Algorithmica manuscript No.
(will be inserted by the editor)

Effect of Gromov-hyperbolicity Parameter on Cuts and Expansions in Graphs and Some Algorithmic Implications

Bhaskar DasGupta · Marek Karpinski · Nasim Mobasheri · Farzane Yahyanejad

Received: date / Accepted: date

Abstract δ-hyperbolic graphs, originally conceived by Gromov in 1987, occur often in many network applications; for fixed δ, such graphs are *simply called hyperbolic graphs* and include non-trivial interesting classes of "non-expander" graphs. The main motivation of this paper is to investigate the effect of the hyperbolicity measure δ on expansion and cut-size bounds on graphs (here δ need *not* be a constant), and the asymptotic ranges of δ for which these results may provide *improved* approximation algorithms for related combinatorial problems. To this effect, we provide *constructive* bounds on node expansions for δ-hyperbolic graphs as a function of δ, and show that many witnesses (subsets of nodes) for such expansions can be computed efficiently even if the witnesses are required to be nested or sufficiently distinct from each other. To the best of our knowledge, these are the first such constructive bounds proven. We also show how to find a large family of s-t cuts with *relatively small* number of cut-edges when s and t are sufficiently far apart. We then provide algorithmic consequences of these bounds and their related proof techniques for two problems for δ-hyperbolic graphs (*where δ is a function f of the number of nodes*,

Bhaskar DasGupta
Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA
Tel.: +312-255-1319
Fax: +312-413-0024
E-mail: bdasgup@uic.edu

Marek Karpinski
Department of Computer Science, University of Bonn, Bonn 53113, Germany E-mail: marek@cs.uni-bonn.de

Nasim Mobasheri
Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA
E-mail: nmobas2@uic.edu

Farzaneh Yahyanejad Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA
E-mail: fyahya2@uic.edu

the exact nature of growth of f being dependent on the particular problem considered).

Keywords Gromov hyperbolicity · Node expansion · Minimum cuts · Approximation algorithms

PACS 02.10.Ox · 89.20.Ff · 02.40.Pc

Mathematics Subject Classification (2000) MSC 68Q25 · MSC 68W25 · MSC 68W40 · MSC 05C85

1 Introduction

Useful insights for many complex systems such as the world-wide web, social networks, metabolic networks, and protein-protein interaction networks can often be obtained by representing them as *parameterized* networks and analyzing them using graph-theoretic tools. Some standard measures used for such investigations include degree based measures (*e.g.*, maximum/minimum/average degree or degree distribution) connectivity based measures (*e.g.*, clustering coefficient, claw-free property, largest cliques or densest sub-graphs), and geodesic based measures (*e.g.*, diameter or betweenness centrality). It is a standard practice in theoretical computer science to investigate and categorize the computational complexities of combinatorial problems in terms of ranges of these parameters. For example:

- Bounded-degree graphs are known to admit improved approximation as opposed to their arbitrary-degree counter-parts for many graph-theoretic problems.
- Claw-free graphs are known to admit improved approximation as opposed to general graphs for graph-theoretic problems such as the maximum independent set problem.

In this paper we consider a *topological* measure called *Gromov-hyperbolicity* (or, simply hyperbolicity for short) for undirected unweighted graphs that has recently received significant attention from researchers in both the graph theory and the network science community. This hyperbolicity measure δ was originally conceived in a somewhat different group-theoretic context by Gromov [20]. The measure was first defined for *infinite* continuous metric space via properties of geodesics [10], but was later also adopted for *finite* graphs. Lately, there have been a surge of theoretical and empirical works measuring and analyzing the hyperbolicity of networks, and many *real-world* networks,such as the following, have been reported (either theoretically or empirically) to be δ-hyperbolic for $\delta = O(1)$:

- "preferential attachment" scale-free networks with appropriate scaling (normalization) [21],
- networks of high power transceivers in a wireless sensor network [2],
- communication networks at the IP layer and at other levels [28], and

► an assorted set of biological and social networks [1].

Moreover, extreme congestion at a small number of nodes in a large traffic network that uses the shortest-path routing was shown in [22] to be caused due to a small value of δ of the network. On the other hand, theoretical investigations have revealed that *expanders*, *vertex-transitive* graphs and (for certain parameter ranges) classical *Erdös-Rényi* random graphs are δ-hyperbolic only for $\delta = \omega(1)$ [6–8, 25].

A major motivation for this paper is a question of the following type[1]:

> "*What is the effect of the hyperbolicity measure δ on expansion and cut-size bounds on graphs (where δ is a free parameter and* not *a necessarily a constant)? For what asymptotic ranges of values of δ can these bounds be used to obtain improved approximation algorithms for related combinatorial problems?*"

Since arbitrarily large δ leads to the class of *all possible* graphs, investigations of this type may eventually provide insights or characterizations of hard graph instances for combinatorial problems via different asymptotic ranges of values of δ. To this effect, in this paper we further investigate the non-expander properties of hyperbolic networks beyond what is shown in [6, 25] and provide constructive proofs of *witnesses* (subsets of nodes) satisfying certain expansion or cut-size bounds. We also provide some algorithmic consequences of these bounds and their related proof techniques for two problems related to cuts and paths for graphs. *A more detailed list of our results is deferred until Section 2 after the basic definitions and notations.*

1.1 Basic Notations and Assumptions

We use the following notations and terminologies throughout the paper. We will simply write log to refer to logarithm base 2. Our basic input is an ordered triple $\langle G, d, \delta \rangle$ denoting the given *connected undirected unweighted* graph $G = (V, E)$ of hyperbolicity δ in which every node has a degree of at most $d > 2$. We will always use the variable m and n to denote the number of edges and the number of nodes, respectively, of the given input graph. *Throughout the paper, we assume that n is always sufficiently large.* For notational convenience, *we will ignore floors and ceilings of fractional values* in our theorems and proofs, *e.g.*, we will simply write $n/3$ instead of $\lfloor n/3 \rfloor$ or $\lceil n/3 \rceil$, since this will have *no* effect on the *asymptotic* nature of the bounds. *We will also make no serious effort to optimize the constants that appear in the bounds in our theorems and proofs.* In addition, the following notations will be used throughout the paper:

► $|\mathcal{P}|$ is the *length* (number of edges) of a path $\mathcal{P}$ of a graph.

[1] This is in contrast to many research works in this area where one studies the properties of δ-hyperbolic graphs assuming δ to be fixed.

- $\overline{u,v}$ is a *shortest path* between nodes u and v. In our proofs, any shortest path can be selected but, once selected, the *same* shortest path *must* be used in the remaining part of the analysis.
- $\text{dist}_H(u,v)$ is the distance (number of edges in a shortest path) between nodes u and v in a graph H (and is ∞ if there is no path between u and v in H).
- $D(H) = \max\limits_{u,v \in V'} \{\text{dist}_H(u,v)\}$ is the *diameter* of the graph $H = (V', E')$. Thus, in particular, for our input $\langle G, d, \delta \rangle$ there exists two nodes p and q such that $\text{dist}_G(p,q) = D(G) \geq \log_d n$.
- For a subset S of nodes of the graph $H = (V', E')$, the *boundary* $\partial_H(S)$ of S is the set of nodes in $V \setminus S$ that are connected to *at least* one node in S, *i.e.*,

  $$\partial_H(S) = \{u \in V' \setminus S \,|\, v \in S \,\&\, \{u,v\} \in E'\}$$

  Similarly, for any subset S of nodes, $\text{cut}_H(S)$ denotes the set of edges of H that have *exactly* one end-point in S.
  The readers should note that our definition of $\partial_H(S)$ involved the set of the nodes, and **not the set of edges**, that are connected to S. Thus, since any subset S containing exactly $|V|/2$ nodes has $|\,\partial_H(S)\,| \leq |V|/2$, we have $0 < \min_{S \subset V \,:\, |S| \leq |V|/2} \{h_H(S)\} \leq 1$ for any graph H.
- $\mathcal{B}_H(u,r)$ is the set of nodes contained in a *ball* of radius r centered at node u in a graph H, *i.e.*, $\mathcal{B}_H(u,r) = \{v \,|\, \text{dist}_H(u,v) \leq r\}$

1.2 Formal Definitions of Gromov-hyperbolicity

Commonly the hyperbolicity measure is defined via geodesic triangles in the following manner.

Definition 1 (δ-hyperbolic graphs via geodesic triangles) A graph G has a (Gromov) hyperbolicity of $\delta = \delta(G)$, or simply is δ-hyperbolic, if and only if for every three ordered triple of shortest paths $(\overline{u,v}, \overline{u,w}, \overline{v,w})$, $\overline{u,v}$ lies in a δ-neighborhood of $\overline{u,w} \cup \overline{v,w}$, i.e., for every node x on $\overline{u,v}$, there exists a node y on $\overline{u,w}$ or $\overline{v,w}$ such that $\text{dist}_G(x,y) \leq \delta$. A δ-hyperbolic graph is simply called a hyperbolic graph if δ is a constant.

Definition 2 (the class of hyperbolic graphs) Let $\mathcal{G}$ be an infinite collection of graphs. Then, $\mathcal{G}$ belongs to the class of hyperbolic graphs if and only if there is an absolute constant $\delta \geq 0$ such that any graph $G \in \mathcal{G}$ is δ-hyperbolic. If $\mathcal{G}$ is a class of hyperbolic graphs then any graph $G \in \mathcal{G}$ is simply referred to as a hyperbolic graph.

There is another alternate but *equivalent* ("up to a constant multiplicative factor") way of defining δ-hyperbolic graphs via the following 4-node conditions.

Definition 3 (equivalent definition of δ-hyperbolic graphs via 4-node conditions) For a set of four nodes u_1, u_2, u_3, u_4, let $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$ be a permutation of $\{1, 2, 3, 4\}$ denoting a rearrangement of the indices of nodes such that

$$\begin{aligned} S_{u_1,u_2,u_3,u_4} = \mathrm{dist}_{u_{\pi_1},u_{\pi_2}} + \mathrm{dist}_{u_{\pi_3},u_{\pi_4}} \\ \leq M_{u_1,u_2,u_3,u_4} = \mathrm{dist}_{u_{\pi_1},u_{\pi_3}} + \mathrm{dist}_{u_{\pi_2},u_{\pi_4}} \\ \leq L_{u_1,u_2,u_3,u_4} = \mathrm{dist}_{u_{\pi_1},u_{\pi_4}} + \mathrm{dist}_{u_{\pi_2},u_{\pi_3}} \end{aligned}$$

and let $\rho_{u_1,u_2,u_3,u_4} = \dfrac{L_{u_1,u_2,u_3,u_4} - M_{u_1,u_2,u_3,u_4}}{2}$. Then, G is δ-hyperbolic if and only if

$$\delta = \delta(G) = \max_{u_1,u_2,u_3,u_4 \in V} \left\{ \rho_{u_1,u_2,u_3,u_4} \right\}.$$

It is well-known (*e.g.*, see [10]) that Definition 1 and Definition 3 of δ-hyperbolicity are equivalent in the sense that they are related by a constant multiplicative factor, *i.e.*, there is an absolute constant $c > 0$ such that if a graph G is δ_1-hyperbolic and δ_2-hyperbolic via Definition 1 and Definition 3, respectively, then $\delta_1/c \leq \delta_2 \leq c\,\delta_1$. Since constant factors are not optimized in our proofs, we will use *either* of the two definitions of hyperbolicity in the sequel as deemed more convenient. Using Definition 3 and casting the resulting computation as a $(\max, \min)$ matrix multiplication problem allows one to compute $\delta(G)$ and a 2-approximation of $\delta(G)$ in $O\left(n^{3.69}\right)$ and in $O\left(n^{2.69}\right)$ time, respectively [17]. Several routing-related problems or the diameter estimation problem become easier if the network is hyperbolic [11–13, 19].

1.2.1 Remarks on Topological Characteristics of Hyperbolicity Measure δ

Even though the hyperbolicity measure $\delta(G)$ is often referred to as a "tree-like" measure, $\delta(G)$ enjoys many non-trivial topological characteristics. For example:

- ⋆ **The "$\delta(G) = o(n)$" property is not hereditary (and thus also not monotone)**. For example, see Fig. 1, which also shows that removing a single node or edge can increase/decrease the value of δ *very sharply*.
- ⋆ **"Close to hyperbolic topology" is not necessarily the same as "close to tree topology"**. For example, *all* bounded-diameter graphs have $\delta = O(1)$ irrespective of whether they are tree or not (however, graphs with $\delta = O(1)$ need *not* be of bounded diameter). In general, even for small δ, the metric induced by a δ-hyperbolic graph may be quite far from a tree metric [11].
- ⋆ **Hyperbolicity is not necessarily the same as tree-width**. A similar popular measure used in both the bioinformatics and theoretical computer science literature is the treewidth measure first introduced by Robertson and Seymour [32]. Many NP-hard problems on general networks in fact allow polynomial-time solutions if restricted to classes of networks with

bounded treewidth [9]. However, as observed in [26] and elsewhere, the two measures are quite different in nature and *not* correlated.

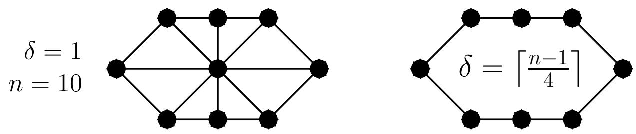

Fig. 1 The "$\delta(G) = o(n)$" property is not hereditary.

Examples of hyperbolic graph classes (*i.e.*, when δ is a constant) include *trees*, *chordal graphs*, *cactus of cliques*, *AT-free* graphs, *link graphs of simple polygons*, and *any* class of graphs with a *fixed* diameter, whereas examples of non-hyperbolic graph classes (*i.e.*, when δ is not a constant) include *expanders*, *simple cycles*, and, for some parameter ranges, the *Erdös-Rényi random graphs*.

Note that if G is δ-hyperbolic then G is also δ'-hyperbolic for any $\delta' > \delta$ (cf. Definition 1). In this paper, to avoid division by zero in terms involving $1/\delta$, we will assume $\delta > 0$. In other words, we will treat a 0-hyperbolic graph (a tree) as a $\frac{1}{2}$-hyperbolic graph in the analysis.

1.3 Relevant Known Results for Gromov Hyperbolicity

We summarize relevant known results that are used in this paper below; many of these results appear in several prior works, *e.g.*, [1, 6, 10, 20, 25]. Fig. 2 pictorially illustrates these results.

Fact 1 (Cylinder removal around a geodesic) [25] *Assume that G is a δ-hyperbolic graph. Let p and q be two nodes of G such that* $\mathrm{dist}_G(p, q) = \beta > 6$*, and let p', q' be nodes on a shortest path between p and q such that* $\mathrm{dist}_G(p, p') = \mathrm{dist}_G(p', q') = \mathrm{dist}_G(q', q) = \beta/3$*. For any $0 < \alpha < 1/4$, let $\mathcal{C}$ be set of nodes at a distance of $\alpha\beta - 1$ of a shortest path $\overline{p', q'}$ between p' and q',* i.e.*, let* $\mathcal{C} = \{u \,|\, \exists\, v \in \overline{p', q'} \colon \mathrm{dist}_G(u, v) = \alpha\beta - 1\}$*. Let $G_{-\mathcal{C}}$ be the graph obtained from G by removing the nodes in $\mathcal{C}$. Then,* $\mathrm{dist}_{G_{-\mathcal{C}}}(p, q) \geq (\beta/60)\, 2^{\alpha\beta/\delta}$.

Fact 2 (Exponential divergence of geodesic rays) [Simplified reformulation of [1, Theorem 10]] *Assume that G is a δ-hyperbolic graph. Suppose that we are given the following:*

- *three integers $\kappa \geq 4$, $\alpha > 0$, $r > 3\kappa\delta$, and*
- *five nodes v, u_1, u_2, u_3, u_4 such that* $\mathrm{dist}_G(v, u_1) = \mathrm{dist}_G(v, u_2) = r$, $\mathrm{dist}_G(u_1, u_2) \geq 3\kappa\delta$, $\mathrm{dist}_G(v, u_3) = \mathrm{dist}_G(v, u_4) = r + \alpha$, *and* $\mathrm{dist}_G(u_1, u_4) = \mathrm{dist}_G(u_2, u_3) = \alpha$.

Consider any path $\mathcal{Q}$ between u_3 and u_4 that does not involve a node in $\bigcup_{0 \leq j \leq r+\alpha} \mathcal{B}_G(v, j)$*. Then, the length $|\mathcal{Q}|$ of the path $\mathcal{Q}$ satisfies* $|\mathcal{Q}| > 2^{\frac{\alpha}{6\delta} + \kappa + 1}$.

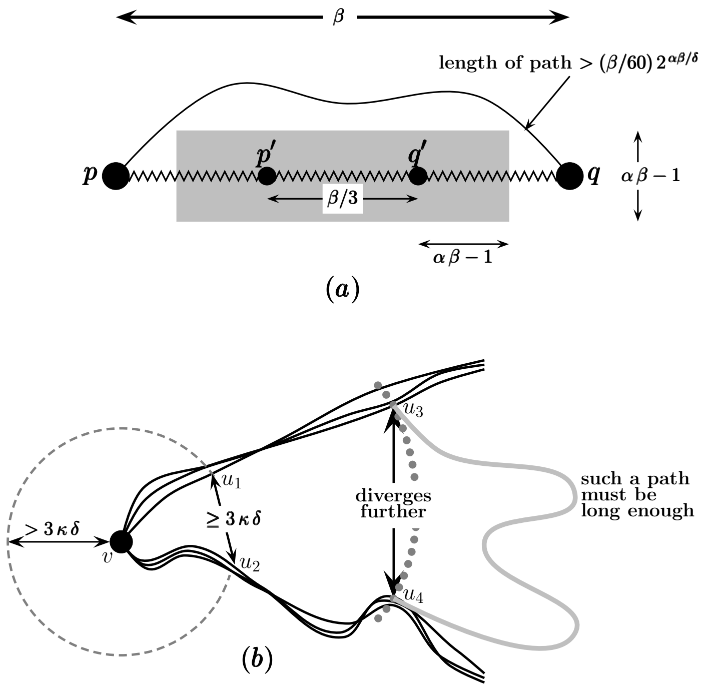

Fig. 2 (a) Illustration of Fact 1. By growing the shaded region and removing nodes in its boundary, one can selectively extract longer paths in the graph. Translating the region slightly does not change this property much. (b) Illustration of Fact 2. Geodesic rays diverging sufficiently cannot connect back without using a sufficiently long path.

2 Overview of Our Results

Before proceeding with formal theorems and proofs, we first provide an informal non-technical intuitive overview of our results.

► Our first two results in Section 3 provide upper bounds for *node expansions* for the triple $\langle G, d, \delta \rangle$ as a function of n, d, and δ. These two results, namely Theorem 1 and Theorem 2, provide absolute bounds and show that many *witnesses* (subset of nodes) satisfying such expansion bounds can be found efficiently in polynomial time satisfying two additional criteria:

- ▷ the witnesses (subsets) form a nested family, or
- ▷ the witnesses have *limited* overlap in the sense that every subset has a certain number of “private” nodes *not* contained in any other subset.

These bounds also imply in an obvious manner corresponding upper bounds for the *edge-expansion* of G and for the smallest non-zero eigenvalue of the Laplacian of G.

In Remark 1, we provide an explanation of the asymptotics of these bounds in comparison to expander-type graphs. For example, if δ is fixed ($i.e.$, G is hyperbolic) then d has to be increased to *at least* $2^{\Omega\left(\sqrt{\log\log n \,/\, \log\log\log n}\right)}$ to get a positive non-zero Cheeger constant, whereas if d is fixed then δ need to be at least $\Omega(\log n)$ to get a positive non-zero Cheeger constant (this last implication also follows from the results in [6, 25]).

- Our last result in Section 3.3, namely Lemma 1, deals with the *absolute* size of s-t cuts in hyperbolic graphs, and shows that a large family of s-t cuts having at most $d^{O(\delta)}$ cut-edges can be found in polynomial time in δ-hyperbolic graphs when d is the maximum degree of any node except s, t and any node within a distance of $35\,\delta$ of s and the distance between s and t is at least $\Omega(\delta \log n)$. This result is later used in designing the approximation algorithm for minimizing bottleneck edges in Section 4.1.
- In Section 4 we discuss some applications of these bounds in designing improved approximation algorithms for two graph-theoretic problems for δ-hyperbolic graphs when δ does not grow too fast as a function of n:
  - We show in Section 4.1 (Lemma 2) that the problem of identifying vulnerable edges in network designs by *minimizing shared edges* admits an improved approximation provided $\delta = o(\log n/\log d)$. We do so by relating it to a hitting set problem for *size-constrained cuts* (Lemma 3) and providing an improved approximation for this latter problem (Lemma 4). We also observe that obvious greedy strategies fail for such problems miserably.
  - In Section 4.2 we provide a polynomial-time solution (Lemma 5) for a type of small-set expansion problem originally proposed by Arora, Barak and Steurer [3] for the case when δ is sub-logarithmic in n.
- Finally, in Section 5 we conclude with some interesting future research questions.

3 Effect of δ on Expansions and Cuts in δ-hyperbolic Graphs

The two results in this section are related to the node (or edge) expansion ratios of a graph that is δ-hyperbolic for some (not necessarily constant) δ. The following definitions are standard in the graph theory literature and repeated here only for the sake of completeness.

Definition 4 (Node and edge expansion ratios of a graph)
(a) The *node expansion ratio* $h_G(S)$ of a subset S of at most $|V|/2$ nodes of a graph $G = (V, E)$ is defined as $h_G(S) = \frac{|\,\partial_G(S)\,|}{|\,S\,|}$. If $h_G(S) > c$ for some constant $c > 0$ and for all subsets S of at most $|V|/2$ nodes then we call G a node-expander.

(b) The edge expansion ratio $g_H(S)$ of a subset S of at most $|V|/2$ nodes of a graph $G = (V, E)$ is defined as $g_G(S) = \frac{|\,\mathrm{cut}_G(S)\,|}{|\,S\,|}$. If $h_G(S) > c$ for some constant $c > 0$ and for all subsets S of at most $|V|/2$ nodes then we call G an edge-expander (or sometimes simply an expander).

Definition 5 (Witness of node or edge expansions) A witness of a node (respectively, edge) expansion bound of c of a graph $G = (V, E)$ is a subset S of at most $|V|/2$ nodes of G such that $h_G(S) \leq c$ (respectively, $g_G(S) \leq c$).

Notation $\mathfrak{h}_G = \min_{S \subset V \,:\, |S| \leq |V|/2} \left\{ h_G(S) \right\}$ will denote the *minimum* node expansion of a graph $G = (V, E)$.

As already observed in Section 1.1, $0 < \mathfrak{h}_G \leq 1$ for any graph G. *All our expansion bounds in this section will be stated for node expansions only.* Since $g_G(S) \leq d\, h_G(S)$ for any graph G whose nodes have a maximum degree of d, our bounds for node expansions translate to some corresponding bounds for the edge expansions as well.

3.1 Nested Family of Witnesses for Node/Edge Expansion

An ordered family of sets $S_1, S_2, \dots, S_\ell$ is called *nested* if $S_1 \subset S_2 \subset \cdots \subset S_\ell$. Our goal in this subsection is to find a large nested family of subsets of nodes with good node expansion bounds.

For two nodes p and q of a graph $G = (V, E)$, a cut S of G that "separates p from q" is a subset S of nodes containing p but not containing q, and the set of cut edges $\mathrm{cut}_G(S, p, q)$ corresponding to the cut S is the set of edges with exactly one end-point in S, *i.e.*,

$$\mathrm{cut}_G(S, p, q) = \left\{ \left\{ u, v \right\} \mid p, u \in S \;\text{ and }\; q, v \in V \setminus S \right\}$$

Recall that d denotes the maximum degree of any node in the given graph G.

Theorem 1 *For any constant $0 < \mu < 1$, the following result holds for $\langle G, d, \delta \rangle$. Let p and q be any two nodes of G and let $\Delta = \mathrm{dist}_G(p, q)$. Then, there exists at least $t = \max\left\{ \frac{\Delta^\mu}{56 \log d}, 1 \right\}$ subsets of nodes $\emptyset \subset S_1 \subset S_2 \subset \cdots \subset S_t \subset V$, each of at most $n/2$ nodes, with the following properties:*

- $\forall\, j \in \{1, 2, \dots, t\}$:
$$h_G(S_j) \leq \min\left\{ 8 \ln(n/2)/\Delta,\ \max\left\{ (1/\Delta)^{1-\mu},\ \frac{500 \ln n}{\Delta\, 2^{\frac{\Delta^\mu}{28\, \delta\, \log(2d)}}} \right\} \right\}.$$
- *All the subsets can be found in a total of $O\left(n^3 \log n + m n^2\right)$ time.*
- *Either all the subsets $S_1, S_2, \dots, S_t$ contain the node p, or all of them contain the node q.*

Corollary 1 *Letting p and q be two nodes such that $\mathrm{dist}_G(p, q) = D(G) = D$ realizes the diameter of the graph G, we get the bound:*

$$h_G(S_j) \leq \min\left\{ \frac{8 \ln(n/2)}{D},\ \max\left\{ \left(\frac{1}{D}\right)^{1-\mu},\ (500 \ln n) / \left(D\, 2^{\frac{D^\mu}{28\, \delta\, \log(2d)}} \right) \right\} \right\}$$

Since $D > \log n / \log d$, the above bound implies:

$$\mathfrak{h}_G < \max\left\{ (\log d / \log n)^{1-\mu},\ (500 \log d) / \left(2^{\log^\mu n / \left(28\, \delta \log^{1+\mu}(2d)\right)} \right) \right\} \qquad (1)$$

Remark 1 The following observations may help the reader to understand the asymptotic nature of the bound in (1).

(a) The first component of the bound is $O\left(1/\log^{1-\mu} n\right)$ for fixed d, and is $\Omega(1)$ only when $d = \Omega(n)$.

(b) To better understand the second component of the bound, consider the following cases (recall that $\mathfrak{h}_G = \Omega(1)$ for an expander):

- ▶ Suppose that the given graph is a hyperbolic graph of constant maximum degree, *i.e.*, both δ and d are constants. In that case,

$$(500 \log d) / \left(2^{\frac{\log^{\mu} n}{28\,\delta \log^{1+\mu}(2d)}} \right) = O\left(1/\left(2^{O(1)\ \log^{\mu} n}\right)\right) = O\left(1/\text{polylog}(n)\right)$$

- ▶ Suppose that the given graph is hyperbolic but the maximum degree d is arbitrary. In that case,

$$(500 \log d) / \left(2^{\frac{\log^{\mu} n}{28\,\delta \log^{1+\mu}(2d)}} \right) = O\left(\log d / \left(2^{O(1) \log^{\mu} n / \log^{1+\mu} d}\right)\right)$$
$$= O\left(\log d / \text{polylog}(n)^{1/\log^{1+\mu} d}\right)$$

  and thus d has to be increased to at least $2^{\Omega\left(\sqrt{\log\log n / \log\log\log n}\right)}$ to get a constant upper bound.

- ▶ Suppose that the given graph has a constant maximum degree but not necessarily hyperbolic (i.e., δ is arbitrary). In that case,

$$(500 \log d) / \left(2^{(\log^{\mu} n)/\left(28\,\delta \log^{1+\mu}(2d)\right)} \right) = O\left(1/2^{O(1) \log^{\mu} n/\delta}\right)$$

  and thus δ need to be at least $\Omega\left(\log^{\mu} n\right)$ to get a constant upper bound.

3.1.1 Proof of Theorem 1

Proof of the main bounds in Theorem 1 uses the same cylinder or ball removing techniques as used in [6, 25] in showing that hyperbolic graphs are not expanders. However, several technical complications arise when we try to find these witnesses while optimizing the corresponding expansion bounds. The time-complexity of finding our witnesses are discussed at the very end of our proof.

(I) **Proof of the easy part of the bound, *i.e.*,** $h_G(S_j) \leq (8 \ln(n/2))/\Delta$

This proof is straightforward and *provided for the sake of completeness.* Assume that $\Delta > (8 \ln(n/2))^{1/\mu}$ since otherwise there is no need to prove this bound. Assume, without loss of generality, that $\left|\mathcal{B}_G(p, \Delta/2)\right| \leq \min\left\{ \left|\mathcal{B}_G(p, \Delta/2)\right|, \left|\mathcal{B}_G(q, \Delta/2)\right| \right\} \leq n/2$. Consider the sequence of balls $\mathcal{B}_G(p, r)$ for $r = 0, 1, 2 \ldots, \Delta/2$. Then it follows that

$$n/2 > |\mathcal{B}_G(p,\Delta/2)| \geq \prod_{\ell=0}^{(\Delta/2)-1}\big(1+h_G(\mathcal{B}_G(p,\ell))\big)$$
$$\geq \prod_{\ell=0}^{(\Delta/2)-1} \mathbf{e}^{h_G(\mathcal{B}_G(p,\ell))/2} = \mathbf{e}^{\sum_{\ell=0}^{(\Delta/2)-1} h_G(\mathcal{B}_G(p,\ell))/2}$$
$$\Rightarrow \ln(n/2) > \sum_{\ell=0}^{(\Delta/2)-1} h_G(\mathcal{B}_G(p,\ell))/2 \Rightarrow \frac{\sum_{\ell=0}^{(\Delta/2)-1} h_G(\mathcal{B}_G(p,\ell))}{\Delta/2} < \frac{4\ln(n/2)}{\Delta}$$

By a simple averaging argument, there must now exist $\Delta/4 > \max\left\{\frac{\Delta^\mu}{56\log d}, 1\right\}$ distinct balls (subsets of nodes) $\mathcal{B}_G(p,r_1) \subset \mathcal{B}_G(p,r_2) \subset \cdots \subset \mathcal{B}_G(p,r_{\Delta/4})$ such that $|\mathcal{B}_G(p,r_j)| < (8\ln(n/2))/\Delta$ for $j=1,2,\ldots,\Delta/4$. It is straightforward to see that these balls can be found within the desired time complexity bound.

(II) **Proof of the difficult part of the bound,**
i.e., $h_G(S_j) \leq \max\left\{(1/\Delta)^{1-\mu}, \frac{500\ln n}{\Delta\, 2^{\frac{\Delta^\mu}{28\,\delta\,\log(2d)}}}\right\}$

(II-a) **The easy case of** $\Delta = O(1)$

If $\Delta = c$ for any some constant $c \geq 1$ (independent of n) then, since $\delta \geq 1/2$, $d > 1$ and n is sufficiently large, we have $(500\ln n)/\left(\Delta\, 2^{\Delta^\mu/(28\,\delta\,\log(2d))}\right) > (500\ln n)/\left(\Delta\, 2^{(1/14)\Delta^\mu}\right) > 1$. Thus, any subset of $n/2$ nodes containing p satisfies the claimed bound, and the number of such subsets is $\binom{\frac{n}{2}-1}{n-2} \gg t$.

(II-b) **The case of** $\Delta = \omega(1)$

Otherwise, assume that $D(n) = \omega(1)$, *i.e.*, $\lim_{n\to\infty} D(n) > c$ for any constant c. Let p', q' be nodes on a shortest path between p and q such that $\mathrm{dist}_G(p,p') = \mathrm{dist}_G(p',q') = \mathrm{dist}_G(q',q) = \Delta/3$. The following initial value of the parameter α is crucial to our analysis[2]:

$$\alpha = \alpha_0 = 1/\left(7\,\Delta^{1-\mu}\,\log(2d)\right) \tag{2}$$

Note that $0 < \alpha_0 < 1/4$. Let $\mathcal{C}$ be set of nodes at a distance of $\lfloor\alpha\Delta\rfloor > \alpha\Delta - 1$ of a shortest path $\overline{p',q'}$ between p' and q'. Thus,

$$\mathcal{C} = \left\{u \mid \exists\, v \in \overline{p',q'} \colon \mathrm{dist}_G(u,v) = \lceil\alpha\Delta\rceil\right\} \Rightarrow |\,\mathcal{C}\,| \leq (\Delta/3)\, d^{\lfloor\alpha\Delta\rfloor} < (\Delta/3)\, d^{\alpha\Delta} \tag{3}$$

Let $G_{-\mathcal{C}}$ be the graph obtained from G by removing the nodes in $\mathcal{C}$. Fact 1 implies:

$$\mathrm{dist}_{G_{-\mathcal{C}}}(p,q) \geq (\Delta/60)\, 2^{\alpha\Delta/\delta} \tag{4}$$

[2] We will later need to vary the value of α in our analysis.

Let $\mathcal{B}_G(p,r)$ be the ball of radius r centered at node p in G with $|\,\mathcal{B}_G(p,r)\,| \leq n/2$, and let $\overline{\boldsymbol{h}}(p,j) \stackrel{\text{def}}{=} \left(\sum_{\ell=0}^{j-1} \mathcal{B}_G(p,\ell)\right)/j$. Then, since $|\,\mathcal{B}_G(p,0)\,| = 1$ and $\frac{|\mathcal{B}_G(p,r)|}{|\mathcal{B}_G(p,r-1)|} = 1 + h_G\left(\mathcal{B}_G(p,r-1)\right)$, we have

$$|\,\mathcal{B}_G(p,r)\,| = \prod_{j=0}^{r-1}\left(1 + h_G\left(\mathcal{B}_G(p,j)\right)\right)$$
$$\geq \prod_{j=0}^{r-1} \mathbf{e}^{h_G(\mathcal{B}_G(p,j))/2} = \mathbf{e}^{\sum_{j=0}^{r-1} h_G(\mathcal{B}_G(p,j))/2} = \mathbf{e}^{r\,\overline{\boldsymbol{h}}(p,r)/2} \quad (5)$$

Assume without loss of generality that[3]

$$\left|\,\mathcal{B}_{G_{-\mathcal{C}}}\left(p, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)\,\right| \leq \left|\,\mathcal{B}_{G_{-\mathcal{C}}}\left(q, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)\,\right| \leq \left(n - |\mathcal{C}|\right)/2 < \frac{n}{2} \quad (6)$$

Case 1: There is a set of t distinct indices $\{i_1, i_2, \ldots, i_t\} \subseteq \{0, 1, \ldots, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\}$ such that, $i_1 < i_2 < \cdots < i_t$ and, for all $1 \leq s \leq t$, $h_G\left(\,\mathcal{B}_G(p,i_s)\,\right) = h_G\left(\,\mathcal{B}_{G_{-\mathcal{C}}}(p,i_s)\,\right) \leq (1/\Delta)^{1-\mu}$ (see Fig. 3 (a)). Then, the subsets $\mathcal{B}_G(p,i_1) \subset \mathcal{B}_G(p,i_2) \subset \cdots \subset \mathcal{B}_G(p,i_t)$ satisfy our claim.

Case 2: Case 1 does not hold. In this case, we have

$$\sum_{\ell=0}^{(\Delta/3)-\alpha\Delta-1} h_G\left(\mathcal{B}_G(p,\ell)\right) > \left(\left(\mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right) - (t-1)\right)\,(1/\Delta)^{1-\mu}$$
$$> \left((\Delta/3) - \alpha\Delta - t\right)(1/\Delta)^{1-\mu} > \Delta^{\mu}/4 \quad (7)$$

Let r_p be the *least* integer such that $\mathcal{B}_{G_{-\mathcal{C}}}(p, r_p) = \mathcal{B}_{G_{-\mathcal{C}}}(p, r_p+1)$. Since G is a connected graph and, for all $r \leq (\Delta/3) - \alpha\Delta$ we have $\mathcal{B}_G(p,r) \cap \mathcal{C} = \emptyset \equiv \mathcal{B}_{G_{-\mathcal{C}}}(p,r) = \mathcal{B}_G(p,r)$ we have $r_p \geq (\Delta/3) - \alpha\Delta$ (see Fig. 3 (a)).

Failure of the current strategy

Note that it is possible that r_p is precisely $(\Delta/3) - \alpha\Delta$ or not too much above it (this could happen when p is disconnected from q in $G_{-\mathcal{C}}$). *Consequently, we may not be able to use our current technique of enlarging the ball $\mathcal{B}_{G_{-\mathcal{C}}}(p,r)$ for r beyond $(\Delta/3) - \alpha\Delta$ to get the required number of subsets of nodes as claimed in the theorem.* A further complication arises because, for $r > (\Delta/3) - \alpha\Delta$, expansion of the balls $\mathcal{B}_{G_{-\mathcal{C}}}(p,r)$ in $G_{-\mathcal{C}}$ may differ from that in G, *i.e.*, $h_G\left(\,\mathcal{B}_{G_{-\mathcal{C}}}(p,r)\,\right)$ *need not be the same as* $h_{G_{-\mathcal{C}}}\left(\,\mathcal{B}_{G_{-\mathcal{C}}}(p,r)\,\right)$.

Rectifying the current strategy

[3] Note that if there is no path between nodes p and q in $G_{-\mathcal{C}}$ then $\mathrm{dist}_{G_{-\mathcal{C}}}(p,q) = \infty$ and $\mathcal{B}_{G_{-\mathcal{C}}}\left(p, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)$ and $\mathcal{B}_{G_{-\mathcal{C}}}\left(q, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)$ contain all the nodes reachable from p and q, respectively, in $G_{-\mathcal{C}}$.

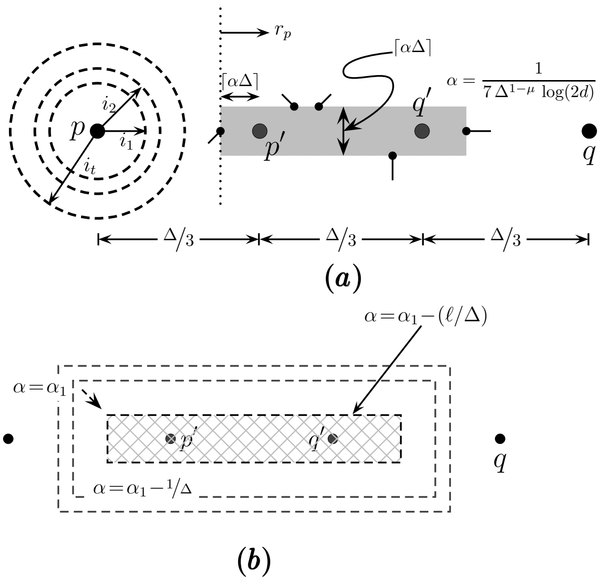

Fig. 3 Illustration of various cases in the proof of Theorem 1. (a) Case 1. Nodes on the boundary of the lightly shaded region belong to $\mathcal{C}_{\alpha_1\Delta}$. ($b$) Case 2. Nodes on the boundary of the lightly cross-hatched region belong to $\mathcal{C}_{\alpha_1\Delta-\ell}$.

We now change our strategy in the following manner. Let us write r_p as $r_{p,\alpha\Delta}$ to show its dependence on $\alpha\Delta$ and let $\alpha_1 = \frac{1}{14\,\Delta^{1-\mu}\,\log(2d)}$. Vary α from $\alpha = \alpha_1$ to $\alpha = \alpha_1/2$ in steps of $-1/\Delta$, and consider the sequence of values $r_{p,\,\alpha_1\Delta}, r_{p,\,\alpha_1\Delta-1}, \dots, r_{p,\,\alpha_1\Delta/2}$. Let $\mathcal{C}_{\alpha_1\Delta-\ell}$ denote the set of nodes in $\mathcal{C}$ when α is set equal to $\alpha_1 - (\ell/\Delta)$ for $\ell = 0, 1, 2, \dots, \alpha_1\Delta/2$ (see Fig. 3 (b)). Consider the two sets of nodes $\mathcal{C}_{\alpha_1\Delta-\ell}$ and $\mathcal{C}_{\alpha_1\Delta-\ell'}$ with $\ell < \ell'$. Obviously, $\mathcal{C}_{\alpha_1\Delta-\ell} \neq \mathcal{C}_{\alpha_1\Delta-\ell'}$ for any $\ell \neq \ell'$.

Case 2.1 (relatively easier case): Removal of each of the set of nodes $\mathcal{C}_{\alpha_1\Delta}, \mathcal{C}_{\alpha_1\Delta-1}, \dots, \mathcal{C}_{(\alpha_1\Delta)/2}$ disconnects p from q in the corresponding graphs $G_{-\mathcal{C}_{\alpha_1\Delta}}, G_{-\mathcal{C}_{\alpha_1\Delta-1}}, \dots, G_{-\mathcal{C}_{(\alpha_1\Delta)/2}}$, respectively.

Then, for any $0 \leq \ell \leq (\alpha_1\Delta)/2$, we have

$$r_{p,\alpha_1\Delta-\ell} \geq (\Delta/3) - \alpha_1\Delta + \ell \geq (\Delta/3) - \alpha_1\Delta$$

$$\left|\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right| > \left|\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta}}}\!\left(p, \frac{\Delta}{3} - \alpha_1\Delta\right)\right| \underset{\text{by (5)}}{\geq} \mathbf{e}^{\frac{1}{2}\sum_{j=0}^{(\Delta/3)-\alpha_1 D-1} h_G(\mathcal{B}_G(p,j))} \underset{\text{by (7)}}{>} \mathbf{e}^{\Delta^{\mu}/8}$$

$$\left|\,\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right)\right| \leq |\,\mathcal{C}_{\alpha_1\Delta-\ell}\,| \leq |\,\mathcal{C}_{\alpha_1\Delta}\,| \underset{\text{by (3)}}{<} (\Delta/3)\, d^{\,\alpha_1\Delta}$$

$$h_G\left(\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right) = \frac{\left|\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right)\right|}{\left|\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right|}$$
$$< \frac{(\Delta/3)\, d^{\alpha_1\Delta}}{\mathbf{e}^{\Delta^\mu/8}} = \frac{(\Delta/3)\, d^{\frac{\Delta^\mu}{14\log(2d)}}}{\mathbf{e}^{\Delta^\mu/8}} < \frac{(\Delta/3)\,\left(d^{1/\log d}\right)^{\Delta^\mu/14}}{\mathbf{e}^{\Delta^\mu/8}}$$
$$= \frac{(\Delta/3)\, 2^{\Delta^\mu/14}}{\mathbf{e}^{\Delta^\mu/8}} < \frac{\Delta/3}{2^{\Delta^\mu/20}} < \left(\frac{1}{\Delta}\right)^{1-\mu}, \text{ since } \mu>0 \text{ and } \Delta=\omega(1) \quad (8)$$

Inequality (8) implies that there is a set of $1+(\alpha_1\Delta)/2 = 1+(\Delta^\mu)/(28\log(2d)) > \Delta^\mu/(56\log d)$ subsets of nodes $\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta}}}(p, r_{p,\alpha_1\Delta}) \subset \mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-1}}}(p, r_{p,\alpha_1\Delta-1}) \subset \cdots \subset \mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta/2}}}(p, r_{p,\alpha_1\Delta/2})$ such that each such subset $\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})$ has $h_G\left(\mathcal{B}_{G_{-\mathcal{C}_{\alpha_1\Delta-\ell}}}(p, r_{p,\alpha_1\Delta-\ell})\right) < (1/\Delta)^{1-\mu}$. This proves our claim.

Case 2.2 (the difficult case): Case 2.1 does not hold.

This means that there exists an index $0 \leq t \leq (\alpha_1\Delta)/2$ such that the removal of the set of nodes in $\mathcal{C}_{\alpha_1\Delta-t}$ does *not* disconnect p from q in the corresponding graphs $G_{-\mathcal{C}_{\alpha_1\Delta-t}}$. This implies $r_{p,\alpha_1\Delta-t} > \text{dist}_{G_{-\mathcal{C}_{\alpha_1\Delta-t}}}(p,q)/2$. For notational convenience, we will denote $\mathcal{C}_{\alpha_1\Delta-t}$ and $G_{-\mathcal{C}_{\alpha_1\Delta-t}}$ simply by $\mathcal{C}$ and $G_{-\mathcal{C}}$, respectively. We redefine $\alpha_0 = \alpha_1 - (t/\Delta)$ such that $\alpha_1\Delta - t = \alpha_0\,\Delta$. Note that $\alpha_1/2 \leq \alpha_0 \leq \alpha_1$.

First goal: show that our selection of α_0 ensures that removal of nodes in $\mathcal{C}$ does not decrease the expansion of the balls $\mathcal{B}_{G_{-\mathcal{C}}}(p,r)$ in the new graph $G_{-\mathcal{C}}$ by more than a constant factor.

First, note that the goal is trivially achieved if If $r \leq (\Delta/3) - \alpha_0\Delta$ since for all $r \leq (\Delta/3) - \alpha_0\Delta$ we have $h_{G_{-\mathcal{C}}}\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r)\right) = h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r)\right)$. Thus, assume that $r > (\Delta/3) - \alpha_0\Delta$. To satisfy our goal, it *suffices* if we can show the following assertion:

$$\forall\, (\Delta/3) - \alpha_0\Delta < r \leq \text{dist}_{G_{-\mathcal{C}}(p,q)}/2 : h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right) > (1/\Delta)^{1-\mu} \Rightarrow$$
$$h_{G_{-\mathcal{C}}}\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right) \geq h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)/2 \quad (9)$$

We verify (9) as shown below. First, note that:

$$h_{G_{-\mathcal{C}}}\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right) \geq h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)/2$$
$$\equiv \frac{\left|\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right| - \left|\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\cap\mathcal{C}\right|}{\left|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right|} \geq \frac{h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)}{2}$$
$$\Leftarrow \frac{\left|\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right| - |\mathcal{C}|}{\left|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right|} \geq h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)/2$$
$$\equiv \frac{\left|\partial_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right|}{\left|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right|} - \frac{|\mathcal{C}|}{\left|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right|} \geq h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)/2$$
$$\equiv h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right) - \frac{|\mathcal{C}|}{\left|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right|} \geq h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)/2$$

$$\equiv \frac{2\,|\mathcal{C}|}{h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)} \le \left|\,\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\,\right|$$

$$\Leftarrow \frac{2\,|\mathcal{C}|}{h_G\big(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\big)} \le \mathbf{e}^{\frac{\Delta^\mu}{8}}, \text{ since } \big|\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\big| \ge \left|\mathcal{B}_{G_{-\mathcal{C}}}\left(p, \frac{\Delta}{3}-\alpha_0\Delta\right)\right|$$

$$= \left|\,\mathcal{B}_G\left(p, (\Delta/3)-\alpha_0\Delta\right)\,\right|$$

$$\ge \left|\,\mathcal{B}_G\left(p, (\Delta/3)-\alpha_1\Delta\right)\,\right| > \mathbf{e}^{\Delta^\mu/8}$$

$$\Leftarrow \left((\Delta/3)\,d^{\,\alpha_0\Delta}\right)\left(2/\left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right)\right) \le \mathbf{e}^{\Delta^\mu/8}, \text{ since } |\mathcal{C}| < (\Delta/3)\,d^{\,\alpha_0\Delta}$$

$$\equiv (\Delta^\mu/8) \ge \ln\Delta + \alpha_0\,\Delta\ln d - \ln(3/2) - \ln\left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right)$$

$$\Leftarrow \frac{\Delta^\mu}{8} \ge \ln\Delta + \alpha_1\,\Delta\ln d - \ln\tfrac{3}{2} - \ln\left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right), \text{ since } \alpha_0 \le \alpha_1$$

$$\Leftarrow \alpha_1 \le \frac{(\Delta^\mu/8) - \ln\Delta + \ln\left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right)}{\Delta\ln d} \tag{10}$$

Now, if $h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right) > (1/\Delta)^{1-\mu}$ then since $\Delta = \omega(1)$ we have:

$$\frac{\Delta^\mu}{8} - \ln\Delta + \ln\left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,r-1)\right)\right) > \frac{\Delta^\mu}{8} - \ln\Delta - (1-\mu)\ln\Delta > (\Delta^\mu/7)$$

Thus, Inequality (10) is satisfied by our selection of $\alpha_1 = 1/\left(14\,\Delta^{1-\mu}\,\log(2d)\right)$. This verifies (9) and satisfies our first goal.

Second goal: Use the first goal and the fact that $\mathrm{dist}_{G_{-\mathcal{C}}(p,q)}$ is large enough to find the desired subsets.

First assume that there exists a set of $t = \max\{1,\, D^\mu/(56\log d)\}$ indices $i_1 < i_2 < \cdots < i_t$ in $\left\{\frac{\Delta}{3} - \alpha_0\Delta + 1, \frac{\Delta}{3} - \alpha_0\Delta + 2, \ldots, (\mathrm{dist}_{G_{-\mathcal{C}}}(p,q))/2\right\}$ such that

$$\forall\, 1 \le s \le t:\; h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,i_s)\right) \le (1/\Delta)^{1-\mu} \tag{11}$$

Obviously, the existence of these subsets $\mathcal{B}_{G_{-\mathcal{C}}}(p,i_1) \subset \mathcal{B}_{G_{-\mathcal{C}}}(p,i_2) \subset \cdots \subset \mathcal{B}_{G_{-\mathcal{C}}}(p,i_t)$ proves our claim. Otherwise, there are no sets of t indices that satisfy (11). This implies that there exists a set of $\xi = \left(\mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right) - ((\Delta/3) - \alpha_0\Delta) - (t-1)$ distinct indices $j_1, j_2, \ldots, j_\xi$ in $\left\{(\Delta/3) - \alpha_0\Delta + 1, (\Delta/3) - \alpha_0\Delta + 2, \ldots, (\mathrm{dist}_{G_{-\mathcal{C}}}(p,q))/2\right\}$ such that

$$\forall\, 1 \le s \le \xi:\; h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,j_s)\right) > (1/\Delta)^{1-\mu} \underset{\text{by (9)}}{\Rightarrow}$$

$$\forall\, 1 \le s \le \xi:\; h_{G_{-\mathcal{C}}}\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,j_s)\right) \ge h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,j_s)\right)/2 \tag{12}$$

This in turn implies

$$\left|\,\mathcal{B}_{G_{-\mathcal{C}}}\left(p, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)\,\right|$$

$$> \left(\prod_{j=0}^{(\Delta/3)-\alpha_0\Delta-1}\left(1 + h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,j)\right)\right)\right)\left(\prod_{j=(\Delta/3)-\alpha_0\Delta}^{(\Delta/3)-\alpha_0\Delta+\xi-1}\left(1 + \left(h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p,j)\right)/2\right)\right)\right)$$

using (12)

$$> \left(\prod_{j=0}^{(\varDelta/3)-\alpha_0\varDelta-1} \mathbf{e}^{h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))\,/\,2} \right) \left(\prod_{j=(\varDelta/3)-\alpha_0\varDelta}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} \mathbf{e}^{h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))\,/\,4} \right)$$

$$= \left(\mathbf{e}^{\sum_{j=0}^{(\varDelta/3)-\alpha_0\varDelta-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))/2} \right) \left(\mathbf{e}^{\sum_{j=(\varDelta/3)-\alpha_0\varDelta}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))/4} \right) > \mathbf{e}^{\sum_{j=0}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))/4} \quad (13)$$

Using (13) and our specific choice of the node p (over node q), we have

$$n/2 > \left| \mathcal{B}_{G_{-\mathcal{C}}}\left(p, \mathrm{dist}_{G_{-\mathcal{C}}}(p,q)/2\right) \right| > \mathbf{e}^{\sum_{j=0}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))/4} \Rightarrow$$

$$\sum_{j=0}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j)) < 4\ln n \quad (14)$$

We now claim that there *must* exist a set of $t = \varDelta^{\mu}/(56\log d)$ distinct indices $i_1 < i_2 < \cdots < i_t$ in $\{0, 1, \ldots, (\varDelta/3) - \alpha_0\varDelta + \xi - 1\}$ such that

$$\forall\, 1 \le s \le t: \; h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p, i_s)\right) \le (500\ln n)/\left(\varDelta\, 2^{\,\varDelta^{\mu}/(28\delta\log(2d))}\right) \quad (15)$$

The existence of these indices will obviously prove our claim. Suppose, for the sake of contradiction, that this is *not* the case. Together with (14) this implies:

$$4\ln n > \sum_{j=0}^{(\varDelta/3)-\alpha_0\varDelta+\xi-1} h_G(\mathcal{B}_{G_{-\mathcal{C}}}(p,j))$$

$$\underset{\text{by (15)}}{>} \left(\frac{\varDelta}{3} - \alpha_0\varDelta + \xi - \frac{\varDelta^{\mu}}{56\log d} + 1 \right) \left((500\ln n)/\left(\varDelta\, 2^{\,\varDelta^{\mu}/(28\delta\log(2d))}\right) \right)$$

$$\Rightarrow \left(\tfrac{\mathrm{dist}_{G_{-\mathcal{C}}}(p,q)}{2} - \max\left\{1, \tfrac{\varDelta^{\mu}}{28\log d}\right\} \right) \left((500\ln n)/\left(\varDelta\, 2^{\frac{\varDelta^{\mu}}{28\delta\log(2d)}}\right) \right) < 4\ln n,$$

substituting the values of t and ξ

$$\Rightarrow \left(\frac{\varDelta}{120}\, 2^{\frac{\alpha_1\varDelta}{2\delta}} - \max\left\{1, \frac{\varDelta^{\mu}}{28\log d}\right\} \right) \left((500\ln n)/\left(\varDelta\, 2^{\frac{\varDelta^{\mu}}{28\delta\log(2d)}}\right) \right) < 4\ln n,$$

by (4) and since $\alpha_1/2 \le \alpha_0$

$$\equiv \left(\frac{\varDelta}{120}\, 2^{\frac{\varDelta^{\mu}}{28\delta\log(2d)}} - \max\left\{1, \frac{\varDelta^{\mu}}{28\log d}\right\} \right) \left(125/\left(\varDelta\, 2^{\,\varDelta^{\mu}/(28\delta\log(2d))}\right) \right) < 1$$

$$\Rightarrow \left((\varDelta/121)\, 2^{\,\varDelta^{\mu}/(28\delta\log(2d))} \right) \left(125/\left(\varDelta\, 2^{\,\varDelta^{\mu}/(28\delta\log(2d))}\right) \right) < 1$$

$$\equiv {}^{125}/_{121} < 1, \quad \text{since } \varDelta = \omega(1) \quad (16)$$

Since (16) is false, there must exist a set of t distinct indices $i_1 < i_2 < \cdots < i_t$ such that (15) holds and the corresponding sets $\mathcal{B}_{G_{-\mathcal{C}}}(p, i_1) \subset \mathcal{B}_{G_{-\mathcal{C}}}(p, i_2) \subset \cdots \subset \mathcal{B}_{G_{-\mathcal{C}}}(p, i_t)$ prove our claim.

(III) **Time complexity for finding each witness**

It should be clear that we can find each witness provided we can implement the following steps:

– Find two nodes p and q such that $\text{dist}_G(p, q) = \Delta$ in $O\left(n^2 \log n + mn\right)$ time.
– Using breadth-first-search (BFS), find the two nodes p', q' as in the proof in $O(m + n)$ time.
– There are at most $\alpha_1 \Delta/2 = \Delta^\mu/(28 \log(2d)) < n$ possible values of α considered in the proof. For each α, the following steps are needed:
  – Use BFS find the set of nodes $\mathcal{C}$ in $O\left(n^2 + mn\right)$ time.
  – Compute $G_{-\mathcal{C}}$ in $O(m + n)$ time.
  – Use BFS to compute $\mathcal{B}_{G_{-\mathcal{C}}}(p, r)$ for every $0 \leq r \leq \text{dist}_{G_{-\mathcal{C}}}(p, q)/2$ in $O(m + n)$ time.
  – Compute $h_G\left(\mathcal{B}_{G_{-\mathcal{C}}}(p, r)\right)$ for every $0 \leq r \leq \text{dist}_{G_{-\mathcal{C}}}(p, q)/2$ in $O(n^2 + mn)$ time, and select a subset of nodes with a minimum expansion.

3.2 Family of Witnesses of Node/Edge Expansion With Limited Mutual Overlaps

The result in the previous section provided a nested family of cuts of small expansion that separated node p from node q. However, pairs of subsets in this family may differ by as few as just *one* node. In some applications, one may need to generate a family of cuts that are sufficiently different from each other, *i.e.*, they are either disjoint or have limited overlap. The following theorem addresses this question.

Theorem 2 *Let p and q be any two nodes of G and let $\Delta = \text{dist}_G(p, q) > 8$. Then, for any constant $0 < \mu < 1$ and for any positive integer $\tau < \Delta/\left(\left(42\,\delta \log(2d) \log(2\Delta)\right)^{1/\mu}\right)$ the following results hold for $\langle G, d, \delta \rangle$: there exists $\lfloor \tau/4 \rfloor$ distinct collections of subsets of nodes $\emptyset \subset \mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_{\lfloor \tau/4 \rfloor} \subset 2^V$ such that*

- ► $\forall\, j \in \{1, \ldots, \lfloor \tau/4 \rfloor\}\ \forall\, S \in \mathcal{F}_j$:
$$h_G(S) \leq \max \left\{ \left(\frac{1}{(\Delta/\tau)}\right)^{1-\mu}, \frac{360 \log n}{(\Delta/\tau)\, 2^{\frac{(\Delta/\tau)^\mu}{7\,\delta\, \log(2d)}}} \right\}.$$
- ► *Each collection $\mathcal{F}_j$ has at least $t_j = \max\left\{\frac{(\Delta/\tau)^\mu}{56 \log d}, 1\right\}$ subsets $V_{j,1}, \ldots, V_{j,t_j}$ that form a nested family, i.e., $\emptyset \subset V_{j,1} \subset V_{j,2} \subset \cdots \subset V_{j,t_j} \subset V$.*
- ► *All the subsets in each $\mathcal{F}_j$ can be found in a total of $O\left(n^3 \log n + mn^2\right)$ time.*
- ► (limited overlap claim) *For every pair of subsets $V_{i,k} \in \mathcal{F}_i$ and $V_{j,k'} \in \mathcal{F}_j$ with $i \neq j$, either $V_{i,k} \cap V_{j,k'} = \emptyset$ or at least $\Delta/(2\,\tau)$ nodes in each subset do not belong to the other subset.*

Remark 2 Consider a bounded-degree hyperbolic graph, *i.e.*, assume that δ and d are constants. Setting $\tau = \Delta^{1/2}$ gives $\Omega(\Delta^{1/2})$ nested families of subsets of nodes, with each family having at least $\Omega(\Delta^{1/2})$ subsets each of maximum node expansion $(1/\Delta)^{(1-\mu)/2}$, such that every pairwise non-disjoint subsets from different families have at least $\Omega(\Delta^{1/2})$ private nodes.

Proof. Select $\tau \leq \Delta/4$ such that τ satisfies the following:

$$\Delta/(60\,\tau)\,2^{\,(\,(\Delta/\tau)^{\mu}\,)/(28\,\delta\,\log(2d))} > (\Delta/\tau) + 2\,\Delta \tag{17}$$

Note that $\tau \geq (42\,\delta\,\log(2d)\,\log(2\Delta)\,)^{1/\mu}/\Delta$ satisfies (17) since

$$\begin{aligned}
&\Delta/(60\,\tau)\,2^{\,(\,(\Delta/\tau)^{\mu}\,)/(28\,\delta\,\log(2d))} > (\Delta/\tau) + 2\,\Delta \\
&\quad\Rightarrow\ (\Delta/\tau)^{\mu} > 28\,\delta\,\log(2d)\,\log(60+120\,\tau) \underset{\text{since } \tau<\Delta/4}{>} 168\,\delta\,\log(2d)\,\log(2\Delta) \\
&\qquad\qquad\qquad\qquad\Rightarrow\ \tau < \Delta/\left(\left(42\,\delta\,\log(2d)\,\log(2\Delta)\,\right)^{1/\mu}\right)
\end{aligned}$$

Let $(p=p_1,p_2,\ldots,p_{\tau+1}=q)$ be an ordered sequence of $\tau+1$ nodes such that $\mathrm{dist}_G\,(p_i,p_{i+1}) = \Delta/\tau$ for $i=1,2,\ldots,\tau$. Applying Theorem 1 for each pair (p_i,p_{i+1}), we get a nested family $\emptyset \subset \mathcal{F}_i \subset 2^V$ of subsets of nodes such that $t_i = |\,\mathcal{F}_i\,| \geq \max\left\{\frac{(\Delta/\tau)^{\mu}}{56\log d},\,1\right\}$ and, for any $V_{i,k} \in \mathcal{F}_i$, $h_G\,(V_{i,k}) \leq \max\left\{(1/(\Delta/\tau))^{1-\mu},\ (360\log n)/\left((\Delta/\tau)\,2^{\,(\Delta/\tau)^{\mu}/(7\,\delta\,\log(2d))}\right)\right\}$. Recall that the subset of nodes $V_{i,k}$ was constructed in Theorem 1 in the following manner (see Fig. 4 for an illustration):

- Let ℓ_i and r_i be two nodes on a shortest path $\overline{p_i,p_{i+1}}$ such that $\mathrm{dist}_G\,(p_i,\ell_i) = \mathrm{dist}_G\,(\ell_i,r_i) = \mathrm{dist}_G\,(r_i,p_{i+1}) = \mathrm{dist}_G\,(p_i,p_{i+1})\,/3$.
- For some $1/\left(28\,(\Delta/\tau)^{1-\mu}\log(2d)\right) \leq \alpha_{i,k} \leq 1/\left(14\,(\Delta/\tau)^{1-\mu}\log(2d)\right) < 1/4$, construct the graph $G_{-\mathcal{C}_{i,k}}$ obtained by removing the set of nodes $\mathcal{C}_{i,k}$ which are *exactly* at a distance of $\lceil \alpha_{i,k}\,\mathrm{dist}_G\,(p_i,p_{i+1})\rceil$ from some node of the shortest path $\overline{\ell_i,r_i}$.
- The subset $V_{i,k}$ is then the ball $\mathcal{B}_{G_{-\mathcal{C}_{i,k}}}(y_i,a_{i,k})$ for some $a_{i,k} \in \left[0,\,\mathrm{dist}_{G_{-\mathcal{C}_{i,k}}}(p_i,p_{i+1})/2\right]$ and for some $y_i \in \{p_i,p_{i+1}\}$. If $y_i = p_i$ then we call the collection of subsets $\mathcal{F}_i$ "left handed", otherwise we call $\mathcal{F}_i$ "right handed".

We can partition the set of τ collections $\mathcal{F}_1,\ldots,\mathcal{F}_\tau$ into four groups depending on whether the subscript j of $\mathcal{F}_j$ is odd or even, and whether $\mathcal{F}_j$ is left handed or right handed. One of these 4 groups must at least $\lfloor \tau/4 \rfloor$ family of subsets. Suppose, without loss of generality, that this happens for the collection of families that contains $\mathcal{F}_{i,k}$ when i is even and $\mathcal{F}_{i,k}$ is left handed (the other cases are similar). We now show that subsets in this collection that belong to different families do satisfy the *limited overlap* claim.

Consider an arbitrary set in the above-mentioned collection of the form $V_{i,k} = \mathcal{B}_{G_{-\mathcal{C}_{i,k}}}(p_i,a_{i,k})$ with even i. Let $\mathfrak{C}_{i,k}$ denote the nodes in the *interior* of the closed cylinder of nodes in G which are at a distance of *at most* $\lceil \alpha_{i,k}\,\mathrm{dist}_G\,(p_i,p_{i+1})\rceil$ from some node of the shortest path $\overline{\ell_i,r_i}$, *i.e.*, let $\mathfrak{C}_{i,k} = \left\{u \,|\, \exists\, v \in \overline{\ell_i,r_i} \colon \mathrm{dist}_G(u,v) \leq \lceil \alpha_{i,k}\,\mathrm{dist}_G\,(p_i,p_{i+1})\rceil\,\right\}$ (see Fig. 4). Let $V_{j,k'} = \mathcal{B}_{G_{-\mathcal{C}_{j,k'}}}(p_j,a_{j,k'})$ be a set in another family $\mathcal{F}_j$ with even $j \neq i$ (see Fig. 4). Assume, without loss of generality, that i is smaller than j, *i.e.*, $i \leq j-2$ (the other case is similar).

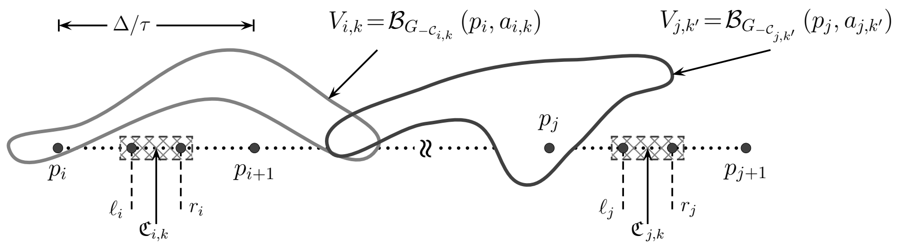

Fig. 4 Illustration of various quantities related to the proof of Theorem 2. Nodes within the lightly cross-hatched region belong to $\mathfrak{C}_{i,k}$ and $\mathfrak{C}_{j,k'}$. Note that $\mathcal{B}_{G-\mathcal{C}_{i,k}}(p_i, a_{i,k})$ and $\mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, a_{j,k'})$ need not be balls in the original graph G.

Proposition 1 $\mathfrak{C}_{i,k} \cap \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, \Delta/(2\tau)) = \emptyset$.

Proof. Assume for the sake of contradiction that $u \in \mathfrak{C}_{i,k} \cap \mathcal{B}_{G-\mathcal{C}_{j,k'}}\left(p_j, \frac{\Delta}{2\tau}\right)$. Since $u \in \mathfrak{C}_{i,k}$, there is a node $v \in \overline{\ell_i, r_i}$ such that

$$\text{dist}_G(v,u) \leq \lceil \alpha_{i,k} \, \text{dist}_G(p_i, p_{i+1}) \rceil < \text{dist}_G(p_i, p_{i+1})/4 = \Delta/(4\tau)$$

Thus,

$$u \in \mathcal{B}_{G-\mathcal{C}_{j,k'}}\left(p_j, \tfrac{\Delta}{2\tau}\right) \Rightarrow \text{dist}_{G-\mathcal{C}_{j,k'}}(u, p_j) \leq \tfrac{\Delta}{2\tau} \Rightarrow \text{dist}_G(u, p_j) \leq \tfrac{\Delta}{2\tau}$$
$$\Rightarrow \text{dist}_G(v, p_j) \leq \text{dist}_G(v,u) + \text{dist}_G(u, p_j) < \Delta/(4\tau) + \Delta/(2\tau) < \Delta/\tau$$

which contradicts the fact that $\text{dist}_G(v, p_j) > \text{dist}_G(p_{i+1}, p_j) = \Delta/\tau$. ❑

Proposition 2 $\text{dist}_{G-\mathcal{C}_{j,k'}}(u, p_j) > \Delta/(2\tau)$ *for any node* $u \in V_{i,k} \cap V_{j,k'} = \mathcal{B}_{G-\mathcal{C}_{i,k}}(p_i, a_{i,k}) \cap \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, a_{j,k'})$.

Proof. Assume for the sake of contradiction that $z = \text{dist}_{G-\mathcal{C}_{j,k'}}(u, p_j) \leq \Delta/(2\tau)$. Since $u \in V_{i,k} = \mathcal{B}_{G-\mathcal{C}_{i,k}}(p_i, a_{i,k})$, this implies

$$\text{dist}_{G-\mathcal{C}_{i,k}}(p_{i,k}, u) \leq a_{i,k} \leq \text{dist}_{G-\mathcal{C}_{i,k}}(p_i, p_{i+1})/2$$

Since $u \in V_{j,k'} = \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, a_{j,k'})$ for some $a_{j,k'}$, this implies $u \in \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, z)$. Since $z \leq \Delta/(2\tau)$, by Proposition 1 $\mathfrak{C}_{i,k} \cap \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, z) = \emptyset$, and therefore

$$\Delta/(2\tau) \geq z = \text{dist}_{G-\mathcal{C}_{j,k'}}(u, p_j) \underset{\text{since } \mathfrak{C}_{i,k} \cap \mathcal{B}_{G-\mathcal{C}_{j,k'}}(p_j, z) = \emptyset}{=} \text{dist}_{G-\mathcal{C}_{i,k} \cup \mathcal{C}_{j,k'}}(u, p_j) \geq \text{dist}_{G-\mathcal{C}_{i,k}}(u, p_j)$$

which in turn implies

$$\text{dist}_{G-\mathcal{C}_{i,k}}(p_i, p_j) \leq \text{dist}_{G-\mathcal{C}_{i,k}}(p_i, u) + \text{dist}_{G-\mathcal{C}_{i,k}}(u, p_j)$$

$$\leq \mathrm{dist}_{G_{-\mathcal{C}_{i,k}}}(p_i, p_{i+1})/2 + \Delta/(2\,\tau) \quad (18)$$

Since the Hausdorff distance between the two shortest paths $\overline{\ell_i, r_i}$ and $\overline{p_j, p_{j+1}}$ is at least $(j-i-1)\frac{\Delta}{\tau}+\Delta/(3\,\tau) > \alpha_{i,k}\,\mathrm{dist}_G\,(p_i, p_{i+1})$ and $\mathrm{dist}_{\,G_{-\mathcal{C}_{i,k}}}(p_j, p_{i+1}) = (j-i)\Delta/\tau < \Delta$, we have

$$\begin{aligned}\mathrm{dist}_{\,G_{-\mathcal{C}_{i,k}}}(p_i, p_{i+1}) &\leq \mathrm{dist}_{\,G_{-\mathcal{C}_{i,k}}}(p_i, p_j) + \mathrm{dist}_{\,G_{-\mathcal{C}_{i,k}}}(p_j, p_{i+1}) \\ &\underset{\text{by (18)}}{\leq} \mathrm{dist}_{G_{-\mathcal{C}_{i,k}}}(p_i, p_{i+1})/2 + \Delta/(2\,\tau) + \Delta \\ &\Rightarrow \mathrm{dist}_{\,G_{-\mathcal{C}_{i,k}}}(p_i, p_{i+1}) \leq \Delta/\tau + 2\Delta \quad (19)\end{aligned}$$

On the other hand, by Fact 1:

$$\mathrm{dist}_{\,G_{-\mathcal{C}_i}}(p_i, p_{i+1}) \geq \Delta/(60\,\tau)\,2^{\,(\alpha_{i,k}\,\Delta)/(\delta\,\tau)} \geq \Delta/(60\,\tau)\,2^{\,(\,(\Delta/\tau)^{\mu}\,)/(28\,\delta\,\log(2d))} \quad (20)$$

Inequalities (19) and (20) together imply

$$\Delta/(60\,\tau)\,2^{\,(\,(\Delta/\tau)^{\mu}\,)/(28\,\delta\,\log(2d))} \leq (\Delta/\tau) + 2\,\Delta \quad (21)$$

Inequality (21) contradicts Inequality (17). ❑

To complete the proof of limited overlap claim, suppose that $V_{i,k} \cap V_{j,k'} \neq \emptyset$ and let $u \in V_{i,k} \cap V_{j,k'}$. Proposition 2 implies that $V_{j,k'} \supset \mathcal{B}_{G_{-\mathcal{C}_{j,k'}}}(p_j, \Delta/(2\,\tau))$, $u \notin \mathcal{B}_{G_{-\mathcal{C}_{j,k'}}}(p_j, \Delta/(2\,\tau))$, and thus there are at least $\Delta/(2\,\tau)$ node on a shortest path in $G_{-\mathcal{C}_{j,k'}}$ from p_j to a node at a distance of $\Delta/(2\,\tau)$ from p_j that are not in $V_{i,k}$. ❑

3.3 Family of Mutually Disjoint Cuts

Recall that, given two distinct nodes $s, t \in V$ of a graph $G = (V, E)$, a cut in G that separates s from t (or, simply a "s-t cut") $\mathrm{cut}_G(S, s, t)$ is a subset of nodes S that disconnects s from t. The *cut-edges* $\mathcal{E}_G(S, s, t)$ (resp., *cut-nodes* $\mathcal{V}_G(S, s, t)$) corresponding to this cut is the set of edges with one end-point in S (resp., the end-points of these cut-edges that belong to S), *i.e.*,

$$\mathcal{E}_G(S, s, t) = \{\,\{u, v\} \mid u \in S,\ v \in V \setminus S,\ \{u, v\} \in E\,\},$$
$$\mathcal{V}_G(S, s, t) = \{\,u \mid u \in S,\ v \in V \setminus S,\ \{u, v\} \in E\,\}$$

Note that in the following lemma d is the maximum degree of any node "except s, t and any node within a distance of $35\,\delta$ of s" (degrees of these nodes may be arbitrary).

Lemma 1 *Suppose that the following holds for our given $\langle G, d, \delta\rangle$:*

► *s and t are two nodes of G such that $\mathrm{dist}_G(s, t) > 48\,\delta + 8\,\delta \log n$, and*

► *d is the maximum degree of any node except s, t and any node within a distance of $35\,\delta$ of s (degrees of these nodes may be arbitrary).*

Then, there exists a set of at least $\frac{\mathrm{dist}_G(s,t)-8\,\delta\log n}{50\,\delta} = \Omega\left(\mathrm{dist}_G(s,t)\right)$ *(node and edge) disjoint cuts such that each such cut has at most* $d^{\,12\delta+1}$ *cut edges.*

Remark 3 Suppose that G is hyperbolic (i.e., δ is a constant), d is a constant, and s and t be two nodes such that $\mathrm{dist}_G(s,t) > 48\delta + 8\delta\log n = \Omega(\log n)$. Lemma 1 then implies that there are $\Omega\left(\mathrm{dist}_G(s,t)\right)$ s-t cuts each having $O(1)$ edges. If, on the other hand, $\delta = O(\log\log n)$, then such cuts have polylog(n) edges.

Remark 4 The bound in Lemma 1 is obviously meaningful only if $\delta = o(\log_d n)$. If $\delta = \Omega(\log_d n)$, then δ-hyperbolic graphs include expanders and thus many small-size cuts may not exist in general.

Proof. Recall that we may assume that $\delta \geq 1/2$. We start by doing a BFS starting from node s. Let $\mathcal{L}_i$ be the sets of nodes at the i^{th} level (*i.e.*, $\forall\, u \in \mathcal{L}_i\colon \mathrm{dist}_G\,(s,u) = i$); obviously $t \in \mathcal{L}_{\mathrm{dist}_G(s,t)}$. Assume $\mathrm{dist}_G(s,t) > 48\,\delta + 8\,\delta\,\log n$, and consider two arbitrary paths $\mathcal{P}_1$ and $\mathcal{P}_2$ between s and t passing through two nodes $v_1, v_2 \in \mathcal{L}_j$ for some $48\,\delta \leq j \leq \mathrm{dist}_G(s,t) - 7\,\delta\,\log n$.

We first claim that $\mathrm{dist}_G\,(v_1,v_2) < 12\,\delta$. Suppose, for the sake of contradiction, suppose that $\mathrm{dist}_G\,(v_1,v_2) \geq 12\delta$. Let v_1' and v_2' be the first node in level $\mathcal{L}_{j+6\delta\log n}$ visited by $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Since both $\mathcal{P}_1$ and $\mathcal{P}_2$ are paths between s and t and $j + 6\,\delta\,\log n < \mathrm{dist}_G(s,t)$ implies $\mathcal{L}_{j+6\delta\log n+1} \neq \emptyset$, there must be a path $\mathcal{P}_3$ between v_1' and v_2' through t using nodes not in $\bigcup_{0\,\leq\,\ell\,\leq\,j+6\,\delta\log n} \mathcal{L}_\ell$. We show that this is impossible by Fact 2. Set the parameters in Fact 2 in the following manner: $\kappa = 4$, $\alpha = 6\,\delta\,\log n$, $r = j > 12\,\kappa\,\delta = 48\,\delta$; $u_1 = v_1$, $u_2 = v_2$, $u_4 = v_1'$, and $u_3 = v_2'$. Then the length of $\mathcal{P}_3$ satisfies $|\,\mathcal{P}_3\,| > 2^{\log n+5} > n$ which is impossible since $|\,\mathcal{P}_3\,| < n$.

We next claim that, for any arbitrary node in level $v \in \mathcal{L}_j$ lying on a path between s and t, $\mathcal{B}_G\,(v,12\delta)$ provides an s-t cut $\mathrm{cut}_G\,(\mathcal{B}_G\,(v,12\delta)\,,s,t)$ having at most $\mathcal{E}_G\,(\mathcal{B}_G\,(v,12\delta)\,,s,t) \leq d^{\,12\delta+1}$ edges. To see this, consider any path $\mathcal{P}$ between s and t and let u be the first node in $\mathcal{L}_j$ visited by the path. Then, $\mathrm{dist}_G(u,v) \leq 12\delta$ and thus $v \in \mathcal{B}_G\,(v,12\delta)$. Since nodes in $\mathcal{B}_G\,(v,12\delta)$ are at a distance of at least 35δ from s and $t \notin \mathcal{B}_G\,(v,12\delta)$, d is the maximum degree of any node in $\mathcal{B}_G\,(v,12\delta)$ and it follows that $\mathcal{E}_G\,(\mathcal{B}_G\,(v,12\delta)\,,s,t) \leq d\,\partial_G\,(\mathcal{B}_G\,(v,12\delta-1)) \leq d^{\,12\delta+1}$.

We can now finish the proof of our lemma in the following way. Assume that $\mathrm{dist}_G(s,t) > 48\,\delta + 8\,\delta\,\log n$. Consider the levels $\mathcal{L}_j$ for $j \in \left\{\, 50\delta, 100\delta, 150\delta, \ldots, \frac{\mathrm{dist}_G(s,t)-8\delta\log n}{50\delta}\,\right\}$. For each such level $\mathcal{L}_j$, select a node v_j that is on a path between s and t and consider the subset of edges in $\mathrm{cut}_G\big(\mathcal{B}_G\big(v_j,12\delta\big),s,t\big)$. Then, $\mathrm{cut}_G\,(\mathcal{B}_G\,(v_j,12\delta)\,,s,t)$ over all j provides our family of s-t cuts. The number of such cuts is at least $(\mathrm{dist}_G(s,t) - 8\delta\log n)/(50\delta)$. To see why these cuts are node and edge disjoint, note that $\mathcal{E}_G(\mathcal{B}_G(v_j,12\delta),s,t) \cap \mathcal{E}_G(\mathcal{B}_G(v_\ell,12\delta),s,t) = \emptyset$ and $\mathcal{V}_G\,(\mathcal{B}_G\,(v_j,12\delta)\,,s,t) \cap \mathcal{V}_G\,(\mathcal{B}_G\,(v_\ell,12\delta)\,,s,t) = \emptyset$ for any $j \neq \ell$ since $\mathrm{dist}_G\,(v_j,v_\ell) > 30\delta$. ❑

4 Algorithmic Applications

In this section, we consider a few algorithmic applications of the bounds and proof techniques we showed in the previous section.

4.1 Network Design Application: Minimizing Bottleneck Edges

In this section we consider the following problem.

Problem 1 (Unweighted Uncapacitated Minimum Vulnerability problem (Uumv) [4, 27, 36]) The input to this problem a graph $G = (V, E)$, two nodes $s, t \in V$, and two positive integers $0 < r < \kappa$. The goal is to find a set of κ paths between s and t that minimizes the number of "shared edges", where an edge is called shared if it is in more than r of these κ paths between s and t. When $r = 1$, the Uumv problem is called the "minimum shared edges" (Mse) problem.

We will use the notation $\mathsf{OPT}_{\text{UUMV}}(G, s, t, r, \kappa)$ to denote the number of shared edges in an optimal solution of an instance of Uumv. Uumv has applications in several communication network design problems (see [34–36] for further details). The following computational complexity results are known regarding Uumv and Mse for a graph with n nodes and m edges (see [4, 27]):

- Mse does not admit a $2^{\log^{1-\varepsilon} n}$-approximation for any constant $\varepsilon > 0$ unless $\mathsf{NP} \subseteq \text{DTIME}\big(n^{\log\log n}\big)$.
- Uumv admits a $\lfloor \kappa/(r+1) \rfloor$-approximation. However, no non-trivial approximation of Uumv that depends on m and/or n only is currently known.
- Mse admits a $\min\left\{ n^{3/4},\, m^{1/2} \right\}$-approximation.

4.1.1 Greedy Fails for Uumv *or* Mse *Even for Hyperbolic Graphs (i.e., Graphs With Constant δ)*

Several routing problems have been looked at for hyperbolic graphs (*i.e.*, constant δ) in the literature before (*e.g.*, see [15, 23]) and, for these problems, it is often seen that simple greedy strategies *do* work. However, that is unfortunately not the case with Uumv or Mse. For example, one obvious greedy strategy that can be designed is as follows.

```
(* Greedy strategy *)
Repeat κ times
    Select a new path between s and t that shares a minimum number
        of edges with the already selected paths
```

The above greedy strategy can be arbitrarily bad even when $r = 1$, $\delta \leq 5/2$ and every node except s and t has degree at most three as illustrated in Fig. 5; even qualifying the greedy step by selecting a shortest path among those that increase the number of shared edges the least does not lead to a better solution.

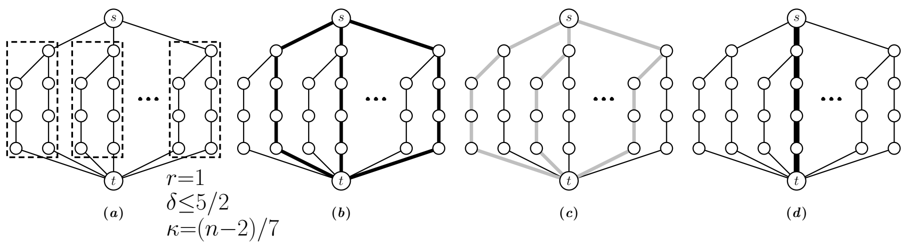

Fig. 5 A bad example for the obvious greedy strategy. (a) The given graph in which every node except s and t has degree at most 3 and $\delta \leq 5/2$. (b) Greedy first selects the $(n-2)/14$ edge-disjoint shortest paths shown in thick black. (c) Greedy then selects the shortest paths shown in light gray one by one, each of which increases the number of shared edges by one more. Thus, greedy uses $(n-2)/7$ shared edges. (d) An optimal solution uses only 5 edges, *i.e.*, $\mathsf{OPT}_{\text{UUMV}}(G, s, t, 1, \kappa) = 5$.

4.1.2 Improved Approximations for Uumv *or* Mse *for* δ *Up To* $o(\log n/\log d)$

Note that in the following lemma d is the maximum degree of any node "except s, t and any node within a distance of $35\,\delta$ of s" (degrees of these nodes may be arbitrary). For up to $\delta = o(\log n/\log d)$, the lemma provides the first non-trivial approximation of Uumv as a function of n only (independent of κ) and improves upon the currently best $\min\left\{n^{3/4},\, m^{1/2}\right\}$-approximation of Mse for arbitrary graphs.

Lemma 2 *Let d is the maximum degree of any node except s, t and any node within a distance of $35\,\delta$ of s (degrees of these nodes may be arbitrary). Then,* Uumv *(and, consequently also* Mse*) for a δ-hyperbolic graph G can be approximated within a factor of* $O\left(\max\left\{\log n,\, d^{\,O(\delta)}\right\}\right)$.

Remark 5 Thus for fixed d Lemma 2 provides improved approximation as long as $\delta = o(\log n)$. **Note that our approximation ratio is independent of the value of κ.** Also note that $\delta = \Omega(\log n)$ allows expander graphs as a sub-class of δ-hyperbolic graphs for which Uumv or Mse is expected to be harder to approximate.

Proof of Lemma 2

Our proof strategy has the following two steps:

- ▶ We define a new more general problem which we call the *edge hitting set problem for size constrained cuts* (Ehssc), and show that Uumv (and thus Mse) has the *same approximability properties as* Ehssc by characterizing optimal solutions of Uumv in terms of optimal solutions of Ehssc.
- ▶ We then provide a suitable approximation algorithm for Ehssc.

Problem 2 (Edge hitting set for size-constrained cuts (Ehssc)) The input to Ehssc is a graph $G = (V, E)$, two nodes $s, t \in V$, and a positive integer $0 < k \leq |E|$. Define a size-constrained s-t cut to be a s-t cut S such that the number of cut-edges $\text{cut}_G(S, s, t)$ is at most k. The goal of Ehssc is to find a hitting set of minimum cardinality for all size-constrained s-t cuts of G, i.e., find $\widetilde{E} \subset E$ such that $|\,\widetilde{E}\,|$ is minimum and

$$\forall\, s \in S \subset V \setminus \{t\}\colon\ |\,\mathcal{E}_G(S, s, t)\,| \leq k \ \Rightarrow\ \mathcal{E}_G(S, s, t) \cap \widetilde{E} \neq \emptyset$$

We will use the notation $E_{\text{Ehssc}}(G, s, t, k)$ to denote an optimal solution containing $\mathsf{OPT}_{\text{Ehssc}}(G, s, t, k)$ edges of an instance of Ehssc.

Lemma 3 (Relating Ehssc to Uumv)
$\mathsf{OPT}_{\text{Uumv}}(G, s, t, r, \kappa) = \mathsf{OPT}_{\text{Ehssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1)$.

Proof. Note that *any* feasible solution for Uumv *must* contain at least one edge from every collection of cut-edges $\mathcal{E}_G(S, s, t)$ satisfying $|\,\mathcal{E}_G(S, s, t)\,| \leq \lceil \kappa/r \rceil - 1$, since otherwise the number of paths going from $\mathcal{E}_G(S, s, t)$ to $V \setminus \mathcal{E}_G(S, s, t)$ is at most $r\left(\left\lceil \frac{\kappa}{r} \right\rceil - 1\right) < \kappa$. Thus we get $\mathsf{OPT}_{\text{Uumv}}(G, s, t, r, \kappa) \geq \mathsf{OPT}_{\text{Ehssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1)$.

On the other hand, $\mathsf{OPT}_{\text{Uumv}}(G, s, t, r, \kappa) \leq \mathsf{OPT}_{\text{Ehssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1)$ can be argued as follows. Consider the set of edges $E_{\text{Ehssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1)$ in an optimal hitting set, and set the capacity $c(e)$ of every edge e of G as

$$c(e) = \begin{cases} \infty, \text{ if } e \in E_{\text{Hssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1) \\ r, \text{ otherwise} \end{cases}$$

The value of the minimum s-t cut for G is then at least $\min\,\{\infty,\ r \times \lceil \kappa/r \rceil\} \geq \kappa$ which implies (by the standard max-flow-min-cut theorem) the existence of κ flows each of unit value. The paths taken by these κ flows provide our desired κ paths for Uumv. Note that at most r paths go through any edge e with $c(e) \neq \infty$ and thus $\mathsf{OPT}_{\text{Uumv}}(G, s, t, r, \kappa) \leq \left|\,\{e \,|\, c(e) \neq \infty\}\,\right| = \mathsf{OPT}_{\text{Ehssc}}\,(G, s, t, \lceil \kappa/r \rceil - 1)$. ❑

Now, we turn to providing a suitable approximation algorithm for Ehssc. Of course, Ehssc has the following obvious *exponential-size* LP-relaxation since it is after all a hitting set problem:

$$\begin{array}{l} \textit{minimize } \sum_{e \in E} x_e \quad \textit{subject to} \\ \qquad \forall\, s \in S \subset V \setminus \{t\} \text{ such that } \text{cut}_G(S, s, t) \leq k \ \ : \ \ \sum_{e \,\in\, \mathcal{E}_G(S,s,t)} x_e \geq 1 \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall\, e \in E \ \ : \ \ x_e \geq 0 \end{array}$$

Intuitively, there are at least two reasons why such a LP-relaxation may not be of sufficient interest. Firstly, known results may imply a large integrality gap. Secondly, it is even not very clear if the LP-relaxation can be solved exactly in a time efficient manner. Instead, we will exploit the hyperbolicity property and use Lemma 1 to derive our approximation algorithm.

Lemma 4 (Approximation algorithm for EHSSC) EHSSC *admits a* $O\left(\max\left\{\delta\,\log n,\ d^{\,O(\delta)}\right\}\right)$*-approximation.*

Proof. Our algorithm for EHSSC can be summarized as follows:

Algorithm for EHSSC

If $k \leq d^{\,12\delta+1}$ **then**
  $\mathcal{A} \leftarrow \emptyset$, $j \leftarrow 0$, set the capacity $c(e)$ of every edge e to 1
  while there exists a s-t cut of capacity at most k **do**
    $j \leftarrow j+1$, let $\mathcal{F}_j$ be the edges of a s-t cut of capacity at most k
    $\mathcal{A} \leftarrow \mathcal{A} \cup \mathcal{F}_j$, set $c(e)=\infty$ for every edge $e \in \mathcal{F}_j$
  return $\mathcal{A}$ as the solution
else (* $k > d^{\,12\delta+1}$ *)
  return all the edges in a shortest path between s and t as the solution $\mathcal{A}$

The following case analysis of the algorithm shows the desired approximation bound.

Case 1: $k \leq d^{\,12\delta+1}$. Let $\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_\ell$ be the sets whose edges were added to $\mathcal{A}$; thus, $|\mathcal{A}| \leq k\,\ell$. Since $|\mathcal{F}_j| \leq k$ and $\mathcal{F}_j \cap \mathcal{F}_{j'} = \emptyset$ for $j \neq j'$, $\mathsf{OPT}_{\text{EHSSC}}(G, s, t, k) \geq \ell$, thus providing an approximation bound of $k \leq d^{\,12\delta+1}$.

Case 2: $k > d^{\,12\delta+1}$ **and** $\text{dist}_G(s,t) \leq 48\,\delta + 8\,\delta\,\log n$. Since $\mathsf{OPT}_{\text{EHSSC}}(G, s, t, k) \geq 1$, this provides a $O(\delta\,\log n)$-approximation.

Case 3: $k > d^{\,12\delta+1}$ **and** $\text{dist}_G(s,t) > 48\,\delta + 8\,\delta\,\log n$. Use Lemma 1 to find a collection $S_1, S_2, \dots, S_\ell$ of $\ell = (\text{dist}_G(s,t) - 8\,\delta\,\log n)/(50\,\delta)$ edge and node disjoint s-t cuts. Since $\text{cut}_G\,(S_j, s, t) \leq d^{\,12\delta+1} < k$, any valid solution of EHSSC must select *at least* one edge from $\mathcal{E}_G\,(S_j, s, t)$. Since the cuts are edge and node disjoint, it follows that

$$\mathsf{OPT}_{\text{EHSSC}}(G, s, t, k) \geq \left(\text{dist}_G(s,t) - 8\,\delta\,\log n\right)/(50\,\delta)$$

Since we return all the edges in a shortest path between s and t as the solution, the approximation ratio achieved is $\text{dist}_G(s,t)/\left(\frac{\text{dist}_G(s,t)-8\,\delta\,\log n}{50\,\delta}\right) < 100\delta$. ❑

4.2 Application to the Small Set Expansion Problem

The *small set expansion* (SSE) problem was studied by Arora, Barak and Steurer in [3] (and also by several other researchers such as [5, 18, 29–31]) in an attempt to understand the computational difficulties surrounding the Unique Games Conjecture (UGC). To define SSE, we will also use the *normalized* edge-expansion of a graph which is defined as follows [14]. For a subset of nodes S of a graph G, let $\mathsf{vol}_G(S)$ denote the sum of degrees of the nodes in G. Then, the normalized edge expansion ratio $\Phi_G(S)$ of a subset S of nodes of at most $|V|/2$ nodes of G is defined as $\Phi_G(S) = \text{cut}_G(S)/\mathsf{vol}_G(S)$. Since we will deal with only d-regular graphs in this subsection, $\Phi_G(S)$ will simplify to $\text{cut}_G(S)/(d\,|S|\,)$.

Definition 6 ((SSE Problem) [a case of [3, Theorem 2.1], rewritten as a problem] Suppose that we are given a d-regular graph $G = (V, E)$ for some fixed d, and suppose G has a subset of at most ζn nodes S, for some constant $0 < \zeta < 1/2$, such that $\Phi_G(S) \le \varepsilon$ for some constant $0 < \varepsilon \le 1$. Then, find as efficiently as possible a subset S' of at most ζn nodes such that $\Phi_G(S) \le \eta\,\varepsilon$ for some "universal constant" $\eta > 0$.

In general, computing a very good approximation of the SSE problem seems to be quite hard; the approximation ratio of the algorithm presented in [30] roughly deteriorates proportional to $\sqrt{\log(1/\zeta)}$, and a $O(1)$-approximation described in [5] works only if the graph excludes two specific minors. The authors in [3] showed how to design a sub-exponential time (*i.e.*, $O\left(2^{c\,n}\right)$ time for some constant $c < 1$) algorithm for the above problem. As they remark, expander like graphs are somewhat easier instances of SSE for their algorithm, and it takes some non-trivial technical effort to handle the "non-expander" graphs. Note that *the class of δ-hyperbolic graphs for $\delta = o(\log n)$ is a non-trivial proper subclass of non-expander graphs.* We show that SSE (as defined in Definition 6) can be solved in polynomial time for such a proper subclass of non-expanders.

Lemma 5 (polynomial time solution of SSE for δ-hyperbolic graphs when δ is sub-logarithmic and d is sub-linear) *Suppose that G is a d-regular δ-hyperbolic graph. Then the* SSE *problem for G can be solved in polynomial time provided d and δ satisfy:*

$$d \le 2^{\log^{(1/3)-\rho} n} \quad \text{and} \quad \delta \le \log^{\rho} n \quad \text{for some constant } 0 < \rho < 1/3$$

Remark 6 Computing the minimum node expansion ratio of a graph is in general NP-hard and is in fact SSE-hard to approximate within a ratio of $C\sqrt{\mathfrak{h}_G \log d}$ for some constant $C > 0$ [24]. Since we show that SSE is polynomial-time solvable for δ-hyperbolic graphs for some parameter ranges, the hardness result of [24] does not directly apply for graph classes that belong to these cases, and thus additional arguments may be needed to establish similar hardness results for these classes of graphs.

Proof. Our proof is quite similar to that used for Theorems 1. But, instead of looking for smallest possible non-expansion bounds, we now relax the search and allow us to consider subsets of nodes whose expansion is just enough to satisfy the requirement. This relaxation helps us to ensure the size requirement of the subset we need to find.

We will use the construction in the proof of Theorem 1 in this proof, so *we urge the readers to familiarize themselves with the details of that proof before reading the current proof.* Note that $h_G(S) \le \varepsilon$ implies $\Phi_G(S) \le d\,h_G(S)/d \le$

ε. We select the nodes p and q such that $\Delta = \text{dist}_G(p,q) = \log_d n = \log n / \log d$, and set $\mu = 1/2$. Note that $(360 \log n) / \left(\Delta\, 2^{\,\Delta^\mu/(28\,\delta\, \log(2d))}\right) < (1/\Delta)^{1-\mu}$ since

$$\begin{aligned}
&(360 \log n) / \left(\Delta\, 2^{\,\Delta^\mu/(28\,\delta\, \log(2d))}\right) < (1/\Delta)^{1-\mu} \\
&\qquad \Leftarrow\ (360 \log d) / \left(2^{\,(\log n)^{1/2}/\left(56\,\delta\,(\log d)^{3/2}\right)}\right) < (\log d/\log n)^{1/2} \\
&\qquad \Leftarrow\ 9 + \log\log n/2 < \left((\log n)^{1/2}\right) / \left(56\, \log^{(1-\rho)/2} n\right) - \log^{(1-\rho)/2} n
\end{aligned}$$

and the last inequality clearly holds for sufficiently large n.

First, suppose that there exists $0 \leq r \leq \frac{\Delta}{3} - \alpha\Delta$ such that $h_G\big(\mathcal{B}_{G_{-\mathcal{C}}}(p,r)\big) = h_G\left(\mathcal{B}_G(p,r)\right) \leq \varepsilon$. We return $S' = \mathcal{B}_G(p,r)$ as our solution, To verify the size requirement, note that

$$\begin{aligned}
&|\,\mathcal{B}_G(p,r)\,| \leq |\,\mathcal{B}_G\left(p,(\Delta/3) - \alpha\,\Delta\right)| < |\,\mathcal{B}_G\left(p,\Delta/3\right)| \\
&\qquad\qquad < \sum_{i=0}^{\Delta/3} d^{\,i} < d^{(\Delta/3)+1} = d\,n^{1/3} < \zeta\, n \qquad (22)
\end{aligned}$$

where the last inequality follows since $d \leq 2^{\,\log^{(1/3)-\rho} n}$ and ζ is a constant.

Otherwise, no such r exists, and this implies

$$\begin{aligned}
&|\,\mathcal{B}_G\left(p,(\Delta/3) - \alpha\,\Delta\right)| \geq (1+\varepsilon)^{(\Delta/3)-\alpha\,\Delta} \\
&\qquad\qquad > (1+\varepsilon)^{\Delta/4} \geq \mathbf{e}^{\varepsilon\Delta/8} = \mathbf{e}^{\varepsilon\,\log_d n/8} = n^{\varepsilon\,\log_d \mathbf{e}/8}
\end{aligned}$$

Now there are two major cases as follows.

Case 1: there exists at least one path between p and q in $G_{-\mathcal{C}}$.

We know that $\text{dist}_{G_{-\mathcal{C}}}(p,q) \geq (\Delta/60)2^{\,\alpha\,\Delta/\delta}$ and (by choice of p) $\left|\,\mathcal{B}_{G_{-\mathcal{C}}}\left(p, \text{dist}_{G_{-\mathcal{C}}}(p,q)/2\right)\,\right| < n/2$. Let $p = u_0, u_1, \ldots, u_{t-1}, u_t = q$ be the nodes in successive order on a shortest path from p to q of length $t = \text{dist}_{G_{-\mathcal{C}}}(p,q)$. Perform a BFS starting from p in $G_{-\mathcal{C}}$, and let $\mathcal{L}_i$ be the sets of nodes at the i^{th} level (*i.e.*, $\forall\, u \in \mathcal{L}_i$: $\text{dist}_{G_{-\mathcal{C}}}(p,u) = i$). Note that $\left|\,\bigcup_{j=0}^{t/2} \mathcal{L}_j\,\right| \leq n/2$. Consider the levels $\mathcal{L}_0, \mathcal{L}_1, \ldots, \mathcal{L}_{t/2}$, and partition the ordered sequence of integers $0, 1, 2, \ldots, t/2$ into consecutive blocks $\Delta_0, \Delta_1, \ldots, \Delta_{(1+(t/2)\,)/\kappa - 1}$ each of length $\eta = (8/\varepsilon)\,\ln n$, *i.e.*,

$$\underbrace{0,1,\ldots,\eta-1}_{\Delta_0},\ \underbrace{\eta,\eta+1,\ldots,2\eta-1}_{\Delta_1},\ldots\ldots,\underbrace{(t/2)-\eta+1,(t/2)-\eta+2,\ldots,(t/2)}_{\Delta_{(1+(t/2)\,)/\eta-1}}$$

We claim that for every Δ_i, there exists an index i^* within Δ_i (*i.e.*, there exists an index $i\,\eta \leq i^* \leq i\,\eta + \eta - 1$) such that $h_G\left(\mathcal{L}_{i^*}\right) \leq \varepsilon$. Suppose for the sake of contradiction that this is not true. Then, it follows that

$$\forall\, i\,\eta \leq j \leq i\,\eta + \eta - 1 :\ h_{G_{-\mathcal{C}}}\left(\mathcal{L}_j\right) \geq h_G\left(\mathcal{L}_j\right)/2 > \varepsilon/2$$

$$\Rightarrow |\mathcal{L}_{i\,\eta+\eta-1}| > |\,\mathcal{L}_{i\,\eta}\,|\,(1+(\varepsilon/2)\,)^{\eta} \geq (1+(\varepsilon/2)\,)^{(8/\varepsilon)\,\ln n} \geq \mathbf{e}^{(\varepsilon/4)\,((8/\varepsilon)\,\ln n)} = n^2 > n$$

which contradicts the fact that $\left|\,\bigcup_{j=0}^{t/2}\mathcal{L}_j\,\right| \leq n/2$. Since $\sum_{i=0}^{(1+(t/2)\,)/\kappa-1} |\,\mathcal{L}_{i^*}\,| < n/2$, there exists a set $\mathcal{L}_{k^*}$ such that $h_G\,(\mathcal{L}_{k^*}\,) \leq \varepsilon$ and

$$|\,\mathcal{L}_{k^*}\,| < \frac{n/2}{(1+(t/2)\,)\,/\kappa} < n\kappa/t < (8\,n\,\ln n)/\left(\varepsilon\,(\varDelta/60)\,2^{\varDelta^{1/2}/(7\,\delta\,\log(2d))}\right) \leq \left(480\,n\log^{(1/3)-\rho} n\right)/\left(\varepsilon\,2^{\,(\log^{\rho/2} n)/14}\right) < \zeta\,n$$

Case 2: there is no path between p and q in $G_{-\mathcal{C}}$.

In this case, we return $\mathcal{B}_{G_{-\mathcal{C}}}\,(p,(\varDelta/3)-\alpha\,\varDelta) = \mathcal{B}_G\,(p,(\varDelta/3)-\alpha\,\varDelta)$ as our solution. The size requirement follows since $|\,\mathcal{B}_G\,(p,(\varDelta/3)-\alpha\,\varDelta)\,| < \zeta\,n$ was shown in (22). Note that nodes in $\mathcal{B}_G\,(p,(\varDelta/3)-\alpha\,\varDelta)$ can only be connected to nodes in $\mathcal{C}$, and thus

$$h_G\,(\mathcal{B}_G\,(p,(\varDelta/3)-\alpha\,\varDelta)\,) \leq |\,\mathcal{C}\,|\,/\,|\,\mathcal{B}_G\,(p,(\varDelta/3)-\alpha\,\varDelta)\,| \leq \left(\,(\varDelta/3)d^{\,\alpha\varDelta}\right)/\left(n^{\varepsilon\log_d \mathbf{e}/8}\right) < n^{\alpha-(\varepsilon\log_d \mathbf{e}/8)}\log n < n^{1/(7\,\varDelta^{1/2}\,\log(2d))\,-\,(\varepsilon/(8\ln d)\,)}\log n < \varepsilon$$

where the penultimate inequality follows since $\varDelta = \omega(1)$.

In all cases, the desired subset of nodes can be found in $O\left(n^2\log n\right)$ time.
❑

5 Conclusion and Open Problems

In this paper we have provided the first known non-trivial bounds on expansions and cut-sizes for graphs as a function of the hyperbolicity measure δ, and have shown how these bounds and their related proof techniques lead to improved algorithms for two related combinatorial problems. We hope that these results sill stimulate further research in characterizing the computational complexities of related combinatorial problems over asymptotic ranges of δ. In addition to the usual future research of improving our bounds, the following interesting research questions remain:

- Can one use Lemma 5 or similar results to get a polynomial-time solution of UGC for some asymptotic ranges of δ? An obvious recursive application using the approach in [3] encounters a hurdle since hyperbolicity is not a hereditary property (cf. Section 1.2.1), *i.e.*, removal of nodes or edges may change δ sharply; however, it is conceivable that a more clever approach may succeed.

► Can our bounds on expansions and cut-sizes be used to get an improved approximation for the multicut problem [33, Problem 18.1] provided $\delta = o(\log n)$?

Acknowledgements B. DasGupta, N. Mobasheri and F. Yahyanejad thankfully acknowledge supported from NSF grant IIS-1160995 for this research. M. Karpinski was supported in part by DFG grants. The problem of investigating expansion properties of δ-hyperbolic graphs was raised originally to some of the authors by A. Wigderson.

References

1. R. Albert, B. DasGupta and N. Mobasheri. *Topological implications of negative curvature for biological and social networks*, Physical Review E, 89 (3), 032811, 2014.
2. F. Ariaei, M. Lou, E. Jonckeere, B. Krishnamachari, and M. Zuniga. *Curvature of sensor network: clustering coefficient*, EURASIP Journal on Wireless Communications and Networking, 213185, 2008.
3. S. Arora, B. Barak and D. Steurer. *Subexponential Algorithms for Unique Games and Related Problems*, 51^{st} annual IEEE Symposium on Foundations of Computer Science, 563-572, 2010.
4. S. Assadi, E. Emamjomeh-Zadeh, A. Norouzi-Fard, S. Yazdanbod and H. Zarrabi-Zadeh. *The Minimum Vulnerability Problem*, Algorithmica, 70, 718-731, 2014.
5. N. Bansal, U. Feige, R. Krauthgamer, K. Makarychev, V. Nagarajan, J. Naor and R. Schwartz. *Min-max Graph Partitioning and Small Set Expansion*, 52^{nd} Annual IEEE Symposium on Foundations of Computer Science, 17-26, 2011.
6. I. Benjamini. *Expanders are not hyperbolic*, Israel Journal of Mathematics, 108, 33-36, 1998.
7. I. Benjamini, C. Hoppen, E. Ofek, P. Pralat and N. Wormald. *Geodesics and almost geodesic cycles in random regular graphs*, Journal of Graph Theory, 66, 115136, 2011.
8. I. Benjamini and O. Schramm. *Finite transitive graph embedding into a hyperbolic metric space must stretch or squeeze*, Geometric Aspects of Functional Analysis, 123-126, Springer, 2012.
9. H. L. Bodlaender. *Dynamic programming on graphs with bounded treewidth*, in Lecture Notes in Computer Science 317, T. Lepistö and A. Salomaa (Eds.), 105-118, Springer, 1988.
10. M. R. Bridson and A. Haefliger. *Metric Spaces of Non-Positive Curvature*, Springer, 1999.
11. V. Chepoi, F. F. Dragan, B. Estellon, M. Habib and Y. Vaxès. *Diameters, centers, and approximating trees of δ-hyperbolic geodesic spaces and graphs*, proceedings of the 24^{th} Annual Symposium on Computational geometry, 59-68, 2008.
12. V. Chepoi, F. F. Dragan, B. Estellon, M. Habib, Y. Vaxès and Y. Xiang. *Additive spanners and distance and routing labeling schemes for δ-hyperbolic graphs*, Algorithmica, 62(3-4), 713-732, 2012.
13. V. Chepoi and B. Estellon. *Packing and covering δ-hyperbolic spaces by balls*, in Lecture Notes in Computer Science 4627, M. Charikar, K. Jansen, O. Reingold and J. D. P. Rolim (Eds.), 59-73, Springer, 2007.
14. F. R. K. Chung. *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, 92, 1997.
15. D. Eppstein and M. T. Goodrich. *Succinct Greedy Geometric Routing Using Hyperbolic Geometry*, IEEE Transactions on Computers, 60(11), 1571-1580, 2011.
16. G. Even, G. Kortsarz and W. Slany. *On network design problems: fixed cost flows and the covering steiner problem*, ACM Transaction on Algorithms, 1(1), 74-101, 2005.
17. H. Fournier, A. Ismail, and A. Vigneron. *Computing the Gromov hyperbolicity of a discrete metric space*, Information Processing Letters, 115, 6-8, 576579, 2015.
18. R. Gandhi and G. Kortsarz. *On edge expansion problems and the small set expansion conjecture*, Discrete Applied Mathematics, 194, 93-101, 2015.

19. C. Gavoille and O. Ly. *Distance labeling in hyperbolic graphs*, in Lecture Notes in Computer Science 3827, X. Deng and D.-Z. Du (Eds.), 1071-1079, Springer, 2005.
20. M. Gromov. Hyperbolic groups, *Essays in group theory*, 8, 75-263, 1987.
21. E. Jonckheere, P. Lohsoonthorn, and F. Bonahon. *Scaled Gromov hyperbolic graphs*, Journal of Graph Theory, 57(2), 157-180, 2007.
22. E. Jonckheerea, M. Loua, F. Bonahona and Y. Baryshnikova. *Euclidean versus hyperbolic congestion in idealized versus experimental networks*, Internet Mathematics, 7(1), 2011.
23. R. Kleinberg. *Geographic Routing Using Hyperbolic Space*, 26^{th} IEEE International Conference on Computer Communications, 1902-1909, 2007.
24. A. Louis, P. Raghavendra and S. Vempala. *The Complexity of Approximating Vertex Expansion*, 54^{th} IEEE Annual Symposium on Foundations of Computer Science, 360-369, 2013.
25. A. Malyshev. *Expanders are order diameter non-hyperbolic*, arXiv:1501.07904, 2015.
26. F. de Montgolfier, M. Soto and L. Viennot. *Treewidth and Hyperbolicity of the Internet*, proceedings of the 10^{th} IEEE International Symposium on Networking Computing and Applications, 25-32, 2011.
27. M. T. Omran, J.-R. Sack and H. Zarrabi-Zadeh. *Finding paths with minimum shared edges*, Journal of Combinatorial Optimization, 26(4), 709-722, 2013.
28. F. Papadopoulos, D. Krioukov, M. Boguna and A. Vahdat. *Greedy Forwarding in Dynamic Scale-Free Networks Embedded in Hyperbolic Metric Spaces*, IEEE INFOCOM, 1-9, 2010.
29. P. Raghavendra and D. Steurer. *Graph Expansion and the Unique Games Conjecture*, 45^{th} ACM Symposium on Theory of Computing, 755-764, 2010.
30. P. Raghavendra, D. Steurer and P. Tetali. *Approximations for the isoperimetric and spectral profile of graphs and related parameters*, 45^{th} ACM Symposium on Theory of Computing, 631-640, 2010.
31. P. Raghavendra, D. Steurer and M. Tulsiani. *Reductions Between Expansion Problems*, IEEE Conference on Computational Complexity, 64-73, 2012.
32. N. Robertson and P. D. Seymour. *Graph minors. i. excluding a forest*. Journal of Combinatorial Theory, Series B, 35(1), 39-61, 1983.
33. V. Vazirani. *Approximation Algorithms*, Springer-Verlag, 2001.
34. J. Wang, M. Yang, B. Yang and S. Q. Zheng. *Dual-homing based scalable partial multicast protection*, IEEE Transaction on Computers, 55 (9), 1130-1141, 2006.
35. B. Yang, M. Yang, J. Wang and S. Q. Zheng. *Minimum cost paths subject to minimum vulnerability for reliable communications*, 8^{th} International Symposium on Parallel Architectures, Algorithms and Networks, 334-339, 2005.
36. S. Q. Zheng, J. Wang, B. Yang and M. Yang. *Minimum-cost multiple paths subject to minimum link and node sharing in a network*, IEEE/ACM Transaction on Networking, 18 (5), 1436-1449, 2010.