\
&\qquad\quad +\mu_\Sigma\big(\{T\in\mathfrak{T}(\Sigma(G)):\ [x,y]\nsubseteq T\}\big)\cdot \mathrm{diam}(\Sigma(G))\\
&\quad\overset{(170)}{\lesssim} \quad \big(1+\delta\,\mathrm{diam}(G)\big) d_{\Sigma(G)}(x,y),
\end{aligned}$$

which is the desired inequality (152).

Case 2. We have $e\in E_G(\Gamma)$. Then arguing as in (170) with the use of (167) replaced by the use of (168),

$$\mu_\Sigma\big(\{T\in\mathfrak{T}(\Sigma(G)):\ [x,y]\nsubseteq T\}\big)\leqslant \frac{d_{\Sigma(G)}(x,y)}{|C_e|}. \tag{171}$$

Moreover, if $T_0\in\mathfrak{T}_{\mathrm{good}}(G)$ then, since T_0 contains all but one of the edges of C_e, we have $d_T(x,y)\leqslant |C_e|$. Since μ is supported on $\mathfrak{T}_{\mathrm{good}}(G)$, we conclude that

$$\begin{aligned}
&\int_{\mathfrak{T}(\Sigma(G))} d_T(x,y) d\mu_\Sigma(T)\\
&\quad\leqslant \quad \mu_\Sigma\big(\{T\in\mathfrak{T}(\Sigma(G)):\ [x,y]\subseteq T\}\big)\cdot d_{\Sigma(G)}(x,y)\\
&\qquad\quad +\mu_\Sigma\big(\{T\in\mathfrak{T}(\Sigma(G)):\ [x,y]\nsubseteq T\}\big)\cdot |C_e|\\
&\quad\overset{(171)}{\leqslant} \quad 2d_{\Sigma(G)}(x,y),
\end{aligned}$$

proving the desired inequality (152) in Case 2 as well. □

Remark 6.15. Using Lemma 6.10 one can prove that if $\delta\leqslant \mathrm{diam}(G)/15$ then G can contain at most one cycle, implying that $c_1(G)=1$.

7. Geometric properties of random regular graphs

Fix $n,d\in\mathbb{N}$. While we are interested in proving estimates about the probability distribution $\mathcal{G}_{n,d}$, it is often simpler to argue about a more tractable probability distribution on $\mathbb{G}_n$, denoted $\mathcal{P}_{n,d}$ and called the *pairing model*; see [Wor99, Sec. 2] and the references therein.

To define $\mathcal{P}_{n,d}$ assume from now on that nd is even and let M be a uniformly random perfect matching of the set $\{1,\ldots,n\}\times\{1,\ldots,d\}$ (recall that a perfect matching is a partition into subsets of size 2). Now define a random d-regular graph G with $V_G=\{1,\ldots,n\}$ and

$$\forall\, i,j\in\{1,\ldots,n\},\qquad E_G(i,j)\stackrel{\mathrm{def}}{=}\sum_{a=1}^{d}\sum_{b=1}^{d}\mathbf{1}_{\{\{(i,a),(j,b)\}\in M\}}.$$

In other words, the edges of G are obtained by "projecting" the perfect matching M onto $\{1,\ldots,n\}$, i.e., whenever $\{(i,a),(j,b)\}\in M$ we add an edge joining i and j in G. The graph G that is obtained in this way from a perfect matching M of $\{1,\ldots,n\}\times\{1,\ldots,d\}$ is denoted $G(M)$.

The probability measures $\mathcal{G}_{n,d}$ and $\mathcal{P}_{n,d}$ are *contiguous* in the following sense. As explained in [Wor99, Sec. 2.2], there exists $\alpha(d)\in(0,\infty)$ such that

$$\forall\, A\in \mathbb{G}_n,\qquad \mathcal{G}_{n,d}(A)\leqslant \alpha(d)\cdot \mathcal{P}_{n,d}(A). \tag{172}$$

By [Wor99], for $d\lesssim \sqrt[3]{n}$ we have $\log\alpha(d)\lesssim d^2$, but we will not explicitly state the dependence on d from now on.

There is another standard probability distribution on n-vertex simple graphs: the Erdös-Rényi model $G(n,p)$ for $p\in(0,1)$. The statements below on the sparsity of subsets of random graphs have been proved in [ABLT06] for the $G(n,p)$ model, and here we need to extend them to the $\mathcal{G}_{n,d}$ model. This leads to some technical changes, but the essence of the argument is the same as in [ABLT06].

The following lemma is well-known, and it follows from much more precise estimates that are available in the literature (see e.g. [McK81]). We include its straightforward proof for the sake of completeness.

Lemma 7.1. *Let F be a set of unordered pairs of vertices in $\{1,\ldots,n\}$ with $|F|<nd/4$. Then*

$$\mathcal{P}_{n,d}\left(\{G\in\mathbb{G}_n:\ F\subseteq E_G\}\right)\leqslant\left(\frac{2d}{n}\right)^{|F|}. \tag{173}$$

Proof. Write $k=|F|$ and $F=\{f_1,\ldots,f_k\}$. Let M be a uniformly random matching of $\{1,\ldots,n\}\times\{1,\ldots,d\}$. For every $\ell\in\{1,\ldots,k\}$ write $f_\ell=\{i_\ell,j_\ell\}$.

We claim that for every $\ell\in\{1,\ldots,k\}$ we have

$$\Pr\left[f_\ell\in G(M)\big|f_1,\ldots,f_{\ell-1}\in G(M)\right]\leqslant\frac{2d}{n}, \tag{174}$$

where Pr is the uniform probability on the matching M. Once (174) is proved, the desired estimate (173) would follow because by the definition of $\mathcal{P}_{n,d}$ we have

$$\mathcal{P}_{n,d}\left(\{G\in\mathbb{G}_n:\ F\subseteq E_G\}\right)\\ =\prod_{\ell=1}^{k}\Pr\left[f_\ell\in G(M)\big|f_1,\ldots,f_{\ell-1}\in G(M)\right]\overset{(174)}{\leqslant}\left(\frac{2d}{n}\right)^{k}.$$

To verify (174) observe that in order for $\{i_\ell,j_\ell\}$ to be in $E_{G(M)}$ there must be $a,b\in\{1,\ldots,d\}$ such that $\{(i_\ell,a),(j_\ell,b)\}\in M$. For every fixed $a \in \{1,\dots,d\}$, if we know that $f_1,\dots,f_{\ell-1} \in G(M)$ then there are $nd-2(\ell-1)-1$ possible elements of $\{1,\dots,n\}\times\{1,\dots,d\}$ that can be matched by M to (i_ℓ,a). We are assuming that $\ell \leqslant k < nd/4$, so that $nd-2(\ell-1)-1 \geqslant nd/2$. We therefore have at least $nd/2$ pairs in $\{1,\dots,n\}\times\{1,\dots,d\}$ that can be matched to (i_ℓ,a), from which at most d can project to the edge f_ℓ. The probability for this to happen is therefore at most $2/n$. The desired estimate (174) now follows since there are at most d possible values of a. □

Definition 7.2. For every $\varepsilon,\delta \in (0,\infty)$ denote by $\mathcal{S}_{\varepsilon,\delta}$ the set of all graphs G with the property that $|E_G(S)| < (1+\delta)|S|$ for every $S \subseteq V_G$ satisfying $|S| \leqslant |V_G|^{1-\varepsilon}$.

Lemma 7.3 below is similar to [ABLT06, Lem 2.8], which treats the same question for Erdös-Rényi graphs (see also [ALN$^+$12, Lem. 3.10]).

Lemma 7.3. *For $\varepsilon \in (0,1)$ and two integers $n,d \geqslant 3$ define*

$$\delta \stackrel{\mathrm{def}}{=} \frac{7\log d}{\varepsilon \log n}. \tag{175}$$

Then

$$1-\mathcal{G}_{n,d}\left(\mathcal{S}_{\varepsilon,\delta}\right) \lesssim_{d,\varepsilon} \frac{1}{n^{1-\varepsilon}}. \tag{176}$$

Proof. By adjusting the implicit constant in (176) if necessary, we may assume in the computations below that n is larger than an appropriate constant that may depend only on ε and d. In particular, we assume throughout that $n > d^{7/\varepsilon}$, or equivalently that $\delta \in (0,1)$.

For every $k \in \{3,\dots,n\}$ define $\mathcal{B}_\delta^k$ to be the set of all graphs G for which there exists $S \subseteq V_G$ with $|S| = k$ and $|E_G(S)| \geqslant (1+\delta)k$. The complement of the event $\mathcal{S}_{\varepsilon,\delta}$ is the union of $\mathcal{B}_\delta^k$ for $k \in \{3,\dots,\lfloor n^{1-\varepsilon}\rfloor\}$. Therefore, due to (172) we have

$$1-\mathcal{G}_{n,d}\left(\mathcal{S}_{\varepsilon,\delta}\right) \lesssim_d 1-\mathcal{P}_{n,d}\left(\mathcal{S}_{\varepsilon,\delta}\right) \leqslant \sum_{k=3}^{\lfloor n^{1-\varepsilon}\rfloor} \mathcal{P}_{n,d}\left(\mathcal{B}_\delta^k\right). \tag{177}$$

Since our assumption $n > d^{7/\varepsilon}$ implies that $k \leqslant nd/16$, by Lemma 7.1,

$$\begin{aligned}
\mathcal{P}_{n,d}\left(\mathcal{B}_\delta^k\right) &\leqslant \binom{n}{k}\binom{\binom{k}{2}}{\lceil(1+\delta)k\rceil}\left(\frac{2d}{n}\right)^{\lceil(1+\delta)k\rceil}\\
&\leqslant \left(\frac{en}{k}\right)^k\left(\frac{ek(k-1)d}{n\lceil(1+\delta)k\rceil}\right)^{\lceil(1+\delta)k\rceil} \qquad (178)\\
&\leqslant \left(\frac{en}{k}\left(\frac{e(k-1)d}{(1+\delta)n}\right)^{1+\delta}\right)^k \leqslant \left(\frac{e^3k^\delta d^{1+\delta}}{n^\delta}\right)^k. \qquad (179)
\end{aligned}$$

In (178) we used the standard estimate $\binom{m}{\ell} \leqslant \left(\frac{em}{\ell}\right)^{\ell}$, which holds for every $m \in \mathbb{N}$ and $\ell \in \{1, \ldots, m\}$. In the penultimate inequality of (179) we used the fact that the function $x \mapsto (ek(k-1)d/(nx))^x$ is decreasing when $x \geqslant dk(k-1)/n$, and $\lceil (1+\delta)k \rceil \geqslant (1+\delta)k \geqslant dk(k-1)/n$ by our assumption $n > d^{7/\varepsilon}$.

If $k < 1/\delta$ then $\lceil (1+\delta)k \rceil = k+1$, and therefore (178) implies that $\mathcal{P}_{n,d}\left(\mathcal{B}_\delta^k\right) \lesssim k(e^2d)^{k+1}/n$. Using this estimate in combination with (179) for $k \geqslant 1/\delta$, it follows from (177) that

$$\begin{aligned}1-\mathcal{G}_{n,d}\left(\mathcal{S}_{\varepsilon,\delta}\right) &\lesssim_d \frac{1}{n}\int_3^{1+1/\delta} x(e^2d)^{x+1}dx + \int_{1/\delta}^{n^{1-\varepsilon}+1}\left(\frac{e^3x^\delta d^{1+\delta}}{n^\delta}\right)^x dx\\ &\lesssim_d \frac{(e^2d)^{1/\delta}}{n\delta} + \int_{1/\delta}^{2n^{1-\varepsilon}}\left(\frac{e^3x^\delta d^{1+\delta}}{n^\delta}\right)^x dx. \end{aligned} \tag{180}$$

To estimate the final integral in (180), observe first that due to (175), its integrand is less than 1, i.e, $e^3 2^\delta n^{(1-\varepsilon)\delta} d^{1+\delta}/n^\delta < 1$. Since $\delta \leqslant 1$ and $n^{\varepsilon\delta} = d^7$, this would follow from $d^7 > 2e^3d^2$, which is true since $d \geqslant 3$. Now fix $M \in [1/\delta, 2n^{1-\varepsilon}]$ and proceed as follows.

$$\begin{aligned}&\int_{1/\delta}^{2n^{1-\varepsilon}}\left(\frac{e^3x^\delta d^{1+\delta}}{n^\delta}\right)^x dx\\ &\leqslant \int_{1/\delta}^{M}\left(\frac{e^3M^\delta d^{1+\delta}}{n^\delta}\right)^x dx + \int_{M}^{2n^{1-\varepsilon}}\left(\frac{e^3(2n^{(1-\varepsilon)})^\delta d^{1+\delta}}{n^\delta}\right)^x dx\\ &\leqslant \int_{1/\delta}^{\infty}\left(\frac{e^3M^\delta d^{1+\delta}}{n^\delta}\right)^x dx + \int_{M}^{\infty}\left(\frac{2^\delta e^3 d^{1+\delta}}{n^{\varepsilon\delta}}\right)^x dx\\ &= \frac{\left(e^3M^\delta d^{1+\delta}n^{-\delta}\right)^{1/\delta}}{\log\left(n^\delta e^{-3}M^{-\delta}d^{-1-\delta}\right)} + \frac{\left(2^\delta e^3 d^{1+\delta} n^{-\varepsilon\delta}\right)^M}{\log\left(n^{\varepsilon\delta}2^{-\delta}e^{-3}d^{-1-\delta}\right)}\\ &\lesssim_d \frac{M(e^3d)^{1/\delta}}{n} + \frac{1}{d^M}, \end{aligned} \tag{181}$$

where in (181) we used the fact that $n^{\varepsilon\delta} = d^7$ and $d \geqslant 3$. Choosing $M = \log n$ and substituting (181) into (180) while using (175) and the fact that $e^3d \leqslant d^4$ (because $d \geqslant 3$), we conclude that

$$1-\mathcal{G}_{n,d}\left(\mathcal{S}_{\varepsilon,\delta}\right) \lesssim_d \frac{d^{4/\delta}\log n}{n} \overset{(175)}{=} \frac{n^{4\varepsilon/7}\log n}{n} \lesssim_\varepsilon \frac{1}{n^{1-\varepsilon}}. \qquad \square$$

The final preparatory lemma that we will need about graphs sampled from $\mathcal{G}_{n,d}$ is that they have only a few short cycles. Specifically, Lemma 7.4 below asserts that the probability that a graph sampled from $\mathcal{G}_{n,d}$ has less than $\sqrt{n}$ cycles of length $\lceil (\log_{d-1} n)/7 \rceil$ is at least

$1 - c(d)/\sqrt[3]{n}$, where $c(d) \in (0,\infty)$ may depend only on d. This is a standard fact but we include its simple proof here since we could not locate the statement below in the literature. Precise asymptotics of the expected numbers of short cycles of a graph sampled from $\mathcal{G}_{n,d}$ were obtained in [MWW04], and Lemma 7.4 follows from the estimates of [MWW04] by Markov's inequality.

Lemma 7.4. *For every two integers $n, d \geqslant 3$ we have*

$$\mathcal{G}_{n,d}\left(\left\{G \in \mathbb{G}_n : \left|\mathfrak{C}_{\lceil(\log_{d-1} n)/7\rceil}(G)\right| \geqslant \sqrt{n}\right\}\right) \lesssim_d \frac{1}{\sqrt[3]{n}}.$$

(Recall that the set of cycles of length less than $t \in \mathbb{N}$ in a graph G was denoted in (148) *by $\mathfrak{C}_t(G)$.)*

Proof. For $r \in \mathbb{N}$ write $X_r(G) \stackrel{\mathrm{def}}{=} |\mathfrak{C}_{r+1}(G)| - |\mathfrak{C}_r(G)|$, i.e., $X_r(G)$ is the number of cycles of length r in G. By [MWW04, Eq. (2.2)], for $r \leqslant (\log_{d-1} n)/7$ we have the expectation bound $\int_{\mathbb{G}_n} X_r d\mathcal{G}_{n,d} \lesssim (d-1)^r$. Hence by Markov's inequality, for every integer $r \leqslant (\log_{d-1} n)/7$,

$$\mathcal{G}_{n,d}\left(\left\{G \in \mathbb{G}_n : X_r(G) \geqslant \frac{\sqrt{n}}{\lceil(\log_{d-1} n)/7\rceil}\right\}\right) \lesssim \frac{(d-1)^{\lceil(\log_{d-1} n)/7\rceil}\log_{d-1} n}{\sqrt{n}} \lesssim_d \frac{\log n}{n^{5/14}}.$$

Consequently,

$$\begin{aligned}
&\mathcal{G}_{n,d}\left(\left\{G \in \mathbb{G}_n : \left|\mathfrak{C}_{\lceil(\log_{d-1} n)/7\rceil}(G)\right| \geqslant \sqrt{n}\right\}\right)\\
&= \mathcal{G}_{n,d}\left(\left\{G \in \mathbb{G}_n : \sum_{r=3}^{\lceil(\log_{d-1} n)/7\rceil} X_r(G) \geqslant \sqrt{n}\right\}\right)\\
&\leqslant \sum_{r=3}^{\lceil(\log_{d-1} n)/7\rceil} \mathcal{G}_{n,d}\left(\left\{G \in \mathbb{G}_n : X_r(G) \geqslant \frac{\sqrt{n}}{\lceil(\log_{d-1} n)/7\rceil}\right\}\right)\\
&\lesssim_d \frac{(\log n)^2}{n^{5/14}} \lesssim \frac{1}{\sqrt[3]{n}}.
\end{aligned}$$

□

7.1. **Proof of Lemma 3.12.** By lemma 7.3 we know that if G is sampled from $\mathcal{G}_{n,d}$ then with high probability for every $S \subseteq V_G$ with $|S| \leqslant n^{1-\varepsilon}$ the graph $(S, E_G(S))$ is $(1+\delta)$-sparse. One is then tempted to use Corollary 6.6 in order to embed the metric space (S, d_G) into L_1, but this is problematic since the metric that the graph $(S, E_G(S))$ induces on S can be very different from the restriction of d_G to S (in fact, $(S, E_G(S))$ may not even be connected). Following an idea of [ALN+12], the following lemma will be used to remedy this matter.

Lemma 7.5. *Fix $n, d, r \in \mathbb{N}$ and let G be a connected n-vertex graph whose maximum degree is d. Suppose that $S, T \subseteq V_G$ satisfy*

$$S \subseteq \bigcup_{u\in T} B_G(u,r), \tag{182}$$

where $B_G(u,r) \stackrel{\mathrm{def}}{=} \{v \in V_G : \ d_G(u,v) \leqslant r\}$ denotes the ball of radius r and center u in the metric space (V_G, d_G). Then there exists $U \subseteq V_G$ with $S \subseteq U$ such that

$$|U| \leqslant |T|\left(d(d-1)^{3r-1} + \mathrm{diam}(G)\right), \tag{183}$$

and if we let H denote the graph $(U, E_G(U))$ then

$$\mathrm{diam}(H) \leqslant 6r + 2\,\mathrm{diam}(G),$$

and for every $a, b \in S$,

$$d_G(a,b) \leqslant d_H(a,b) \leqslant 2\left(\frac{\mathrm{diam}(G)}{r} + 1\right) d_G(a,b). \tag{184}$$

Proof. For every $x, y \in V_G$ let $P_{x,y} \subseteq V_G$ be an arbitrary shortest path joining x and y in G. Fix $z \in V_G$ and define

$$U \stackrel{\mathrm{def}}{=} \left(\bigcup_{x\in T} B_G(x,3r)\right) \bigcup \left(\bigcup_{x\in T} P_{x,z}\right).$$

Then,

$$\begin{aligned} |U| &\leqslant \sum_{x\in T} |B_G(x,3r)| + \sum_{x\in T} |P_{x,z}| \\ &\leqslant |T|d(d-1)^{3r-1} + \sum_{x\in T}(d_G(x,z)-1) + 1 \qquad (185) \\ &\leqslant |T|d(d-1)^{3r-1} + |T|(\mathrm{diam}(G)-1) + 1, \end{aligned}$$

where in (185) we used the fact that since G has maximum degree d, for every $k \in \mathbb{N}$ and $w \in G$ we have $|B_G(w,k)| \leqslant d(d-1)^{k-1}$. The fact that $\mathrm{diam}(H) \leqslant 6r + 2\,\mathrm{diam}(G)$ is immediate from the definition of U.

Having proved (183), we proceed to prove (184). Because H is a subgraph of G, the leftmost inequality in (184) holds true for every $a, b \in V_G$. So, fixing $a, b \in S$, it remains to prove the rightmost inequality in (184). To do so we distinguish between two cases.

Case 1. There exists $x \in T$ such that $a \in B_G(x,r)$ and $b \in B_G(x,2r)$. Then $d_G(a,b) \leqslant d_G(a,x) + d_G(b,x) \leqslant 3r$. Thus $|P_{a,b}| \leqslant 3r+1$. Write $P_{a,b} = \{w_0 = a, w_1, \ldots, w_m = b\}$ with $m \leqslant 3r$ and $\{w_{i-1}, w_i\} \in E_G$ for every $i \in \{1, \ldots, m\}$. For every $i \in \{0, \ldots, 2r\}$ we have $d_G(w_i, a) \leqslant 2r$, and therefore $d_G(w_i, x) \leqslant d_G(w_i, a) + d_G(x, a) \leqslant 3r$. Similarly, for every

$i \in \{2r+1, \ldots, m\}$ we have $d_G(w_i, b) \leqslant r$, and therefore $d_G(w_i, x) \leqslant 3r$. This shows that $P_{a,b} \subseteq B_G(x, 3r) \subseteq U$, and hence $d_G(a, b) = d_H(a, b)$.

Case 2. There exist distinct $x, y \in T$ such that $a \in B_G(x, r)$ and $b \in B_G(y, r) \smallsetminus B_G(x, 2r)$. Then

$$d_G(a, b) \geqslant d_G(x, b) - d_G(x, a) \geqslant r. \tag{186}$$

Since $a \in B_G(x, r)$ we have $P_{a,x} \subseteq B_G(x, r) \subseteq U$. For the same reason, $P_{b,y} \subseteq U$. Also, by the definition of U we have $P_{x,z}, P_{y,z} \subseteq U$. By considering the path $P_{a,x} \cup P_{x,z} \cup P_{y,z} \cup P_{b,y} \subseteq U$ we see that

$$\begin{aligned} d_H(a, b) &\leqslant d_G(a, x) + d_G(x, z) + d_G(y, z) + d_G(b, y) \\ &\leqslant 2r + 2\,\mathrm{diam}(G) \overset{(186)}{\leqslant} 2\left(1 + \frac{\mathrm{diam}(G)}{r}\right) d_G(a, b), \end{aligned}$$

completing the verification of (184) in this case as well.

Note that due to (182) one of the above two cases must occur, so the proof of (184) is complete. □

Definition 7.6. Fix two integers $d, t \geqslant 3$ and a d-regular simple graph G. For every cycle C of G fix an arbitrary edge $e_C \in E_G(C)$. Denote

$$I_G^t \stackrel{\mathrm{def}}{=} \{e_C : \ C \in \mathfrak{C}_t(G)\}.$$

Thus I_G^t contains a single representative edges from each cycle of G of length less than t (formally, I_G depends also on the choices of the edges e_C, but we fix such a choice once and for all and do not indicate this dependence explicitly). Denote the graph $(V_G, E_G \smallsetminus I_G^t)$ by L_G^t. Note that by definition we have $\mathrm{girth}(L_G^t) \geqslant t$.

Before proceeding we record for future use two simple lemmas about the concepts that were introduced in Definition 7.6.

Lemma 7.7. *Fix three integers $r, t, d \geqslant 3$ and write*

$$\eta \stackrel{\mathrm{def}}{=} \frac{1}{r + 2t - 3}. \tag{187}$$

Suppose that G is a connected d-regular graph such that for every subset $S \subseteq V_G$ we have

$$|S| \leqslant d(d-1)^{\frac{t+r}{2}-1} \cdot |\mathfrak{C}_t(G)| \implies |E_G(S)| \leqslant (1+\eta)|S|. \tag{188}$$

Then for every distinct cycles $C_1, C_2 \in \mathfrak{C}_t(G)$ we have $d_G(C_1, C_2) \geqslant r$. Moreover, the graph L_G^t is connected and

$$\mathrm{diam}\left(L_G^t\right) \leqslant \frac{t + r - 1}{r + 1} \cdot \mathrm{diam}(G) + \frac{r(t-2)}{r+1}. \tag{189}$$

Proof. Define a set of edges $F \subseteq E_G$ by

$$F \stackrel{\mathrm{def}}{=} \bigcup_{e\in I_G^t} \left\{ f \in E_G :\ d_G(e,f) \leqslant \left\lfloor \frac{t-1}{2} \right\rfloor + \left\lfloor \frac{r-1}{2} \right\rfloor \right\}. \tag{190}$$

Also, let $U \stackrel{\mathrm{def}}{=} \bigcup_{f\in F} f \subseteq V_G$ be the set of vertices that belong to an edge in F. Let H be the graph (U,F). For every $C \in \mathfrak{C}_t(G)$ choose an arbitrary vertex $v_C \in e_C$. By the definition of F, for every $u \in U$ there exists $C \in \mathfrak{C}_t(G)$ such that

$$d_G(u, v_C) \leqslant 1 + \left\lfloor \frac{t-1}{2} \right\rfloor + \left\lfloor \frac{r-1}{2} \right\rfloor \leqslant \frac{t+r}{2}.$$

Consequently,

$$|U| \leqslant \sum_{C\in\mathfrak{C}_t(G)} \left| B_G\left(v_C, \frac{t+r}{2}\right)\right| \leqslant d(d-1)^{\frac{t+r}{2}-1}|\mathfrak{C}_t(G)|.$$

By our assumption (188), the graph H is $(1+\eta)$-sparse. An application of Lemma 6.10 to the connected components of H shows that if C_1 and C_2 are distinct cycles of H of length less than t then either they are contained in different connected components of H or

$$d_H(C_1,C_2) \geqslant 1 + \frac{1}{\eta} - 2(t-1) \stackrel{(187)}{=} r. \tag{191}$$

It follows from (191) that any distinct cycles $C_1, C_2 \in \mathfrak{C}_t(G)$ also satisfy $d_G(C_1,C_2) \geqslant r$. Indeed, observe first that for every $C \in \mathfrak{C}_t(G)$ we have $C \subseteq U$. This is true because the diameter (in the metric d_G) of C is at most $\lfloor (|C|-1)/2 \rfloor \leqslant \lfloor (t-1)/2 \rfloor$, and therefore by (190) we have $E_G(C) \subseteq F$. This also shows that $\mathfrak{C}_t(G) = \mathfrak{C}_t(H)$. Now suppose for the sake of obtaining a contradiction that $C_1, C_2 \in \mathfrak{C}_t(G)$ and $0 < d_G(C_1,C_2) \leqslant r-1$. Then there exist $u_1,\ldots,u_r \in V_G$ such that $u_1 \in C_1$, $u_r \in C_2$ and $\{u_{i-1},u_i\} \in E_G$ for every $i \in \{2,\ldots,r\}$. We have $d_G(u_1, e_{C_1}) \leqslant \lfloor (t-1)/2 \rfloor$ and $d_G(u_r, e_{C_2}) \leqslant \lfloor (t-1)/2 \rfloor$. This implies that $d_G(u_i, C_1 \cup C_2) \leqslant \lfloor (t-1)/2 \rfloor + \lfloor (r-1)/2 \rfloor$ for every $i \in \{1,\ldots,r\}$, and therefore by (190) we have $\{u_{i-1},u_i\} \in F$ for every $i \in \{2,\ldots,r\}$, implying that $d_H(C_1,C_2) \leqslant r-1$, in contradiction to (191).

To prove (189), take $u,v \in V_G$ and write $d_G(u,v) = m \leqslant \mathrm{diam}(G)$. Let $\{w_0 = w, w_1, \ldots, w_m = v\} \subseteq V_G$ be distinct vertices such that $\{w_{i-1}, w_i\} \in E_G$ for every $i \in \{1,\ldots,m\}$. We proceed to examine the edges on this path that are not in $E_{L_G^t}$. Thus, suppose that $j_1,\ldots,j_k \in \{0,\ldots,m-1\}$ are such that for every $\ell \in \{1,\ldots,k\}$ we have $\{w_{j_\ell}, w_{j_\ell+1}\} = e_{C_\ell}$ for some $C_\ell \in \mathfrak{C}_t(G)$, and $\{w_{i-1},w_i\} \neq e_C$ for every $i \in \{1,\ldots,m\} \smallsetminus \{j_1,\ldots,j_k\}$ and every $C \in \mathfrak{C}_t(G)$. Since we proved that distinct cycles in $\mathfrak{C}_t(G)$ are at distance at least r (in the metric d_G), it follows that $j_{\ell+1}-j_\ell \geqslant r+1$ for every $\ell \in \{1,\ldots,k-1\}$. Hence $m \geqslant (k-1)(r+1)+1$, or equivalently $k \leqslant (m+r)/(r+1)$. Now, for every $\ell \in \{1,\ldots,k\}$ replace each edge $\{w_{j_\ell}, w_{j_\ell+1}\}$ by the path $C_\ell \setminus \{e_{C_\ell}\}$, whose length is at most $t-2$. By the definition of L_G^t, we thus obtain a new path joining u and v, all of whose edges are edges of L_G^t, and its length is at most $m+k(t-2)$. By substituting the above upper bound on k we conclude that

$$d_H(u,v) \leqslant m + \frac{(m+r)(t-2)}{r+1} \leqslant \frac{t+r-1}{r+1}\cdot \operatorname{diam}(G) + \frac{r(t-2)}{r+1}. \qquad \square$$

Lemma 7.8. *Fix three integers $M,t,d \geqslant 3$ and suppose that G is a connected d-regular simple graph such that $d_G(C_1,C_2) > 8M$ for every distinct $C_1,C_2 \in \mathfrak{C}_t(G)$. Suppose also that*

$$\min_{\substack{S\subseteq G\\ 0<|S|\leqslant n/2}} \frac{E_G(S,V_G\setminus S)}{|S|} \geqslant \frac{1}{M}. \tag{192}$$

Then

$$\min_{\substack{S\subseteq G\\ 0<|S|\leqslant n/2}} \frac{E_{L_G^t}(S,V_G\setminus S)}{|S|} \geqslant \frac{1}{4M}.$$

Proof. Fix $S \subseteq V_G$ with $0<|S|\leqslant |V_G|/2$. Write

$$\mathscr{C} \stackrel{\mathrm{def}}{=} \{C \in \mathfrak{C}_t(G):\ e_C \cap S \neq \emptyset\},$$

and for every $C \in \mathscr{C}$ choose an arbitrary endpoint $u_C \in e_C \cap S$. Define

$$T \stackrel{\mathrm{def}}{=} \{u_C\}_{C\in\mathscr{C}} \subseteq V_G.$$

If $|S| \geqslant 2M|T|$ then

$$\frac{E_{L_G^t}(S,V_G\setminus S)}{|S|} \geqslant \frac{E_G(S,V_G\setminus S)-|T|}{|S|} \stackrel{(192)}{\geqslant} \frac{1}{2M}.$$

So, we may assume from now on that $|S| < 2M|T|$. Because every distinct $x,y\in T$ satisfy $d_G(x,y) > 8M$, the balls $\{B_G(x,4M)\}_{x\in T}$ are disjoint. This implies that if we define

$$W \stackrel{\mathrm{def}}{=} \{x \in T:\ |B_G(x,4M)\cap S| < 4M\},$$

then $|W| > |T|/2$, since otherwise $|S| \geqslant 4M|T\setminus W| \geqslant 2M|T|$.

For every $x \in W$ choose a maximal $k(x) \in \mathbb{N}\cup\{0\}$ for which there exists $w_1(x),\ldots,w_k(x) \in S$ such that

$$\{x,w_1(x)\},\{w_1(x),w_2(x)\},\ldots,\{w_{k-1}(x),w_k(x)\} \in E_{L_G^t} = E_G \setminus I_G^t.$$

In other words, we are considering a maximal path starting from x in the graph $(S, E_{L_G^t}(S))$. Since $|B_G(x,4M)\cap S| < 4M$ it follows that $k(x) \leqslant 4M-1$. Since $w_{k(x)}(x)$ belongs to at most one edge in I_G^t, and $w_{k(x)}(x)$ has $d\geqslant 3$ neighbors in G, it follows that there exists $z(x)\in V_G\smallsetminus\{w_{k(x)-1}(x)\}$ such that $\{w_{k(x)}, z(x)\}\in E_{L_G^t}$. By the maximality of $k(x)$ we necessarily have $\{w_{k(x)}, z(x)\}\in E_{L_G^t}(S, V_G\smallsetminus S)$. Moreover, for distinct $x,y\in W$ we have $\{w_{k(x)}, z(x)\}\neq\{w_{k(y)}, z(y)\}$, since otherwise, because $d_G(x, w_{k(x)}(x)), d_G(y, w_{k(y)}(x))\leqslant 4M-1$ and $d_G(x,z(x)), d_G(y,z(y))\leqslant 4M$, it would follow that $d_G(x,y)\leqslant 8M$.

We have proved that $E_{L_G^t}(S, V_G\smallsetminus S)\geqslant |W| > |T|/2 > |S|/(4M)$, completing the proof of Lemma 7.8. $\square$

Proof of Lemma 3.12. Fix a parameter $M\in(1,\infty)$ that will be determined later. Let $\mathcal{E}_M^1$ denote the family of all connected n-vertex d-regular simple graphs G that satisfy the following conditions

$$\operatorname{diam}(G)\leqslant M\log_d n, \tag{193}$$

and

$$\min_{\substack{S\subseteq G\\ 0<|S|\leqslant n/2}} \frac{E_G(S,V_G\smallsetminus S)}{|S|}\geqslant \frac{1}{M}.$$

By the proofs in [BFdlV82] and [Bol88] we can fix M to be a sufficiently large universal constant such that

$$1-\mathcal{G}_{n,d}\left(\mathcal{E}_M^1\right)\lesssim_d \frac{1}{n}. \tag{194}$$

(The papers [BFdlV82, Bol88] contain much more precise information on the diameter and expansion of random regular graphs, respectively.)

By adjusting the value of the constant $C(d)$ in Lemma 3.12, it suffices to prove the required statement under the assumption $n\geqslant d^{2000M}$. Define $\delta\in(0,1)$ and an integer $t\geqslant 3$ by

$$\delta\stackrel{\mathrm{def}}{=}\frac{21}{\log_d n},\qquad\text{and}\qquad t\stackrel{\mathrm{def}}{=}\left\lfloor\frac{\log_d n}{63}\right\rfloor. \tag{195}$$

Also, set from now on $\varepsilon=1/3$. Let $\mathcal{E}^2$ denote the family of graphs $\mathcal{S}_{\varepsilon,\delta}$ (recall Definition 7.2) and let $\mathcal{E}^3$ denote the family of all connected n-vertex d-regular simple graphs G that satisfy $|\mathfrak{C}_t(G)|\leqslant\sqrt{n}$. By Lemma 7.3 and Lemma 7.4 we have

$$1-\mathcal{G}_{n,d}\left(\mathcal{E}^2\right)\lesssim_d\frac{1}{n^{2/3}}\qquad\text{and}\qquad 1-\mathcal{G}_{n,d}\left(\mathcal{E}^3\right)\lesssim_d\frac{1}{\sqrt[3]{n}} \tag{196}$$

By virtue of (194) and (196), if we denote $\mathscr{L} \stackrel{\mathrm{def}}{=} \mathcal{E}_M^1 \cap \mathcal{E}^2 \cap \mathscr{E}^3$ then

$$1 - \mathcal{G}_{n,d}(\mathscr{L}) \lesssim_d \frac{1}{\sqrt[3]{n}}.$$

We will now show that for a large enough universal constant $K \in (0, \infty)$ the graph family $\mathscr{L}$ can serve as the graph family $\mathscr{L}_K^{n,d}$ of Lemma 3.12.

Take $G \in \mathscr{L}$ and write $I = I_G^t$ and $L = L_G^t$. By the definition of $\mathcal{E}^3$ we have $|I| \leqslant \sqrt{n}$, as required. By the definition of $\mathcal{S}_{\varepsilon,\delta}$, for every $S \subseteq V_G$ with $|S| \leqslant n^{2/3}$ we have $|E_G(S)| \leqslant (1+\delta)|S|$. Since $n \geqslant d^{200}$, by the definitions (195) we have $n^{2/3} \geqslant d(d-1)^{t-1}\sqrt{n}$ and $\delta \leqslant 1/(3t-3)$. Since $|\mathfrak{C}_t(G)| \leqslant \sqrt{n}$, this implies that assumption (188) of Lemma 7.7 is satisfied, with $r = t$. It therefore follows from Lemma 7.7 that $d_G(C_1, C_2) \geqslant t$ for distinct $C_1, C_2 \in \mathfrak{C}_t(G)$ and

$$\mathrm{diam}(L) \stackrel{(189)}{\leqslant} \frac{2t-1}{t+1}\mathrm{diam}(G) + \frac{t(t-2)}{t+1} \stackrel{(193)\wedge(195)}{\leqslant} 3M \log_d n. \tag{197}$$

This shows that assertion (b) of Definition 3.11 holds true provided $K \geqslant 3M$. Since by (195) and the assumption $n \geqslant d^{2000M}$ we have $t > 8M$, we may also use Lemma 7.8 to deduce that

$$\min_{\substack{S \subseteq G \\ 0 < |S| \leqslant n/2}} \frac{E_L(S, V_G \smallsetminus S)}{|S|} \geqslant \frac{1}{4M}.$$

Hence, assertion (c) of Definition 3.11 holds true provided $K \geqslant 4M$.

By the definition of L we have

$$\mathrm{girth}(L) \geqslant t \stackrel{(195)\wedge(197)}{\geqslant} \frac{\mathrm{diam}(G)}{378M}. \tag{198}$$

So, assertion (1) of Definition 3.10 holds true provided $K \geqslant 378M$.

It remains to prove that if K is large enough then the graph L satisfies assertion (2) of Definition 3.10. Suppose that $S_0 \subseteq \Sigma(V_G)$ satisfies $|S_0| \leqslant \sqrt{n}$. Let $S \subseteq V_G$ denote the union of $S_0 \cap V_G$ with the endpoints of all the edges in E_L whose corresponding unit interval in the one-dimensional simplicial complex $\Sigma(L)$ contains a point from $S_0 \smallsetminus V_G$. Thus $|S| \leqslant 2\sqrt{n}$. Apply Lemma 7.5 with $S = T$ and $r = \lfloor (\log_d n)/36 \rfloor$. We obtain $U \subseteq V_G$ with $U \supseteq S$ such that if we let H denote the graph $(U, E_L(U))$ then

$$\mathrm{diam}(H) \leqslant 6r + 2\,\mathrm{diam}(L) \stackrel{(197)}{\leqslant} 4M \log_d n, \tag{199}$$

and for every $a,b\in S$

$$
\begin{aligned}
d_L(a,b) \leqslant d_H(a,b) &\leqslant 2\left(\frac{\operatorname{diam}(L)}{r}+1\right)d_L(a,b)\\
&\overset{(197)}{\leqslant} 2\left(\frac{3M\log_d n}{\lfloor(\log_d n)/36\rfloor}+1\right)d_L(a,b)\lesssim Md_L(a,b). \quad (200)
\end{aligned}
$$

Since $U\supseteq S$, and S contains $S_0\cap V_G$ and the endpoints of all the edges of $\Sigma(L)$ that contain points in $S_0\smallsetminus V_G$, it follows from (200) that

$$\forall\, x,y\in S_0,\qquad d_{\Sigma(L)}(x,y)\leqslant d_{\Sigma(H)}(x,y)\lesssim Md_{\Sigma(L)}(x,y). \qquad (201)$$

Also, by (183), and using the fact that $n\geqslant d^{2000M}$, we have

$$|U|\leqslant 2\sqrt{n}\left(n^{1/12}+\operatorname{diam}(L)\right)\overset{(197)}{\leqslant} 2n^{7/12}+2M\sqrt{n}\cdot\log_d n\leqslant n^{2/3}.$$

Since G belongs to $\mathcal{S}_{\varepsilon,\delta}$, it follows that H is $(1+\delta)$-sparse. By Corollary 6.6 we conclude that

$$c_1\left(\Sigma(H)\right)\lesssim 1+\delta\operatorname{diam}(H)\overset{(195)\wedge(199)}{\leqslant} 1+84M.$$

Due to (201) we therefore have

$$c_1\left(S_0,\frac{2\pi}{\operatorname{girth}(L)}\cdot d_{\Sigma(L)}\right)\lesssim M,$$

and by Corollary 3.8,

$$c_1\left(\mathbf{Cone}\left(S_0,\frac{2\pi}{\operatorname{girth}(L)}\cdot d_{\Sigma(L)}\right)\right)\lesssim M.$$

This concludes the proof of assertion (2) of Definition 3.10 provided K is a sufficiently large multiple of (the universal constant) M. $\square$

7.2. **Proof of Proposition 3.15.** We shall continue using here the notation and assumptions that were used in the proof of Lemma 3.12. Suppose that $G\in\mathscr{L}$. We have already seen that if G is distributed according to $\mathcal{G}_{n,d}$ then this happens with probability at least $1-a(d)/\sqrt[3]{n}$ for some $a(d)\in(0,\infty)$.

Recalling the definition of t in (195), define $A_1,A_2\subseteq\Sigma(G)$ as follows.

$$A_1\overset{\mathrm{def}}{=}\bigcup_{e\in I_G^t}\left\{x\in\Sigma(G):\ d_{\Sigma(G)}(x,e)\leqslant t\right\},$$

and

$$A_2\overset{\mathrm{def}}{=}\Sigma(G)\smallsetminus\bigcup_{e\in I_G^t}\left\{x\in\Sigma(G):\ d_{\Sigma(G)}(x,e)\leqslant\frac{t}{2}\right\}.$$

Then $A_1 \cup A_2 = \Sigma(G)$ and

$$d_{\Sigma(G)}\left(A_1 \smallsetminus A_2, A_2 \smallsetminus A_1\right) \geqslant \frac{t}{2} \overset{(195)}{\gtrsim} \log_d n. \tag{202}$$

As in the proof of Lemma 3.12, denote $I \stackrel{\mathrm{def}}{=} I_G^t$ and $L \stackrel{\mathrm{def}}{=} L_G^t$. We also let $\Phi \subseteq \Sigma(G)$ be the union of all the unit intervals corresponding to the edges in I. Thus A_1 is the t-neighborhood of Φ in $\Sigma(G)$ and A_2 is the complements of the $(t/2)$-neighborhood of Φ in $\Sigma(G)$.

We claim that

$$\forall\, x, y \in A_2, \qquad d_{\Sigma(G)}(x,y) \leqslant d_{\Sigma(L)}(x,y) \leqslant 3d_{\Sigma(G)}(x,y). \tag{203}$$

Since $E_L \subseteq E_G$, only the rightmost inequality in (203) requires proof. Fix $x, y \in A_2$ and let $P_{x,y} : [0, d_{\Sigma(G)}(x,y)] \to \Sigma(G)$ be a geodesic joining x and y in $\Sigma(G)$. If $P_{x,y} \subseteq \Sigma(L)$ then $d_{\Sigma(G)}(x,y) = d_{\Sigma(L)}(x,y)$. The situation is therefore only interesting when $P_{x,y} \cap \Phi \neq \emptyset$. In this case, since the distances of x and y from Φ are at least $t/2$, the length of $P_{x,y}$ must be at least t. Let $J_1, \ldots, J_k \subseteq [0, d_{\Sigma(G)}(x,y)]$ be a maximal collection of intervals such that $P_{x,y}(J_i) \in I$ for every $i \in \{1, \ldots, k\}$. We have shown in the proof of Lemma 3.12 that (using Lemma 7.7) we have $d_{\Sigma(G)}(P_{x,y}(J_i), P_{x,y}(J_j)) \geqslant t$ for every distinct $i, j \in \{1, \ldots, k\}$. This means that the length of $P_{x,y}$ is at least $(k-1)t$, or equivalently that $k \leqslant 1 + d_{\Sigma(G)}(x,y)/t$. Each edge $P_{x,y}(J_i)$ lies on a cycle $C_i \in \mathfrak{C}_t(G)$, and all the other edges of C_i are in E_L. Therefore, if we replace each edge $P_{x,y}(J_i)$ by the path $C_i \smallsetminus P_{x,y}(J_i)$ (whose length is at most $t-1$), we will obtain a path joining x and y in $\Sigma(H)$ of length at most

$$\begin{aligned} d_{\Sigma(G)}(x,y) + k(t-1) &\leqslant d_{\Sigma(G)}(x,y) + \left(1 + \frac{d_{\Sigma(G)}(x,y)}{t}\right)(t-1) \\ &\leqslant 2d_{\Sigma(X)}(x,y) + t \leqslant 3d_{\Sigma(X)}(x,y), \end{aligned}$$

where we used the fact that $d_{\Sigma(G)}(x,y) \geqslant t$. This completes the verification of (203).

The scaling factor σ of Proposition 3.15 will be chosen to be

$$\sigma \stackrel{\mathrm{def}}{=} \frac{2\pi}{\mathrm{girth}(L)} \overset{(197)\wedge(198)}{\asymp} \frac{1}{\log_d n}. \tag{204}$$

with this choice, due to (202) assertion (II) of Proposition 3.15 holds true. We have already proved in Lemma 3.12 that $L \in \mathcal{F}_K$, and therefore $\mathbf{Cone}(\Sigma(L), \sigma d_{\Sigma(L)})$ is isometric to a subset of $\mathscr{X}_K$. By (203) and Fact 3.5 we therefore have

$$c_{\mathscr{X}_K}\left(\mathbf{Cone}(A_2, \sigma d_{\Sigma(G)})\right) \leqslant c_{\mathbf{Cone}(\Sigma(L), \sigma d_{\Sigma(L)})}\left(\mathbf{Cone}(A_2, \sigma d_{\Sigma(G)})\right) \leqslant 3,$$

thus proving assertion (IV) of Proposition 3.15.

To prove assertion (V) of Proposition 3.15 we define an embedding $f:\Sigma(G)\to \mathbf{Cone}\left(\Sigma(G)\right)$ by $f(x)=(1/\sqrt{2},x)$. For every $x,y\in\Sigma(G)$,

$$d_{\mathbf{Cone}\left(\Sigma(G),\sigma d_{\Sigma(G)}\right)}(f(x),f(y))\stackrel{(43)}{=}\sqrt{1-\cos\left(\min\{\pi,\sigma d_{\Sigma(G)}(x,y)\}\right)}$$
$$\stackrel{(48)\wedge(204)}{\asymp}\min\left\{\pi,\frac{d_{\Sigma(G)}(x,y)}{\log_d n}\right\}\stackrel{(193)}{\asymp} d_{\Sigma(G)}(x,y).$$

All that remains is to prove assertion (III) of Proposition 3.15. Define $S=A_1\cap V_G$ and $T=\Phi\cap V_G=\bigcup_{e\in I}e$. Thus $|T|\leqslant 2|I|\leqslant 2\sqrt{n}$. By the definition of A_1, condition (182) of Lemma 7.5 holds true with $r=t$. Consequently, by Lemma 7.5 there exists $U\subseteq V_G$ with $U\supseteq S$ such that if we let H denote the graph $(U,E_G(U))$ then for every $a,b\in S$,

$$d_G(a,b)\leqslant d_H(a,b)\leqslant 2\left(\frac{\operatorname{diam}(G)}{t}+1\right)d_G(a,b)\stackrel{(193)\wedge(195)}{\lesssim} d_G(a,b).$$

Since $U\supseteq S$ and $S=A_1\cap V_G$, it follows that

$$\forall\, x,y\in A_2,\qquad d_{\Sigma(G)}(x,y)\asymp d_{\Sigma(H)}(x,y). \tag{205}$$

Moreover, using the assumption $n\geqslant d^{2000M}$, it follows from (183) that

$$|U|\leqslant |T|\left(d^{3t}+\operatorname{diam}(G)\right)\stackrel{(193)\wedge(195)}{\leqslant}2\sqrt{n}\left(n^{1/21}+M\log_d n\right)\leqslant n^{2/3}.$$

recalling that G belongs to $\mathcal{S}_{\varepsilon,\delta}$, we conclude that H is $(1+\delta)$-sparse, and therefore by Corollary 6.6,

$$c_1\left(A_2,d_{\Sigma(G)}\right)\stackrel{(205)}{\lesssim}c_1\left(\Sigma(H),d_{\Sigma(H)}\right)\lesssim 1+\delta\operatorname{diam}(H)\stackrel{(195)\wedge(199)}{\lesssim}1.$$

Now assertion (III) of Proposition 3.15 follows by Corollary 3.8. □

8. Are two stochastically independent random graphs expanders with respect to each other?

Here we present partial progress towards (a positive solution of) Question 2.4. Proposition 8.1 below is based on ideas of U. Feige.

Proposition 8.1. *For $n\in\mathbb{N}$ even let G,H be two i.i.d. random graphs sampled from $\mathcal{G}_{n,3}$. Then with probability that tends to 1 as $n\to\infty$, for every permutation $\pi\in S_n$ we have*

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_H(\pi(i),\pi(j))^2\lesssim\frac{1}{n}\sum_{\{i,j\}\in E_G}d_H(\pi(i),\pi(j))^2. \tag{206}$$

A positive answer to Question 2.4 would require proving (206) when G, H are independent random 3-regular graphs of possibly different cardinalities, say G sampled from $\mathcal{G}_{n,3}$ and H sampled from $\mathcal{G}_{m,3}$, and with the permutation π of Proposition 8.1 replaced with a general mapping $f : \{1, \ldots, n\} \to \{1, \ldots, m\}$. Note that for fixed $m \in \mathbb{N}$ we have $\gamma_+(G, d_H^2) \lesssim (\log m)^2$ asymptotically almost surely. Indeed, G is asymptotically almost surely an expander, and by Bourgain's embedding theorem [Bou85], the metric space $(\{1, \ldots, m\}, d_H)$ embeds into ℓ_2 with bi-Lipschitz distortion $O(\log m)$. When $m > n$ one can show that asymptotically almost surely

$$\gamma_+(G, d_H)^2 \lesssim \max\left\{1, \left(\frac{\log n}{\log(m/n)}\right)^2\right\}. \tag{207}$$

The validity of (207) follows from the fact that G is asymptotically almost surely an expander, combined with an application of Lemma 7.3 and Corollary 6.6 to the random graph H.

Below we shall use the following version of Chernoff's bound (see e.g [AS08, Thm. A.1.12]). Suppose that $X_1, \ldots, X_n$ are i.i.d. random variables taking values in $\{0, 1\}$ and write $\Pr[X_i = 1] = p$. Then for every $\beta \in (1, \infty)$ we have

$$\Pr\left[\sum_{i=1}^n X_i > \beta p n\right] < \left(\frac{e^{\beta-1}}{\beta^\beta}\right)^{pn}. \tag{208}$$

Proof of Proposition 8.1. By [BFdlV82] with probability that tends to 1 as $n \to \infty$ we have $\operatorname{diam}(H) \lesssim \log n$. The left hand side of (206) therefore asymptotically almost surely satisfies

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n d_H(\pi(i), \pi(j))^2 \lesssim (\log n)^2.$$

Since there are $n!$ possible permutations $\pi \in S_n$, it suffices to prove that there exists a universal constant $c \in (0, \infty)$ such that for every fixed permutation $\pi \in S_n$ we have

$$\mathcal{G}_{n,3} \times \mathcal{G}_{n,3}\left(\left\{G, H \in \mathbb{G}_n : \frac{1}{n}\sum_{\{i,j\} \in E_G} d_H(\pi(i), \pi(j))^2 \geqslant c(\log n)^2\right\}\right) \geqslant 1 - \frac{o(1)}{n!}. \tag{209}$$

Because the distribution of the metric d_H is invariant under permutations, it suffices to prove (209) when π is the identity permutation.

To this end, given a 3-regular graph $H \in \mathbb{G}_n$, consider the following subset of the unordered pairs of elements of $\{1,\ldots,n\}$.

$$N_H \stackrel{\mathrm{def}}{=} \left\{\{i,j\} \in \binom{\{1,\ldots,n\}}{2}:\ d_H(i,j) \leqslant \frac{\log n}{16}\right\}.$$

Since H is 3-regular we have $|N_H| \leqslant \frac{3}{2}n^{17/16}$.

In order to prove (209) (with $c = 1/16$ and π the identity mapping) it suffice to prove that for every fixed 3-regular graph $H \in \mathbb{G}_n$ we have

$$\mathcal{G}_{n,3}\left(\left\{G \in \mathbb{G}_n:\ |E_G \cap N_H| > \frac{4n}{3}\right\}\right) \leqslant \frac{o(1)}{n!}. \tag{210}$$

By (172) it suffices to prove (210) in the pairing model, i.e.,

$$\mathcal{P}_{n,3}\left(\left\{G \in \mathbb{G}_n:\ |E_G \cap N_H| > \frac{4n}{3}\right\}\right) \leqslant \frac{o(1)}{n!}. \tag{211}$$

Assume from now on that n is divisible by 4. The proof for general even n follows mutatis mutandis from the argument below, the only difference being that one needs to round certain numbers to their nearest integers.

Recall that in the pairing model $\mathcal{P}_{n,3}$ we choose a matching M of $P = \{1,\ldots,n\} \times \{1,2,3\}$ uniformly at random, and "project it" onto $\{1,\ldots,n\}$ so as to get a random 3-regular graph $G(M)$. It will be convenient to order the $\binom{3n}{2}$ pairs in $\binom{P}{2}$ arbitrarily. Once this is done, the uniformly random matching M can be obtained by choosing its edges sequentially, where at each step we choose an additional edge uniformly at random from the unused pairs. Fix a 3-regular graph $H \in \mathbb{G}_n$ and let Y_i^H be the $\{0,1\}$-valued random variable that takes the value 1 if and only if ith edge is in N_H. Note that Y_i^H is obtained by flipping a biased coin with success probability p_i which is itself a random variable depending on the first $i-1$ edges of M. However, for $i \leqslant 5n/4$ we have

$$p_i \leqslant \frac{|N_H|}{\binom{3n-2(i-1)}{2}} \leqslant \frac{2|N_H|}{(3n-2i)^2} \leqslant \frac{8|N_H|}{n^2} \leqslant \frac{12n^{17/16}}{n^2} = \frac{12}{n^{15/16}} \stackrel{\mathrm{def}}{=} p.$$

We proceed by a standard coupling argument. Having sampled $Y_1^H,\ldots,Y_{5n/4}^H$, we sample a sequence $X_1,\ldots,X_{5n/4}$ of $\{0,1\}$-valued random variables as follows. If $Y_i^H = 1$ then $X_i = 1$, and if $Y_i^H = 0$ then X_i is obtained by tossing an independent coin with success probability $(p-p_i)/(1-p_i)$. Then $\Pr[X_i = 1] = p$, and because the coins that we used are independent, $X_1,\ldots,X_{5n/4}$ are independent random variables. Moreover, we have the point-wise inequality $X_i \geqslant Y_i^H$ for every

$i \leqslant 5n/4$. Now,

$$\mathcal{P}_{n,3}\left(\left\{G \in \mathbb{G}_n : \ |E_G \cap N_H| > \frac{4n}{3}\right\}\right) = \Pr\left[\sum_{i=1}^{3n/2} Y_i^H > \frac{4n}{3}\right]$$
$$\leqslant \Pr\left[\sum_{i=1}^{5n/4} Y_i^H > \frac{13n}{12}\right] \leqslant \Pr\left[\sum_{i=1}^{5n/4} X_i^H > \frac{13n}{12}\right].$$

Hence by (208) with $\beta = \frac{13n^{15/16}}{180}$, we obtain the bound

$$\mathcal{P}_{n,3}\left(\left\{G \in \mathbb{G}_n : \ |E_G \cap N_H| > \frac{4n}{3}\right\}\right) \leqslant \left(\frac{180e}{13}\right)^{\frac{13n}{12}} n^{-\frac{65}{64}n} = \frac{o(1)}{n!},$$

implying the desired estimate (211). □

Acknowledgements. We are grateful to Uriel Feige, Konstantin Makarychev and Yuval Peres for helpful discussions, and to Jon Kleinberg for formulating Question 2.4. We also thank Takefumi Kondo for sending us a preliminary version of [Kon12]. M. M. was supported by ISF grants 221/07 and 93/11, BSF grant 2010021, and NSF grants CCF-0832797 and DMS-0835373. Part of this work was completed while M. M. was visitor at Microsoft Research and a member of the Institute for Advanced Study at Princeton. A. N. was supported by NSF grant CCF-0832795, BSF grant 2010021, the Packard Foundation and the Simons Foundation. Part of this work was completed while A. N. was a Visiting Fellow at Princeton University.

References

[ABLT06] S. Arora, B. Bollobás, L. Lovász, and I. Tourlakis. Proving integrality gaps without knowing the linear program. *Theory Comput.*, 2:19–51, 2006.

[ABN86] A. D. Aleksandrov, V. N. Berestovskiĭ, and I. G. Nikolaev. Generalized Riemannian spaces. *Uspekhi Mat. Nauk*, 41(3(249)):3–44, 240, 1986.

[ALN+12] S. Arora, L. Lovász, I. Newman, Y. Rabani, Y. Rabinovich, and S. Vempala. Local versus global properties of metric spaces. *SIAM J. Comput.*, 41(1):250–271, 2012.

[AM85] N. Alon and V. D. Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985.

[AS08] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[Ber83] V. N. Berestovskiĭ. Borsuk's problem on metrization of a polyhedron. *Dokl. Akad. Nauk SSSR*, 268(2):273–277, 1983.

[BFdlV82] B. Bollobás and W. Fernandez de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982.

[BGS07] K. Barhum, O. Goldreich, and A. Shraibman. On approximating the average distance between points. In M. Charikar, K. Jansen, O. Reingold, and J. D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. Proccedings of 10th International Workshop APPROX 2007 and 11th International Workshop RANDOM 2007*, volume 4627 of *Lecture Notes in Computer Science*, pages 296–310. Springer, 2007.

[BH99] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[BLMN05] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric Ramsey-type phenomena. *Ann. of Math. (2)*, 162(2):643–709, 2005.

[BMW86] J. Bourgain, V. Milman, and H. Wolfson. On type of metric spaces. *Trans. Amer. Math. Soc.*, 294(1):295–317, 1986.

[Bol88] B. Bollobás. The isoperimetric number of random regular graphs. *European J. Combin.*, 9(3):241–244, 1988.

[Bou85] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[Che70] J. Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Papers dedicated to Salomon Bochner, 1969)*, pages 195–199. Princeton Univ. Press, Princeton, N. J., 1970.

[DL97] M. M. Deza and M. Laurent. *Geometry of cuts and metrics*, volume 15 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1997.

[Enf76] P. Enflo. Uniform homeomorphisms between Banach spaces. In *Séminaire Maurey-Schwartz (1975–1976), Espaces, L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. No. 18*, page 7. Centre Math., École Polytech., Palaiseau, 1976.

[Gro83] M. Gromov. Filling Riemannian manifolds. *J. Differential Geom.*, 18(1):1–147, 1983.

[Gro01] M. Gromov. CAT(κ)-spaces: construction and concentration. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 280(Geom. i Topol. 7):100–140, 299–300, 2001.

[Gro03] M. Gromov. Random walk in random groups. *Geom. Funct. Anal.*, 13(1):73–146, 2003.

[Hei80] S. Heinrich. Ultraproducts in Banach space theory. *J. Reine Angew. Math.*, 313:72–104, 1980.

[HLW06] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.

[IKN12] H. Izeki, T. Kondo, and S. Nayatani. N-step energy of maps and the fixed-point property of random groups. *Groups Geom. Dyn.*, 6(4):701–736, 2012.

[IN05] H. Izeki and S. Nayatani. Combinatorial harmonic maps and discrete-group actions on Hadamard spaces. *Geom. Dedicata*, 114:147–188, 2005.

[Ind99] P. Indyk. Sublinear time algorithms for metric space problems. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 428–432 (electronic). ACM, New York, 1999.

[Jos97] J. Jost. *Nonpositive curvature: geometric and analytic aspects*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997.

[Kap09] M. Kapovich. *Hyperbolic manifolds and discrete groups*. Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2009. Reprint of the 2001 edition.

[KN06] S. Khot and A. Naor. Nonembeddability theorems via Fourier analysis. *Mathematische Annalen*, 334(4):821–852, 2006.

[Kon12] T. Kondo. CAT(0) spaces and expanders. *Math. Z.*, 271(1-2):343–355, 2012.

[KV05] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of the 46th Annual IEEE Conference on Foundations of Computer Science (FOCS 2005)*, pages 53–62. 2005.

[KY06] G. Kasparov and G. Yu. The coarse geometric Novikov conjecture and uniform convexity. *Adv. Math.*, 206(1):1–56, 2006.

[Laf08] V. Lafforgue. Un renforcement de la propriété (T). *Duke Math. J.*, 143(3):559–602, 2008.

[Laf09] V. Lafforgue. Propriété (T) renforcée Banachique et transformation de Fourier rapide. *J. Topol. Anal.*, 1(3):191–206, 2009.

[Laf10] V. Lafforgue. Propriété (T) renforcée et conjecture de Baum-Connes. In *Quanta of maths*, volume 11 of *Clay Math. Proc.*, pages 323–345. Amer. Math. Soc., Providence, RI, 2010.

[LLR95] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[Mar88] G. A. Margulis. Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators. *Problemy Peredachi Informatsii*, 24(1):51–60, 1988.

[Mat97] J. Matoušek. On embedding expanders into ℓ_p spaces. *Israel J. Math.*, 102:189–197, 1997.

[McK81] B. D. McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the Twelfth Southeastern Conference on Combinatorics, Graph Theory and Computing, Vol. II (Baton Rouge, La., 1981)*, volume 33, pages 213–223. 1981. ISSN 0384-9864.

[MN04] M. Mendel and A. Naor. Euclidean quotients of finite metric spaces. *Adv. Math.*, 189(2):451–494, 2004.

[MN10] M. Mendel and A. Naor. Towards a calculus for non-linear spectral gaps. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 236–255. 2010.

[MN13] M. Mendel and A. Naor. Spectral calculus and Lipschitz extension for barycentric metric spaces. *Anal. Geom. Metr. Spaces*, 1:163–199, 2013.

[MN14] M. Mendel and A. Naor. Nonlinear spectral calculus and super-expanders. *Publ. Math. Inst. Hautes Études Sci.*, 119(1):1–95, 2014.

[MS05] P. Mahlmann and C. Schindelhauer. Peer-to-peer networks based on random transformations of connected regular undirected graphs. In *Proceedings of the seventeenth annual ACM symposium on Parallelism in algorithms and architectures*, SPAA'05. 2005.

[MWW04] B. D. McKay, N. C. Wormald, and B. Wysocka. Short cycles in random regular graphs. *Electron. J. Combin.*, 11(1):Research Paper 66, 12 pp. (electronic), 2004.

[Nao10] A. Naor. L_1 embeddings of the Heisenberg group and fast estimation of graph isoperimetry. In *Proceedings of the International Congress of Mathematicians. Volume III*, pages 1549–1575. Hindustan Book Agency, New Delhi, 2010.

[NS11] A. Naor and L. Silberman. Poincaré inequalities, embeddings, and wild groups. *Compos. Math.*, 147(5):1546–1572, 2011.

[Oza04] N. Ozawa. A note on non-amenability of $\mathscr{B}(l_p)$ for $p = 1, 2$. *Internat. J. Math.*, 15(6):557–565, 2004.

[Pan09] P. Pansu. Superrigidité géométique et applications harmoniques. In *Géométries à courbure négative ou nulle, groupes discrets et rigidités*, volume 18 of *Sémin. Congr.*, pages 373–420. Soc. Math. France, Paris, 2009.

[Pic08] M. Pichot. Harmonic analysis from quasi-periodic domains. *Israel J. Math.*, 167:63–90, 2008.

[Pis86] G. Pisier. Probabilistic methods in the geometry of Banach spaces. In *Probability and analysis (Varenna, 1985)*, volume 1206 of *Lecture Notes in Math.*, pages 167–241. Springer, Berlin, 1986.

[Pis10] G. Pisier. Complex interpolation between Hilbert, Banach and operator spaces. *Mem. Amer. Math. Soc.*, 208(978):vi+78, 2010.

[RVW02] 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.

[Sch38] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.

[Stu03] K.-T. Sturm. Probability measures on metric spaces of nonpositive curvature. In *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002)*, volume 338 of *Contemp. Math.*, pages 357–390. Amer. Math. Soc., Providence, RI, 2003.

[Wan98] M.-T. Wang. A fixed point theorem of discrete group actions on Riemannian manifolds. *J. Differential Geom.*, 50(2):249–267, 1998.

[Wan00] M.-T. Wang. Generalized harmonic maps and representations of discrete groups. *Comm. Anal. Geom.*, 8(3):545–563, 2000.

[Woj91] P. Wojtaszczyk. *Banach spaces for analysts*, volume 25 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1991.

[Wor81] N. C. Wormald. The asymptotic connectivity of labelled regular graphs. *J. Combin. Theory Ser. B*, 31(2):156–167, 1981.

[Wor99] N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

[WW75] J. H. Wells and L. R. Williams. *Embeddings and extensions in analysis.* Springer-Verlag, New York, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.

Mathematics and Computer Science Department, The Open University of Israel, 1 University Road, P.O. Box 808 Raanana 43107, Israel
E-mail address: mendelma@gmail.com

Courant Institute, New York University, 251 Mercer Street, New York NY 10012, USA
E-mail address: naor@cims.nyu.edu