Effect of Gromov-Hyperbolicity Parameters on Cuts and Expansions in Graphs and Some Algorithmic Implications

(Revised Version)

Bhaskar DasGupta*  Marek Karpinski†  Nasim Mobasheri*  Farzaneh Yahyanejad*

Abstract

$\delta$-hyperbolic graphs, originally conceived by Gromov in 1987, include non-trivial interesting classes of “non-expander” graphs; for fixed $\delta$, such graphs are simply called hyperbolic graphs. Our goal in this paper is to study the effect of the hyperbolicity measure $\delta$ on expansion and cut-size bounds on graphs (here $\delta$ need not be a constant, i.e., the graph is not necessarily hyperbolic), and investigate up to what values of $\delta$ these results could provide improved approximation algorithms for related combinatorial problems. To this effect, we provide the following results.

• We provide constructive bounds on node expansions and cut-sizes for $\delta$-hyperbolic graphs as a function of $\delta$, and show that witnesses for such non-expansion or cut-size can be computed efficiently in polynomial time. 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 small number of cut-edges when $s$ and $t$ are sufficiently far apart.

• We also provide the following algorithmic consequences of these bounds and their related proof techniques for a few problems related to cuts and paths for $\delta$-hyperbolic graphs (where $\delta$ need not necessarily a constant but may be a function $f$ of the number of nodes, the exact nature of growth of $f$ depends on the problem considered):

  - We provide improved approximation algorithms for minimizing the number of bottleneck edges that arises in network design applications. En route, we also formulate the hitting set problem of size-constrained cuts and show a connection between approximability issues of these two problems.

  - We provide a polynomial-time solution for a type of small-set expansion problem that arises in the investigation of unique games conjecture.

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

*Department of Computer Science, University of Illinois at Chicago, Chicago, IL 60607. Emails: {bdasgup,nmobash2,fyahya2}@uic.edu. Supported by NSF grant IIS-1160995.

†Department of Computer Science, University of Bonn, Bonn 53113. Email: marek@cs.uni-bonn.de. Supported in part by DFG grants.
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. The hyperbolicity measure was originally conceived in a somewhat different group-theoretic context by Gromov in 1987 [20] from an observation that many results concerning the fundamental group of a Riemann surface hold true in a more general context. The measure was first defined for infinite continuous metric space with bounded local geometry via properties of geodesics [10], but was later also adopted for finite graphs. Off late, there has been a surge of theoretical and empirical works measuring and analyzing the hyperbolicity of networks, and many real-world networks have been reported to be hyperbolic. For example, preferential attachment scale-free networks were reported to be hyperbolic with appropriate scaling (normalization) in [21], networks of high power transceivers in a wireless sensor network were empirically observed to be hyperbolic in [2], communication networks at the IP layer and at other levels were empirically observed to be hyperbolic in [28], an assorted set of biological and social networks were empirically observed to be hyperbolic in [1], and 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 hyperbolicity of the network. On the other hand, theoretical investigations have revealed that expanders, vertex-transitive graphs and classical Erdős-Rényi random graphs are not hyperbolic [6–8, 25].

A major motivation for the investigations carried out in this paper is the following\textsuperscript{1}:

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

Since arbitrarily large $\delta$ leads to the class of all possible graphs, our hope is that investigations of this type will provide a characterization of hard graph instances for combinatorial problems via a lower bound on $\delta$. 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) of small expansion or small cut-size. We also provide some algorithmic consequences of these bounds and their related proof techniques for a few problems related to cuts and paths for hyperbolic 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

\textsuperscript{1}This is in contrast to many research works in this area where one studies the properties of $\delta$-hyperbolic graphs assuming $\delta$ to be fixed.
undirected unweighted graph \( G = (V, E) \) of hyperbolicity \( \delta \) 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 \( \lfloor n/3 \rfloor \) 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.
- \( \overline{u,v} \) 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 \( \infty \) if there is no path between \( u \) and \( v \) in \( H \)).
- \( D(H) = \max_{u,v \in V} \{ \text{dist}_H(u,v) \} \) is the diameter of the graph \( H = (V', E') \). Thus, in particular, for our input graph \( G \) 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.,
  \[ \partial_H(S) = \{ u \in V' \setminus S \mid v \in S \& \{u,v\} \in E' \} \]
  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 \( S \).
- \( B_H(u,r) \) is the set of nodes contained in a ball of radius \( r \) centered at node \( u \) in a graph \( H \), i.e., \( B_H(u,r) = \{ v \mid \text{dist}_H(u,v) \leq r \} \)

### 1.2 Formal Definitions of Gromov-hyperbolicity

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

**Definition 1 (\( \delta \)-hyperbolic graphs via geodesic triangles)** A graph \( G \) has a (Gromov) hyperbolicity of \( \delta = \delta(G) \), or simply is \( \delta \)-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 \( \delta \)-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 \( \delta \)-hyperbolic graph is simply called a hyperbolic graph if \( \delta \) 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 \( \delta \)-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 \( \delta \)-hyperbolic graphs via the following 4-node conditions.
Definition 3 (equivalent definition of $\delta$-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

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

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

$$\delta = \max_{u_1,u_2,u_3,u_4 \in V} \{ \rho_{u_1,u_2,u_3,u_4} \}.$$

It is well-known (e.g., see [10]) that Definition 1 and Definition 3 of $\delta$-hyperbolicity are equivalent in the sense that they are related by a constant multiplicative factor, i.e., there is a constant $c > 0$ such that if a graph $G$ is $\delta_1$-hyperbolic and $\delta_2$-hyperbolic via Definition 1 and Definition 3, respectively, then $\frac{1}{c} \delta_1 \leq \delta_2 \leq c \delta_1$. 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 $\delta$ and a 2-approximation of $\delta$ in $O(n^{3.69})$ and in $O(n^{2.69})$ 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 $\delta$

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

The hyperbolicity property is not hereditary (and thus also not monotone). For example, see Fig. 1(b). The examples in Fig. 1(a) and Fig. 1(b) also show that removing a single node or edge can increase/decrease the value of $\delta$ abruptly.

“Close to hyperbolic topology” is not necessarily the same as “close to tree topology”. For example, all bounded-diameter graphs are also hyperbolic graphs irrespective of whether they are tree or not (however, hyperbolic graphs need not be of bounded diameter). In general, even for small $\delta$, the metric induced by a $\delta$-hyperbolic graph may be quite far from a tree metric [11].

Hyperbolicity is not necessarily the same as tree-width. A somewhat related 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.

Examples of hyperbolic graph classes (i.e., when $\delta$ 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 $\delta$ is not a constant) include expanders, simple cycles, and, for some parameter ranges, the Erdős–Rényi random graphs.

Note that if $G$ is $\delta$-hyperbolic then $G$ is also $\delta'$-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 $\frac{1}{\pi}$-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].

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

Fact 2 (Exponential divergence of geodesic rays) [Simplified reformulation of [1, Theorem 10]] Assume that $G$ is a $\delta$-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 $\text{dist}_G(v, u_1) = \text{dist}_G(v, u_1) = r$, $\text{dist}_G(u_1, u_2) \geq 3\kappa \delta$, $\text{dist}_G(v, u_3) = \text{dist}_G(v, u_4) = r + \alpha$, and $\text{dist}_G(u_1, u_4) = \text{dist}_G(u_2, u_3) = \alpha$.

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

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 $(G, d, \delta)$ as a function of $n$, $d$, and $\delta$. These two results, namely Theorem 6 and Theorem 8, 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 (laminar) 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 $G$.

To illustrate the non-trivialness of these bounds, suppose that the maximum degree $d$ and the hyperbolicity value $\delta$ grows asymptotically very slowly with respect to the number of nodes $n$, and the diameter $D$ to be of the order of the minimum possible value of $\log_d n$. In Remark 1, we provide an explanation of the asymptotics of these bounds in comparison to expander-type
graphs. In particular, if \( \delta \) 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
\( \delta \) need to be at least \( \Omega\left(\log n\right) \) 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 9, 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
be found in polynomial time in \( \delta \)-hyperbolic graphs when every node other than \( s \) and \( t \) has a
maximum degree of \( d \) and the distance between \( s \) and \( t \) is at least \( \Omega(\delta \log n) \). This result was 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 \( \delta \)-hyperbolic graphs when \( \delta \) does not grow too
fast as a function of \( n \):

  - We show in Section 4.1 (Lemma 10) 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 11) and providing an improved approximation for this latter problem (Lemma 12).
  - We also observe that obviously greedy strategies fail for such problems miserably.
  - Finally, in Section 4.2 we provide a polynomial-time solution (Lemma 14) for a type of small-
set expansion problem originally proposed by Arora, Barak and Steurer [3] for the case when
\( \delta \) is sub-logarithmic in \( n \).

## 3 Effect of \( \delta \) on Expansions and Cuts in \( \delta \)-hyperbolic Graphs

The two results in this section are related to the node (or edge) expansion ratios of a graph that
is \( \delta \)-hyperbolic for some (not necessarily constant) \( \delta \). 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 \( h_G(S) = \frac{|G_e(S)|}{|S|} \). 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 \( g_H(S) = \frac{|\text{cut}_G(S)|}{|S|} \). 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** \( h_G = \min_{S \subseteq V : |S| \leq |V|/2} \{h_G(S)\} \) will denote the minimum node expansion of a graph \( G = (V, E) \).

Since any subset \( S \) containing exactly \( |V|/2 \) nodes has \( |\partial G(S)| \leq |V|/2 \), \( h_G \) satisfies \( 0 < h_G \leq 1 \)
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

A family of sets $S_1, S_2, \ldots, S_\ell$ 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 $\text{cut}_G(S, p, q)$ corresponding to the cut $S$ is the set of edges with exactly one end-point in $S$, i.e.,

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

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

**Theorem 6** For any constant $0 < \mu < 1$, the following result holds for $(G, d, \delta)$. Let $p$ and $q$ be any two nodes of $G$ and let $\Delta = \text{dist}_G(p, q)$. Then, there exists at least $t = \max\left\{ \frac{\Delta^\mu}{50 \log q}, 1 \right\}$ 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:

- $\forall j \in \{1, 2, \ldots, t\}$: $h_G(S_j) \leq \min\left\{ \frac{8 \ln n}{\Delta}, \max\left\{ \left( \frac{1}{\Delta} \right)^{1-\mu}, \frac{500 \ln n}{\Delta 2^{\frac{\Delta^\mu}{50 \log q}}} \right\} \right\}$.
- All the subsets can be found in a total of $O(n^3 \log n + mn^2)$ time.
- Either all the subsets $S_1, S_2, \ldots, S_t$ contain the node $p$, or all of them contain the node $q$.

**Corollary 7** Letting $p$ and $q$ be two nodes such that $\text{dist}_G(p, q) = D(G) = D$ realizes the diameter of the graph $G$, we get the bound:

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

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

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

(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 $O(1/\log^{1-\mu} n)$ 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 $h_G = \Omega(1)$ for an expander):

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

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

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

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

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

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

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

and thus $\delta$ need to be at least $\Omega(\log n)$ to get a constant upper bound.
3.1.1 Proof of Theorem 6

Proof of the main bounds in Theorem 6 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 \( |B_G(p,\Delta/2)| \leq \min \{ |B_G(p,\Delta/2)|, |B_G(q,\Delta/2)| \} \leq n/2 \). Consider the sequence of balls \( B_G(p, r) \) for \( r = 0, 1, 2, \ldots, \Delta/2 \). Thus it follows that

\[
n/2 > |B_G(p,\Delta/2)| \geq \prod_{\ell=0}^{(\Delta/2)-1} (1 + h_G(B_G(p,\ell))) \geq \prod_{\ell=0}^{(\Delta/2)-1} e^{h_G(B_G(p,\ell))/2} = e^{\sum_{\ell=0}^{(\Delta/2)-1} h_G(B_G(p,\ell))/2} \frac{\Delta}{2} < (4 \ln (n/2))/\Delta
\]

By a simple averaging argument, there must now exist \( \Delta/4 > \max \{\Delta^\mu/(56 \log d), 1\} \) distinct balls (subsets of nodes) \( B_G(p, r_1) \subset B_G(p, r_2) \subset \cdots \subset B_G(p/r_{\Delta/4}) \) such that \( |B_G(p, r_j)| < (8 \ln (n/2))/\Delta \) 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., \( h_G(S_j) \leq \max \left\{ \left( \frac{1}{\Delta} \right)^{1-\mu}, \frac{500 \ln n}{\Delta 2^{28 \delta \log(2d)}} \right\} \)

(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 \( (500 \ln n)/(\Delta 2^{\Delta^\mu/(28 \delta \log(2d)))} > (500 \ln n)/(\Delta 2^{(1/14)\Delta^\mu}) > 1 \). Thus, any subset of \( n/2 \) nodes containing \( p \) satisfies the claimed bound, and the number of such subsets is \( \left( \frac{n}{2} - 1 \right)^t \) \( \Rightarrow t \).

(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 \( \text{dist}_G(p,p') = \text{dist}_G(p',q') = \text{dist}_G(q',q) = \Delta/3 \). The following initial value of the parameter \( \alpha \) is crucial to our analysis:

\[
\alpha = \alpha_0 = 1/(7 \Delta^{1-\mu} \log(2d))
\]

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 \( p', q' \) between \( p' \) and \( q' \). Thus,

\[
\mathcal{C} = \{ u | \exists v \in p', q': \text{dist}_G(u,v) = \lfloor \alpha \Delta \rfloor \} \Rightarrow |\mathcal{C}| \leq (\Delta/3) d^{|\alpha \Delta|} < (\Delta/3) d^{\alpha \Delta}
\]

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

\[
\text{dist}_{G_{-\mathcal{C}}}(p,q) \geq (\Delta/60) 2^{\alpha \Delta/\delta}
\]

\(^2\)We will later need to vary the value of \( \alpha \) in our analysis.
Let $B_G(p, r)$ be the ball of radius $r$ centered at node $p$ in $G$ with $|B_G(p, r)| \leq n/2$, and let $h_{B(p,j)} \overset{\text{def}}{=} \left( \sum_{j=0}^{r-1} \ell \right) / 2$ and $B_{G-C}(q, dist_{G-C}(p, q)/2)$ contain all the nodes reachable from $p$ and $q$, respectively, in $G-C$.

Note that if there is no path between nodes $p$ and $q$ in $G-C$ then dist$_{G-C}(p, q) = \infty$ and $B_{G-C}(p, \text{dist}_{G-C}(p,q)/2)$ and $B_{G-C}(q, \text{dist}_{G-C}(p,q)/2)$ contain all the nodes reachable from $p$ and $q$, respectively, in $G-C$.

Assume without loss of generality that $|B_{G-C}(p, \text{dist}_{G-C}(p,q)/2)| \leq (n-|C|)/2 < n/2$.

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

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

$$\sum_{\ell=0}^{(\Delta/3)-\alpha\Delta-1} h_G(B_G(p, \ell)) > \left( (\text{dist}_{G-C}(p,q)/2) - (t-1) \right) (1/\Delta)^1-\mu > ((\Delta/3) - \alpha\Delta - t) (1/\Delta)^1-\mu > \Delta^\mu/4 \tag{7}$$

Let $r_p$ be the least integer such that $B_{G-C}(p, r_p) = B_{G-C}(p, r_p + 1)$. Since $G$ is a connected graph and, for all $r \leq (\Delta/3) - \alpha\Delta$ we have $B_G(p, r) \cap C = \emptyset \equiv B_{G-C}(p, r) = B_G(p, r)$ we have $r_p \geq (\Delta/3) - \alpha\Delta$ (see Fig. 2(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 $G-C$). Consequently, we may not be able to use our current technique of enlarging the ball $B_{G-C}(p, r)$ 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 $B_{G-C}(p, r)$ in $G-C$ may differ from that in $G$, i.e., $h_G(B_{G-C}(p, r))$ need not be the same as $h_G(B_{G-C}(p, r))$. 

---

Note that if there is no path between nodes $p$ and $q$ in $G-C$ then dist$_{G-C}(p, q) = \infty$ and $B_{G-C}(p, \text{dist}_{G-C}(p,q)/2)$ and $B_{G-C}(q, \text{dist}_{G-C}(p,q)/2)$ contain all the nodes reachable from $p$ and $q$, respectively, in $G-C$. 

---
Rectifying the current strategy

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

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

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

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

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 $C_{\alpha_1 \Delta - t}$ does not disconnect $p$ from $q$ in the corresponding graphs $G - C_{\alpha_1 \Delta - t}$. This implies $r_{p, \alpha_1 \Delta - t} > \text{dist}_{G - C_{\alpha_1 \Delta - t}}(p, q)/2$. For notational convenience, we will denote $C_{\alpha_1 \Delta - t}$ and $G - C_{\alpha_1 \Delta - t}$ simply by $C$ and $G - C$, 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 $\alpha_0$ ensures that removal of nodes in $C$ does not decrease the expansion of the balls $B_{G - C}(p, r)$ in the new graph $G - C$ by more than a constant factor.

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

$$\forall (\Delta / 3) - \alpha_0 \Delta < r \leq \text{dist}_{G - C}(p, q) / 2 :$$

$$h_{G}(B_{G - C}(p, r)) > (1/\Delta)^{1-\mu} \Rightarrow h_{G - C}(B_{G - C}(p, r - 1)) \geq h_{G}(B_{G - C}(p, r - 1)) / 2$$

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

\[
\begin{align*}
    & h_{G-c}(B_{G-c}(p,r-1)) \geq h_G(B_{G-c}(p,r-1))/2 \\
    \iff & \left| \partial G(B_{G-c}(p,r-1)) - |C| \right| \geq h_G(B_{G-c}(p,r-1))/2 \\
    \iff & |\partial G(B_{G-c}(p,r-1))| \leq |C| \geq h_G(B_{G-c}(p,r-1))/2 \\
    \iff & h_G(B_{G-c}(p,r-1)) - |C| \geq h_G(B_{G-c}(p,r-1))/2 \\
    \iff & \frac{2|C|}{h_G(B_{G-c}(p,r-1))} \leq |B_{G-c}(p,r-1)| \\
    \iff & \frac{2|C|}{h_G(B_{G-c}(p,r-1))} \leq e^{\Delta/\mu}, \quad \text{since} \quad |B_{G-c}(p,r-1)| \geq |B_{G-c}(p,(\Delta/3) - \alpha_0 \Delta)| \\
    \iff & |B_{G-c}(p,(\Delta/3) - \alpha_0 \Delta)| \geq |B_{G-c}(p,(\Delta/3) - \alpha_1 \Delta)| > e^{\Delta/\mu} \\
    \iff & ((\Delta/3)d^{\alpha_0 \Delta}) \left(2/\left(h_G(B_{G-c}(p,r-1))\right)\right) \leq e^{\Delta/\mu}, \quad \text{since} \quad |C| < (\Delta/3)d^{\alpha_0 \Delta} \\
    \iff & (\Delta/8) \geq \ln \Delta + \alpha_1 \Delta d - \ln(3/2) - \ln(h_G(B_{G-c}(p,r-1))) \\
    \iff & (\Delta/8) \geq \ln \Delta + \alpha_1 \Delta d - \ln(3/2) - \ln(h_G(B_{G-c}(p,r-1))), \quad \text{since} \quad \alpha_0 \leq \alpha_1 \\
    \iff & \alpha_1 \leq \frac{(\Delta/8) - \ln \Delta + \ln(h_G(B_{G-c}(p,r-1)))}{\Delta \ln d} \tag{10}
\end{align*}
\]

Now, if \( h_G(B_{G-c}(p,r-1)) > (1/\Delta)^{1-\mu} \) then since \( \Delta = \omega(1) \) we have:

\[
(\Delta/8) - \ln \Delta + \ln(h_G(B_{G-c}(p,r-1))) > (\Delta/8) - \ln \Delta - (1 - \mu) \ln \Delta > (\Delta/7)
\]

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

**Second goal:** Use the first goal and the fact that \( \text{dist}_{G-c(p,q)} \) 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 \( \{ (\Delta/3) - \alpha_0 \Delta + 1, (\Delta/3) - \alpha_0 \Delta + 2, \ldots, \text{dist}_{G-c(p,q)}/2 \} \) such that

\[
\forall 1 \leq s \leq t : h_G(B_{G-c}(p,i_s)) \leq (1/\Delta)^{1-\mu} \tag{11}
\]

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

\[
\forall 1 \leq s \leq \xi : h_G(B_{G-c}(p,j_s)) > (1/\Delta)^{1-\mu} \Rightarrow \forall 1 \leq s \leq \xi : h_{G-c}(B_{G-c}(p,j_s)) \geq h_G(B_{G-c}(p,j_s))/2 \tag{12}
\]
This in turn implies
\[
|B_{G-C}(p, \text{dist}_{G-C}(p,q)/2)| > \left(\prod_{j=0}^{(\Delta/3)-\alpha_0\Delta-1} (1 + h_G(B_{G-C}(p,j))) \right) \left(\prod_{j=(\Delta/3)-\alpha_0\Delta}^{(\Delta/3)-\alpha_0\Delta+\xi-1} \left(1 + \left(h_G(B_{G-C}(p,j)) / 2\right)\right)\right)
\]
using (12)
\[
> \left(\prod_{j=0}^{(\Delta/3)-\alpha_0\Delta-1} e^{h_G(B_{G-C}(p,j)) / 2}\right) \left(\prod_{j=(\Delta/3)-\alpha_0\Delta}^{(\Delta/3)-\alpha_0\Delta+\xi-1} \left(1 + \left(h_G(B_{G-C}(p,j)) / 4\right)\right)\right)
\]
\[
= \left(\prod_{j=0}^{(\Delta/3)-\alpha_0\Delta-1} e^{-\alpha_0\Delta / (\Delta/3)-\alpha_0\Delta}\right) \left(\sum_{j=0}^{(\Delta/3)-\alpha_0\Delta+\xi-1} h_G(B_{G-C}(p,j)) / 4\right) > e^{j=0} \sum_{j=0}^{(\Delta/3)-\alpha_0\Delta+\xi-1} h_G(B_{G-C}(p,j)) < 4 \ln n
\]

Using (13) and our specific choice of the node \(p\) (over node \(q\)), we have
\[
n/2 > |B_{G-C}(p, \text{dist}_{G-C}(p,q)/2)| \Rightarrow \sum_{j=0}^{(\Delta/3)-\alpha_0\Delta+\xi-1} h_G(B_{G-C}(p,j)) < 4 \ln n
\]

We now claim that there must exist a set of \(t = \Delta^\mu/(56 \log d)\) distinct indices \(i_1 < i_2 < \ldots < i_t\) in \(\{0,1,\ldots,(\Delta/3) - \alpha_0\Delta + 1\}\) such that
\[
\forall 1 \leq s \leq t : h_G(B_{G-C}(p,i_s)) \leq (500 \ln n) / \left(\Delta 2^{\Delta^\mu/(28\delta \log(2d))}\right)
\]
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:
\[
4 \ln n > \sum_{j=0}^{(\Delta/3)-\alpha_0\Delta+\xi-1} h_G(B_{G-C}(p,j))
\]
\[
> \left((\Delta/3) - \alpha_0\Delta + \xi - (\Delta^\mu/(56 \log d)) + 1\right) \left(500 \ln n / \left(\Delta 2^{\Delta^\mu/(28\delta \log(2d))}\right)\right)
\]
\[
\Rightarrow \left(\text{dist}_{G-C}(p,q)/2\right) - \max\{1, (\Delta^\mu/(28 \log d))\} \left(500 \ln n / \left(\Delta 2^{\Delta^\mu/(28\delta \log(2d))}\right)\right) < 4 \ln n,
\]
substituting the values of \(t\) and \(\xi\)
\[
\Rightarrow \left(\Delta/120\right) \left(2^{(\alpha_1)^2/(2\delta)} - \max\{1, (\Delta^\mu/(28 \log d))\}\right) \left(125 / \left(\Delta 2^{\Delta^\mu/(28\delta \log(2d))}\right)\right) < 4 \ln n,
\]
by (4) and since \(\alpha_1/2 < \alpha_0\)
\[
\equiv \left(\Delta/120\right) \left(2^{\Delta^\mu/(28\delta \log(2d))} - \max\{1, (\Delta^\mu/(28 \log d))\}\right) \left(125 / \left(\Delta 2^{\Delta^\mu/(28\delta \log(2d))}\right)\right) < 1
\]
\[
\Rightarrow \left(\Delta/121\right) \left(2^{\Delta^\mu/(28\delta \log(2d))}\right) < 1 \equiv 125/121 < 1, \text{ since } \Delta = \omega(1)
\]

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 \(B_{G-C}(p,i_1) \subset B_{G-C}(p,i_2) \subset \cdots \subset B_{G-C}(p,i_t)\) 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 \(O(n^2 \log n + mn)\) 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 $\alpha$ considered in the proof.
  For each $\alpha$, the following steps are needed:
  
  \begin{itemize}
    \item Use BFS find the set of nodes $C$ in $O(n^2 + mn)$ time.
    \item Compute $G_{-C}$ in $O(m + n)$ time.
    \item Using BFS, compute $B_{G_{-C}}(p, r)$ for every $0 \leq r \leq \text{dist}_{G_{-C}}(p, q)/2$ in $O(m + n)$ time.
    \item Compute $h_G(B_{G_{-C}}(p, r))$ for every $0 \leq r \leq \text{dist}_{G_{-C}}(p, q)/2$ in $O(n^2 + mn)$ time, and select a subset of nodes with a minimum expansion.
  \end{itemize}

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 8** 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 $(G, d, \delta)$: there exists $\lfloor \tau/4 \rfloor$ distinct collections of subsets of nodes $\emptyset \subset F_1, F_2, \ldots, F_{\lfloor \tau/4 \rfloor} \subset 2^V$ such that

\begin{itemize}
  \item $\forall j \in \{1, \ldots, \lfloor \tau/4 \rfloor\}$ $\forall S \in F_j : h_G(S) \leq \max \left\{\left(\frac{1}{(\Delta/\tau)}\right)^{1-\mu}, \frac{360 \log n}{(\Delta/\tau)^{2\frac{(\Delta/\tau)^\mu}{\log(2d)}}}\right\}$.
  \item Each collection $F_j$ has at least $t_j = \max\left\{\frac{(\Delta/\tau)^\mu}{\log d}, 1\right\}$ subsets $V_{j,1}, \ldots, V_{j,t_j}$ that form a nested family, i.e., $\emptyset \subset V_{j,1} \subset V_{j,2} \subset \cdots \subset V_{j,t_j} \subset V$.
  \item All the subsets in each $F_j$ can be found in a total of $O\left(n^3 \log n + mn^2\right)$ time.
  \item (limited overlap claim) For every pair of subsets $V_{i,k} \in F_i$ and $V_{j,k'} \in F_j$ 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.
\end{itemize}

**Remark 2** Consider a bounded-degree hyperbolic graph, i.e., assume that $\delta$ 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 $(1/\Delta)(1-\mu)^{1/2}$, 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 $\tau$ satisfies the following:

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

(17)

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

$$\Delta/(60 \tau) 2^{((\Delta/\tau)^\mu)/(28 \delta \log(2d))} > (\Delta/\tau) + 2 \Delta$$
\[
(\Delta/\tau)^\mu > 28 \delta \log(2d) \log(60+120 \tau) > 168 \delta \log(2d) \log(2\Delta) \Rightarrow \tau < \Delta/4
\]
since \(\tau < \Delta/4\)

Let \(p = p_1, p_2, \ldots, p_{\tau+1} = q\) be an ordered sequence of \(\tau+1\) nodes such that \(\text{dist}_G(p_1, p_{i+1}) = \Delta/\tau\) for \(i = 1, 2, \ldots, \tau\). Applying Theorem 6 for each pair \((p_i, p_{i+1})\), we get a nested family \(\emptyset \subset F_i \subset 2^V\) of subsets of nodes such that \(t_i = |F_i| \geq \max \left\{1/(\Delta/\tau)^{\mu}, 1\right\}\) and, for any \(V_{i,k} \in F_i\), \(h_G(V_{i,k}) \leq \max \left\{(1/(\Delta/\tau))^{1-\mu}, (360 \log n)/(2(\Delta/\tau)^{\mu}(7 \delta \log(2d)))\right\}\). Recall that the subset of nodes \(V_{i,k}\) was constructed in Theorem 6 in the following manner (see Fig. 3 for an illustration):

- Let \(\ell_i\) and \(r_i\) be two nodes on a shortest path \(p_i, p_{i+1}\) such that \(\text{dist}_G(p_i, \ell_i) = \text{dist}_G(\ell_i, r_i) = \text{dist}_G(r_i, p_{i+1}) = \text{dist}_G(p_i, p_{i+1})/3\).

- For some \(1/(28(\Delta/\tau)^{1-\mu} \log(2d)) \leq (\Delta/\tau)^{1-\mu} \log(2d)) < 1/4\), construct the graph \(G_{C_{i,k}}\) obtained by removing the set of nodes \(C_{i,k}\) which are exactly at a distance of \(\left[\alpha_{i,k} \text{dist}_G(p_i, p_{i+1})\right]\) from some node of the shortest path \(\ell_i, r_i\).

- The subset \(V_{i,k}\) is then the ball \(B_{G_{c_{i,k}}}(y_i, a_{i,k})\) for some \(a_{i,k} \in [0, \text{dist}_G_{c_{i,k}}(p_i, p_{i+1})]/2\) and for some \(y_i \in \{p_i, p_{i+1}\}\). If \(y_i = p_i\) then we call the collection of subsets \(F_i\) “left handed”, otherwise we call \(F_i\) “right handed”.

We can partition the set of \(\tau\) collections \(F_1, \ldots, F_\tau\) into four groups depending on whether the subscript \(j\) of \(F_j\) is odd or even, and whether \(F_j\) 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 \(F_{i,k}\) when \(i\) is even and \(F_{i,k}\) 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.

**Figure 3:** Illustration of various quantities related to the proof of Theorem 8. Nodes within the lightly cross-hatched region belong to \(C_{i,k}\) and \(C_{j,k'}\). Note that \(B_{G-c_{i,k}}(p_i, a_{i,k})\) and \(B_{G-c_{j,k'}}(p_j, a_{j,k'})\) need not be balls in the original graph \(G\).

Consider an arbitrary set in the above-mentioned collection of the form \(V_{i,k} = B_{G-c_{i,k}}(p_i, a_{i,k})\) with even \(i\). Let \(C_{i,k}\) denote the nodes in the interior of the closed cylinder of nodes in \(G\) which are at a distance of at most \(\left[\alpha_{i,k} \text{dist}_G(p_i, p_{i+1})\right]\) from some node of the shortest path \(\ell_i, r_i\), i.e., let \(C_{i,k} = \{u \mid \exists v \in \ell_i, r_i: \text{dist}_G(u, v) \leq \left[\alpha_{i,k} \text{dist}_G(p_i, p_{i+1})\right]\}\) (see Fig. 3). Let \(V_{j,k'} = B_{G-c_{j,k'}}(p_j, a_{j,k'})\) be a set in another family \(F_j\) with even \(j \neq i\) (see Fig. 3). Assume, without loss of generality, that \(i\) is smaller than \(j\), i.e., \(i \leq j - 2\) (the other case is similar).

**Proposition 1** \(C_{i,k} \cap B_{G-c_{j,k'}}(p_j, \Delta/(2\tau)) = \emptyset\).
Proof. Assume for the sake of contradiction that \( u \in C_{i,k} \cap B_{G_{c_{j,k'}}}(p_j, \Delta/(2 \tau)) \neq \emptyset \). Since \( u \in C_{i,k} \), there exists \( v \in \ell_{i,r_i} \) such that \( \text{dist}_G(v, u) \leq \lfloor \alpha_{i,k} \text{dist}_G(p_i, p_{i+1}) \rfloor < \text{dist}_G(p_i, p_{i+1})/4 = \Delta/(4 \tau) \). Thus,

\[
\begin{align*}
u \in B_{G_{c_{j,k'}}}(p_j, \Delta/(2 \tau)) & \Rightarrow \text{dist}_{G_{c_{j,k'}}}(u, p_j) \leq \Delta/(2 \tau) \Rightarrow \text{dist}_{G}(u, p_j) \leq \Delta/(2 \tau) \\
& \Rightarrow \text{dist}_{G}(v, p_j) \leq \text{dist}_{G}(v, u) + \text{dist}_{G}(u, p_j) < \Delta/(4 \tau) + \Delta/(2 \tau) < \Delta/\tau
\end{align*}
\]

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

Proposition 2 \( \text{dist}_{G_{c_{j,k'}}}(u, p_j) > \Delta/(2 \tau) \) for any node \( u \in V_{i,k} \cap V_{j,k'} = B_{G_{c_{i,k'}}}(p_i, a_{i,k}) \cap B_{G_{c_{j,k'}}}(p_j, a_{j,k'}) \).

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

\[
\text{dist}_{G_{c_{i,k}}}(p_{i,k}, u) \leq a_{i,k} \leq \text{dist}_{G_{c_{i,k}}}(p_i, p_{i+1})/2
\]

Since \( u \in V_{j,k'} = B_{G_{c_{j,k'}}}(p_j, a_{j,k'}) \), this implies \( u \in B_{G_{c_{j,k'}}}(p_j, z) \). Since \( z \leq \Delta/(2 \tau) \), by Proposition 1 \( C_{i,k} \cap B_{G_{c_{j,k'}}}(p_j, z) = \emptyset \), and therefore

\[
\Delta/(2 \tau) \geq z = \text{dist}_{G_{c_{j,k'}}}(u, p_j) = \text{dist}_{G_{c_{i,k}} \cup c_{j,k'}}(u, p_j) \geq \text{dist}_{G_{c_{i,k}}}(u, p_j)
\]

which in turn implies

\[
\text{dist}_{G_{c_{i,k}}}(p_{i,k}, p_j) \leq \text{dist}_{G_{c_{i,k}}}(p_i, u) + \text{dist}_{G_{c_{i,k}}}(u, p_j) \leq \text{dist}_{G_{c_{i,k}}}(p_i, p_{i+1})/2 + \Delta/(2 \tau) \tag{18}
\]

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

\[
\text{dist}_{G_{c_{i,k}}}(p_{i,k}, p_{i+1}) \leq \text{dist}_{G_{c_{i,k}}}(p_i, p_j) + \text{dist}_{G_{c_{i,k}}}(p_j, p_{i+1}) \leq \text{dist}_{G_{c_{i,k}}}(p_i, p_{i+1})/2 + \Delta/(2 \tau) + \Delta
\]

by (18)

\[
\Rightarrow \text{dist}_{G_{c_{i,k}}}(p_{i,k}, p_{i+1}) \leq \Delta/\tau + 2 \Delta \tag{19}
\]

On the other hand, by Fact 1:

\[
\text{dist}_{G_{c_i}}(p_i, p_{i+1}) \geq \Delta/(60 \tau) 2^{(\alpha_{i,k} \Delta)/(\delta \tau)} \geq \Delta/(60 \tau) 2^{((\Delta/\tau) \mu)/(28 \delta \log(2d))} \tag{20}
\]

Inequalities (19) and (20) together imply

\[
\Delta/(60 \tau) 2^{((\Delta/\tau) \mu)/(28 \delta \log(2d))} \leq (\Delta/\tau) + 2 \Delta \tag{21}
\]

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 \( V_{j,k'} \cap B_{G_{c_{j,k'}}}(p_j, \Delta/(2 \tau)), u \notin B_{G_{c_{j,k'}}}(p_j, \Delta/(2 \tau)) \), and thus there are at least \( \Delta/(2 \tau) \) node on a shortest path in \( G_{c_{j,k'}} \) 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”) \( \text{cut}_G(S, s, t) \) 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.,

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

Lemma 9 Suppose that the following holds for our given \((G, d, \delta)\):

- \( s \) and \( t \) are two nodes of \( G \) such that \( \text{dist}_G(s, t) > 48\delta + 8\delta \log n \), 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 \((\text{dist}_G(s, t) - 8\delta \log n)/(50\delta) = \Omega(\text{dist}_G(s, t)) \) (node and edge) disjoint cuts such that each such cut has at most \( d^{128^{1+1}} \) cut edges.

Remark 3 Suppose that \( G \) is hyperbolic (i.e., \( \delta \) is a constant), \( d \) is a constant, and \( s \) and \( t \) be two nodes such that \( \text{dist}_G(s, t) > 48\delta + 8\delta \log n = \Omega(\log n) \). Lemma 9 then implies that there are \( \Omega(\text{dist}_G(s, t)) \) s-t cuts each having \( O(1) \) edges. If, on the other hand, \( \delta = O(\log \log n) \), then such cuts have \( \text{polylog}(n) \) edges.

Remark 4 The bound in Lemma 9 is obviously meaningful only if \( \delta = o(\log n) \). If \( \delta = \Omega(\log n) \), then \( \delta \)-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 \( L_i \) be the sets of nodes at the \( i \)th level (i.e., \( \forall u \in L_i : \text{dist}_G(s, u) = i \)); obviously \( t \in L_{\text{dist}_G(s, t)} \). Assume \( \text{dist}_G(s, t) > 48\delta + 8\delta \log n \), and consider two arbitrary paths \( \mathcal{P}_1 \) and \( \mathcal{P}_2 \) between \( s \) and \( t \) passing through two nodes \( v_1, v_2 \in L_j \) for some \( 48\delta \leq j \leq \text{dist}_G(s, t) - 7\delta \log n \).

We first claim that \( \text{dist}_G(v_1, v_2) < 12\delta \). Suppose, for the sake of contradiction, suppose that \( \text{dist}_G(v_1, v_2) \geq 12\delta \). Let \( v'_1 \) and \( v'_2 \) be the first node in level \( L_{j+6\delta \log n} \) 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 \( j + 6\delta \log n < \text{dist}_G(s, t) \) implies \( L_{j+6\delta \log n+1} \neq \emptyset \), there must be a path \( \mathcal{P}_3 \) between \( v'_1 \) and \( v'_2 \) through \( t \) using nodes not in \( \bigcup_{0 \leq \ell \leq j+6\delta \log n} L_{\ell} \). 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 \( v \in L_j \) lying on a path between \( s \) and \( t \), \( B_G(v, 12\delta) \) provides an s-t cut \( \text{cut}_G(B_G(v, 12\delta), s, t) \) having at most \( \mathcal{E}_G(B_G(v, 12\delta), s, t) \leq d^{128^{1+1}} \) edges. To see this, consider any path \( \mathcal{P} \) between \( s \) and \( t \) and let \( u \) be the first node in \( L_j \) visited by the path. Then, \( \text{dist}_G(u, v) \leq 12\delta \) and thus \( u \in B_G(v, 12\delta) \). Since nodes in \( B_G(v, 12\delta) \) are at a distance of at least \( 35\delta \) from \( s \) and \( t \notin B_G(v, 12\delta) \), \( d \) is the maximum degree of any node in \( B_G(v, 12\delta) \) and it follows that \( \mathcal{E}_G(B_G(v, 12\delta), s, t) \leq d\partial_G(B_G(v, 12\delta - 1)) \leq d^{128^{1+1}} \).

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


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, 35])

The input to this problem a graph $G = (V, E)$, two nodes $s, t \in V$, and two positive integers $0 < r < \kappa$. Call an edge “shared” if it is in more than $r$ paths between $s$ and $t$. The goal is to find a set of $k$ paths between $s$ and $t$ that minimizes the number of shared edges.

When $r = 1$, the Uumv problem is called the “minimum shared edges” (Mse) problem.

We will use the notation $\text{OPT}_{\text{Uumv}}(G, s, t, r, \kappa)$ 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 [33–35] 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 $\text{NP} \subseteq \text{DTIME}(n^{\log \log n})$.
- Uumv admits a $\lfloor k/(r + 1) \rfloor$-approximation.
- Mse admits a $\min \{ \lfloor k/2 \rfloor, n^{3/4}, m^{1/2} \}$-approximation.

4.1.1 Greedy Fails for Uumv or Mse Even for Hyperbolic Graphs (i.e., constant $\delta$)

Several routing problems have been looked at for hyperbolic graphs (i.e., constant $\delta$) 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 $\kappa$ 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. 4; 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.

4.1.2 Improved Approximations for Uumv or Mse for $\delta$ 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).
Lemma 10 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, \( \text{UUMV} \) (and, consequently also \( \text{MSE} \)) for a \( \delta \)-hyperorbolic graph \( G \) can be approximated within a factor of \( O \left( \max \{ \log n, d^{O(\delta)} \} \right) \). This improves upon the currently best \( O \left( n^{1/2} \right) \)-approximation for arbitrary graphs provided \( \delta = o \left( \log n / \log d \right) \).

Remark 5 Thus for fixed \( d \) Lemma 10 provides improved approximation as long as \( \delta = o(\log n) \).

Note that our approximation ratio is independent of the value of \( \kappa \). Also note that \( \delta = \Omega(n) \) allows expander graphs as a sub-class of \( \delta \)-hyperorbolic graphs for which \( \text{UUMV} \) is expected to be harder to approximate.

Proof of Lemma 10

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 \( \text{UUMV} \) (and thus \( \text{MSE} \)) has the same approximability properties as EHSSC by characterizing optimal solutions of \( \text{UUMV} \) 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 \( \tilde{E} \subseteq E \) such that \( |\tilde{E}| \) is minimum and

\[
\forall s \in S \subseteq V \setminus \{t\}: \quad |E_G(S, s, t)| \leq k \Rightarrow E_G(S, s, t) \cap \tilde{E} \neq \emptyset
\]

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

Lemma 11 (Relating EHSSC to UUMV) \( \text{OPT}_{\text{UUMV}}(G, s, t, r, \kappa) = \text{OPT}_{\text{EHSSC}}(G, s, t, \lceil \kappa/r \rceil - 1) \).
Note that any feasible solution for $Uumv$ must contain at least one edge from every collection of cut-edges $E_G(S, s, t)$ satisfying $|E_G(S, s, t)| \leq \lceil \kappa/r \rceil - 1$, since otherwise the number of paths going from $E_G(S, s, t)$ to $V \setminus E_G(S, s, t)$ is at most $r \times (\lceil \kappa/r \rceil - 1) < \kappa$. Thus it follows that $\text{OPT}_{Uumv}(G, s, t, r, \kappa) \geq \text{OPT}_{Ehssc}(G, s, t, \lceil \kappa/r \rceil - 1)$.

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

$$c(e) = \begin{cases} \infty, & \text{if } e \in E_{Ehssc}(G, s, t, [\kappa/r] - 1) \\ r, & \text{otherwise} \end{cases}$$

The value of the minimum $s$-$t$ cut for $G$ is then at least min $\{\infty, r \times [\kappa/r]\} \geq \kappa$ which implies (by the max-flow-min-cut theorem) the existence of $\kappa$ flows each of unit value. The paths taken by these $\kappa$ flows provide our desired $\kappa$ paths for $Uumv$. \qed

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:

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

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 9 to derive our approximation algorithm.

**Lemma 12 (Approximation algorithm for $Ehssc$)** $Ehssc$ admits a $O\left(\max\{\delta \log n, d^{O(\delta)}\}\right)$-approximation.

**Proof.** Our algorithm for $Ehssc$ can be summarized as follows:

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

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

**Case 1:** $k \leq d^{12\delta+1}$. Let $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_\ell$ 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 $j \neq j'$, $\text{OPT}_{Ehssc}(G, s, t, k) \geq \ell$, 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 $\text{OPT}_{Ehssc}(G, s, t, k) \geq 1$, 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 9 to find a collection \( S_1, S_2, \ldots, S_\ell \) of \( \ell = \ldots \) cases, and thus additional arguments may be needed to establish similar hardness results these classes of graphs.

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

### 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 \( \text{vol}_G(S) \) 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 \( \Phi_G(S) = \text{cut}_G(S)/\text{vol}_G(S) \). 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 13 (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 \( \zeta n \) nodes \( S \), for some constant \( 0 < \zeta < 1/2 \), such that \( \Phi_G(S) \leq \varepsilon \) for some constant \( 0 < \varepsilon \leq 1 \). Then, find as efficiently as possible a subset \( S' \) of at most \( \zeta n \) nodes such that \( \Phi_G(S') \leq \eta \varepsilon \) 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., \( O(2^{c n}) \) 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 \( \delta \)-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 13) can be solved in polynomial time for such a proper subclass of non-expanders provided the hyperbolicity measure \( \delta \) is a sufficiently slowly growing function of \( n \).

**Lemma 14 (polynomial time solution of SSE for \( \delta \)-hyperbolic graphs when \( \delta \) is sub-logarithmic and \( d \) is sub-linear)** Suppose that \( G \) is a \( d \)-regular \( \delta \)-hyperbolic graph. Then the SSE problem for \( G \) can be solved in polynomial time provided \( d \) and \( \delta \) satisfy:

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

**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 \( C \sqrt{\log d} \) for some constant \( C > 0 \) [24]. Since we show that SSE is polynomial-time solvable for \( \delta \)-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 these classes of graphs.
Proof. Our proof is quite similar to that used for Theorems 6. 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 and proof of Theorem 6 in this proof, so we urge the readers to familiarize themselves with the details of that proof before reading the current proof. Note that $h_G(S) \leq \varepsilon$ implies $\Phi_G(S) \leq d h_G(S)/d \leq \varepsilon$. We select the nodes $p$ and $q$ such that $\Delta = \text{dist}_G(p, q) = \log_q n = \log n/\log d$. Set the constant $\mu$ to be $1/2$. Note that $(360 \log n)/\left(\Delta^2/(28 \delta \log(2d))\right) < (1/\Delta)^{1-\mu}$ since

$$
(360 \log n)/\left(\Delta^2/(28 \delta \log(2d))\right) < (1/\Delta)^{1-\mu}
$$

$$
\leq (360 \log d)/\left(2^{(\log n)^{1/2}/((56 \delta (\log d)^{3/2}))}\right) < (\log d/\log n)^{1/2}
$$

$$
\leq 9 + \log \log n/2 < \left((\log n)^{1/2}/(56 \log(1-\rho)/2 n)\right) - \log(1-\rho)/2 n
$$

and the last inequality clearly holds for sufficiently large $n$.

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

$$
|\mathcal{B}_G(p, r)| \leq |\mathcal{B}_G(p, (\Delta/3) - \alpha \Delta)| < |\mathcal{B}_G(p, \Delta/3)| < \sum_{i=0}^{\Delta/3} d^i < d^{(\Delta/3)+1} = d n^{1/3} < \zeta n
$$

(22)

where the last inequality follows since $d \leq 2^{\log(1/3) - r} n$.

Otherwise, no such $r$ exists, and this implies

$$
|\mathcal{B}_G(p, (\Delta/3) - \alpha \Delta)| \geq (1 + \varepsilon)^{(\Delta/3) - \alpha \Delta} > (1 + \varepsilon)^{\Delta/4} \geq e^{\varepsilon \Delta/8} = e^{\varepsilon \log d n/8} = n^\varepsilon \log d e/8
$$

Now there are two major cases as follows.

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

We know that $\text{dist}_{G_{-c}}(p, q) \geq (\Delta/60)^2 \Delta^2$ and (by choice of $p$) $|\mathcal{B}_{G_{-c}}(p, \text{dist}_{G_{-c}}(p, q)/2)| < n/2$. 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 $t = \text{dist}_{G_{-c}}(p, q)$. Perform a BFS starting from $p$ in $G_{-c}$, and let $\mathcal{L}_i$ be the sets of nodes at the $i$th level (i.e., $\forall u \in \mathcal{L}_j: \text{dist}_{G_{-c}}(p, u) = i$). Note that $\left|\bigcup_{j=0}^{t/2} \mathcal{L}_j\right| \leq n^{\varepsilon}$. 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 $\Delta_0, \Delta_1, \ldots, \Delta_{(t/2)}(\kappa-1)$ each of length $\kappa = (8/\varepsilon) \ln n$, i.e.,

$$
\begin{align*}
0, & \ldots, \kappa - 1, \\
\Delta_0 & , \kappa + 1, \kappa + 2, \ldots, 2\kappa - 1, \\
\Delta_1 & , \ldots, (t/2) - \kappa + 1, (t/2) - \kappa + 2, \ldots, (t/2) \\
\Delta_{(t/2)} & , \kappa - 1
\end{align*}
$$

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

$$
\forall i\kappa \leq j \leq i\kappa + \kappa - 1: h_{G_{-c}}(\mathcal{L}_j) \geq h_G(\mathcal{L}_j)/2 > \varepsilon/2
$$

$$
|\mathcal{L}_{i\kappa + \kappa - 1}| > |\mathcal{L}_{i\kappa}| \left(1 + (\varepsilon/2)\right)^\kappa \geq (1 + (\varepsilon/2))^{(8/\varepsilon) \ln n} \geq e^{(\varepsilon/4)((8/\varepsilon) \ln n)} = n^2 > n
$$
which contradicts the fact that \( \left| \bigcup_{j=0}^{t/2} L_j \right| \leq n/2 \). Since \( \sum_{i=0}^{(1+(t/2))/\kappa-1} \left| L_{i^*} \right| < n/2 \), there exists a set \( L_{k^*} \) such that\( h_{G^*}(L_{k^*}) \leq \varepsilon \) and \( |L_{k^*}| < n/2 \).

\[
\frac{n/2}{(1+(t/2))/\kappa} < n\kappa/t < (8 \ln n)/\left( \varepsilon (\Delta/60) 2^{\Delta^1/2} (7 \delta \log(2d)) \right) \\
\leq \left( 480 n \log^{(1/3)-(\rho)} n \right) / \left( \varepsilon 2 \log^{\rho/2} n/14 \right) < \zeta n
\]

**Case 2:** There is no path between \( p \) and \( q \) in \( G_{-C} \).

In this case, we return \( B_{G_{-C}}(p, (\Delta/3) - \alpha \Delta) = B_G(p, (\Delta/3) - \alpha \Delta) \) as our solution. The size requirement follows since we showed in (22) that \( |B_G(p, (\Delta/3) - \alpha \Delta)| < \zeta n \). Note that nodes in \( B_G(p, (\Delta/3) - \alpha \Delta) \) can only be connected to nodes in \( C \), and thus

\[
h_{G}(B_G(p, (\Delta/3) - \alpha \Delta)) \leq |C| / |B_G(p, (\Delta/3) - \alpha \Delta)| \leq \left( (\Delta/3)d^{\alpha \Delta} \right) / \left( n^{(\rho \log_d e)/8} \right) \\
< n^{\alpha-(\varepsilon \log_d e)/8} \log n < n^{1/7 \Delta^1/2 \log(2d)} - (\varepsilon/(8 \ln d)) \log n < \varepsilon
\]

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

In all cases, the required subset of nodes can be found in \( O(n^2 \log n) \) time. \[\square\]

**Acknowledgements** The problem of investigating expansion properties of \( \delta \)-hyperbolic graphs was raised originally to some of the authors by A. Wigderson.

**References**


