Improved Online Reachability Preservers*

Greg Bodwin and Tuong Le
University of Michigan
{bodwin, tuongle}@umich.edu

Abstract

A *reachability preserver* is a basic kind of graph sparsifier, which preserves the reachability relation of an n-node directed input graph G among a set of given demand pairs P of size $|P| = p$. We give constructions of sparse reachability preservers in the online setting, where G is given on input, the demand pairs $(s,t) \in P$ arrive one at a time, and we must irrevocably add edges to a preserver H to ensure reachability for the pair (s,t) before we can see the next demand pair. Our main results are:

- There is a construction that guarantees a maximum preserver size of

$$|E(H)| \leq O\left(n^{0.72}p^{0.56} + n^{0.6}p^{0.7} + n\right).$$

  This improves polynomially on the previous online upper bound of $O(\min\{np^{0.5}, n^{0.5}p\}) + n$, implicit in the work of Coppersmith and Elkin [SODA '05].

- Given a promise that the demand pairs will satisfy $P \subseteq S \times V$ for some vertex set S of size $|S| =: \sigma$, there is a construction that guarantees a maximum preserver size of

$$|E(H)| \leq O\left((np\sigma)^{1/2} + n\right).$$

  A slightly different construction gives the same result for the setting $P \subseteq V \times S$. This improves polynomially on the previous online upper bound of $O(\sigma n)$ (folklore).

All of these constructions are polynomial time, deterministic, and they do not require knowledge of the values of p, σ, or S. Our techniques also give a small polynomial improvement in the current upper bounds for *offline* reachability preservers, and our results extend to an even stronger model in which we must commit to a path for all possible reachable pairs in G before any demand pairs have been received. As an application, we improve the competitive ratio for Online Unweighted Directed Steiner Forest to $O(n^{3/5+\varepsilon})$, improving on the previous bound of $O(n^{2/3+\varepsilon})$ [Grigorescu, Lin, Quanrud APPROX-RANDOM '21].

*This work was supported by NSF:AF 2153680.

1 Introduction

We study *reachability preservers*, a basic graph sparsifier that has found applications in graph spanners and shortcut sets [25], property testing algorithms [2], flow/cut approximation algorithms [18], Steiner network design algorithms [1], etc (see [8] for more discussion and applications).

Definition 1 (Reachability Preservers)**.** Given a directed graph $G = (V, E)$ and a set of demand pairs $P \subseteq V \times V$, a *reachability preserver* is a subgraph $H \subseteq G$ with the property that, for all $(s, t) \in P$, there is an $s \rightsquigarrow t$ path in H iff there is one in G.

The study of reachability preservers goes back at least to Directed Steiner Network (DSN), a classic NP-hard graph algorithm, which can be phrased as the computational task of computing the sparsest (or minimum weight) reachability preserver of a given instance G, P [26]. More recently they have been intensively studied from an extremal perspective, where the goal is to determine the worst-case number of edges needed for a reachability preserver, typically as a function of n (the number of nodes in the input graph) and $p := |P|$ (the number of demand pairs) [1,5,8,12,13,17,24].

Almost all previous work on reachability preservers operates in the **offline** model, where G, P are given on input and the goal is to construct a sparse preserver H. However, there is also a long line of algorithmic work on **online** Steiner network design [3, 4, 9–11, 21, 23, 24]. Here the model is that G is given on input, and the demand pairs $(s, t) \in P$ arrive one at a time. We must irrevocably add edges to H to preserve reachability for the current demand pair before the next one arrives, and the process can halt at any time without warning. These papers typically try to achieve a small *competitive ratio*, that is, the goal is design an efficient online algorithm that achieves an upper bound on preserver size of the form $|E(H)| \leq OPT \cdot f(n, p)$ where OPT is the best *offline* solution for the given instance G, P and the function f is as small as possible.

We study extremal bounds for reachability preservers in the online model. This problem has been previously explicitly considered only in a stronger non-standard online model (see Theorem 3), or implicitly as an internal ingredient inside the aforementioned algorithms (see Section 1.2). There is one previous paper that implies results in this setting, which is by Coppersmith and Elkin [19], and it more strongly studies *distance* preservers (which must preserve exact distances among demand pairs). Although not explicitly discussed, one can use a folklore reduction of reachability preservers into the setting of DAGs (see Theorem 9), and then apply the analysis of Coppersmith and Elkin to show an online upper bound of

$$|E(H)| \leq O\left(\min\left\{np^{1/2}, n^{1/2}p\right\} + n\right).$$

Meanwhile, in the *source-restricted* setting where the demand pairs satisfy $P \subseteq S \times V$ (or $P \subseteq V \times S$) for some set of source nodes S, we have the online bound

$$|E(H)| \leq O\left(n|S|\right).$$

This construction is folklore; one simply selects paths for demand pairs in a "consistent" fashion [7, 19], meaning that they will all lie within a set of in- and out-trees rooted at the nodes in S, which have $O(n|S|)$ edges in total. This simple construction remains state of the art, and has been used recently as an ingredient in online Steiner network algorithms [21,24], thus motivating further investigation. We note that neither of these constructions require any advance knowledge of P or S (not even their size); this is typically considered to be an essential feature of the online model.

Although there are many more previously-known extremal bounds for offline reachability preservers (see Table 1), they all rely on one or more technical tools that inherently require advance

knowledge of the demand pairs; we discuss these tools in more detail in Section 3.3. We develop a framework in which to analyze online reachability preservers, and we use it to prove two new upper bounds, improving on both of the results mentioned above:

Theorem 1 (Online Pairwise Reachability Preservers)**.** Given an n-node directed graph G, there is an online algorithm that constructs a reachability preserver H of $|P| = p$ total demand pairs of size at most

$$|E(H)| \leq O\left(n^{\frac{2+\alpha}{3+\alpha}+o(1)}p^{\frac{2}{3+\alpha}} + n^{2-2\alpha+o(1)}p^{\alpha} + n\right) < O\left(n^{0.72}p^{0.56} + n^{0.6}p^{0.7} + n\right),$$

where $\alpha \geq 0.7$ is a root of $4x^3 - 13x^2 + 10x - 2$.[1]

Theorem 2 (Online Source-Restricted Reachability Preservers)**.** Given an n-node directed graph $G = (V, E)$, and a promise that the demand pairs will satisfy $P \subseteq S \times V$ for some set of source nodes $S \subseteq V$, there is an online algorithm that constructs a reachability preserver H of p total demand pairs of size at most

$$|E(H)| \leq O\left((n|S|p)^{1/2} + n\right).$$

The same result holds under the promise that $P \subseteq V \times S$, but requires a slightly different algorithm.

All of our construction algorithms are deterministic, run in polynomial time, and do not require any knowledge of P or S (including their size). Theorem 2 ties the state-of-the-art upper bound in the *offline* sourcewise setting [1].

It may also be interesting to compare Theorem 1 to the following lower bound, proved in [8] in a stronger online model:

Theorem 3 ([8])**.** Consider a stronger version of the online model where the graph G is initially empty, and an adversary may add new edges to G in each round before providing the next demand pair. Then there is a strategy for the adversary that guarantees that any online reachability preserver H will have size

$$|E(H)| \geq \Omega\left((np)^{2/3} + n\right).$$

Our bound in Theorem 1 is polynomially better than this one (in exchange for a weaker adversary), and so it formally separates these two models.

1.1 Other Models

A couple of the new tools that we develop for the online setting also turn out to be helpful in the classical offline setting. This yields the following small polynomial improvement in the state of the art:

Theorem 4 (Offline Reachability Preservers)**.** Given an n-node directed graph G and a set of demand pairs P, one can construct in polynomial time a reachability preserver H of size

$$|E(H)| \leq O\left(n^{3/4}p^{1/2} + n^{2-\sqrt{2}+o(1)}p^{1/\sqrt{2}} + n\right) \leq O\left(n^{0.75}p^{0.5} + n^{0.59}p^{0.71} + n\right).$$

[1] This upper bound on $|E(H)|$ is decreasing as α increases, so one gets a correct but slightly suboptimal upper bound by plugging in $\alpha = 0.7$ (the explicit form on the right is a slight overestimate of that bound, with the exponents rounded off). The exact value is $\alpha = 0.70086\ldots$.

Bound on $\lvert E(H)\rvert$		**Offline**	**Online**	**Citation**
$O\big(\min\{np^{1/2}, n^{1/2}p\}$	$+n\big)$	✓	✓	[19] (implicit)
$O\big(n^{2/3}p^{2/3}$	$+n\big)$	✓		[1]
$O\big(\frac{n^2}{2^{\log^* n}}$	$+p\big)$	✓		[1]
$O\big(n^{3/4}p^{1/2} + n^{5/8}p^{11/16}$	$+n\big)$	✓		[8]
$O\big(\frac{p^2}{2^{\log^* p}}$	$+n\big)$	✓		[8]
$O\big(n^{3/4}p^{1/2} + n^{2-\sqrt{2}+o(1)}p^{1/\sqrt{2}}$	$+n\big)$	✓		this paper
$O\big(n^{0.73}p^{0.54} + n^{0.6}p^{0.7}$	$+n\big)$	✓	✓	this paper
$O\big((np\lvert S\rvert)^{1/2}$	$+n\big)$	✓		[1]
$O\big((np\lvert S\rvert)^{1/2}$	$+n\big)$	✓	✓	this paper
For any int $d \geq 1$: $\Omega\big(n^{\frac{2}{d+1}}p^{\frac{d-1}{d}}$	$+n+p\big)$	✓	✓	[1], based on [19]
For $p \geq n^{4/9}\lvert S\rvert^{2/3}$: $\Omega\big(n^{4/5}p^{1/5}\lvert S\rvert^{1/5}$	$+n+p\big)$	✓	✓	[1]

Table 1: The progression of upper and lower bounds on the extremal number of edges needed for an offline/online reachability preserver. When S is present, the bound applies in the source-restricted setting $P \subseteq S \times V$ (or $P \subseteq V \times S$). The terms $2^{\log^* n}$ and $2^{\log^* p}$ reflect the current lower bounds on the Ruzsa-Szemerédi function; see [8] for discussion.

We will also consider the stronger *non-adaptive* version of the online model. In its strongest form, this model would require that the selected path $\pi(s,t)$ added to the preserver for each demand pair (s,t) depends only on the input graph G, and not also on the previous demand pairs or the current preserver H. Equivalently, after receiving the input graph G, we are required to commit to a choice of path for *every* reachable pair (s,t) before *any* demand pairs are received. In addition to the interpretation through the online model, this can be viewed as a parallelization of reachability preservers: it allows one to preprocess a graph G in such a way that, given any set of demand pairs P, we can construct a sparse reachability preserver H by adding a path for all demand pairs $(s,t) \in P$ *in parallel*, without the path-adding process for (s,t) even knowing the other demand pairs in P.

We will show that our upper bounds can *almost* be made non-adaptive, but a little bit of extra information is required. Making them fully non-adaptive is an interesting open problem.

Theorem 5 (Non-Adaptive Reachability Preservers)**.** There are online algorithms satisfying Theorem 1 and Theorem 2 where the selected path for each demand pair (s,t) depends only on the input graph G, the set of source nodes S (for Theorem 2), and

- the index i of the current demand pair, **or**
- the total number p of demand pairs that will arrive.

The first (i-sensitive) part of this theorem strictly strengthens Theorem 1. The second (p-sensitive) part is incomparable in strength to Theorem 1, since it depends on the overall number of demand pairs p, which is unknown in Theorem 1. Both parts are incomparable in strength to Theorem 2, since they require knowledge of the source nodes S which are unknown in Theorem 2.

These are proved in Section 4. For comparison, the results of Coppersmith and Elkin for distance preservers [19] imply non-adaptive reachability preservers of quality $O(\min\{np^{1/2}, n^{1/2}p\} + n)$, but nothing further was previously known.

1.2 Application to Unweighted Directed Steiner Network

The algorithmic problem of computing the sparsest[2] possible reachability preserver of an input G, P is called *Unweighted Directed Steiner Network (UDSN)*. This problem is NP-hard, and even hard to approximate [20], but it is well studied from the standpoint of approximation algorithms. After considerable research effort [1, 6, 14–16, 22], the state-of-the-art is:[3]

Theorem 6 (Offline UDSN [1, 15])**.** There is a randomized polynomial time algorithm for (offline) UDSN that, given an n-node input graph G and a set of p demand pairs P, returns a reachability preserver H of size

$$|E(H)| \leq OPT \cdot O\left(\min\left\{n^{4/7+\varepsilon}, p^{1/2+\varepsilon}\right\}\right)$$

where OPT is the number of edges in the sparsest possible reachability preserver of G, P.

The bound of $O(n^{4/7+\varepsilon})$ follows a proof framework that was originally developed by Chlamtáč, Dinitz, Kortsarz, and Laekhanukit [16], but then replaces a certain key internal ingredient in their proof with an offline source-restricted reachability preserver [1].

It is an interesting open question whether the bound in Theorem 6 can be matched by an online algorithm (in other words, is the *competitive ratio* for UDSN bounded by $O\left(\min\left\{n^{4/7+\varepsilon}, p^{1/2+\varepsilon}\right\}\right)$?). This was partially achieved about ten years ago by Chakrabarty, Ene, Krishnaswamy, and Panigrahi [21], who showed:

Theorem 7 (Online UDSN, parametrized on p [21])**.** There is a randomized polynomial time algorithm for online UDSN that achieves a competitive ratio of $O\left(p^{1/2+\varepsilon}\right)$.

Thus, the remaining question is whether the dependence on n in Theorem 6 can be recovered in the online setting. Here the state of the art is due to, Grigorescu, Lin, and Quanrud, who obtained a competitive ratio of $n^{2/3+\varepsilon}$ [24], roughly following the framework of Chlamtáč et al. [16] but with adaptations for the online setting. As an application of our new online source-restricted preservers, we are able to substitute them into this framework, improving the competitive ratio:

Theorem 8 (Online UDSN, parametrized on n)**.** There is a randomized polynomial time algorithm for online UDSN that achieves a competitive ratio of $O(n^{3/5+\varepsilon})$.

1.3 Organization

- Section 2 recaps some useful technical preliminaries, largely from [8].
- Section 3.1 sets up the algorithms used to select paths for demand pairs as they arrive in the online setting, and proves their basic properties.
- Section 3.2 proves Theorem 2.

[2] More generally, in algorithmic contexts one can consider a weighted version of the problem, where G has edge weights and the goal is to construct a min-weight preserver H.

[3] All work to date has focused on proving an approximation ratio as a function of n *or* as a function of p. It is not clear if a better approximation ratio can be achieved if we consider bounds that can depend on both n and p.

- Section 3.3 informally overviews Theorems 1 and 4. The formal proofs are more technical, so we give them in Appendices A and B, respectively.
- Section 4 proves Theorem 5.
- Section 5 proves Theorem 8.

2 Technical Preliminaries

We will recap some definitions and results from prior work that will be useful in our arguments to follow.

2.1 The DAG reduction

While proving all of our upper bounds, it will be convenient to assume that the input graph is a DAG. That this assumption is without loss of generality comes from the following standard reduction, which we will recap somewhat briefly here:

Theorem 9 (DAG Reduction (folklore))**.** If there is an algorithm that constructs a reachability preserver H of size

$$|E(H)| \leq f(n, p, \sigma)$$

for any p demand pairs using $|S| = \sigma$ source (or sink) nodes in an n-node DAG, then there is an algorithm that constructs a reachability preserver H of size

$$|E(H)| \leq f(n, p, \sigma) + 2n$$

in arbitrary graphs.

Proof. Given an input graph G, compute a strongly connected component (SCC) decomposition. For each component C in the decomposition, choose an arbitrary vertex $v \in C$ and an in- and out-tree rooted at v spanning C. There are $2(|C| - 1)$ edges in these two trees, and hence there are slightly less than $2n$ edges in total, across all trees.

We add all edges in these trees to our reachability preserver H at their first opportunity, and then for the rest of the construction we treat each SCC as a single contracted supernode, yielding a DAG G'. For each demand pair (s, t), we can map the nodes s, t onto the corresponding supernodes in G' and choose paths in G' using our assumed DAG algorithm, to get a reachability preserver H' in G' on $\leq f(n, p, \sigma)$ edges. Each edge (u, v) added to H' can be mapped back to any single edge in G between the set of nodes corresponding to the supernode u and the set of nodes corresponding to the supernode v. Since our in- and out-trees preserve strong among all nodes in each of these sets, this will give a correct preserver in G. □

2.2 Path System Definitions

A *path system* is a pair $S = (V, \Pi)$, where V is a set of vertices and Π is a set of nonempty vertex sequences called *paths*. We will next recap some basic definitions; see also [8] for more discussion. Note that, even when the vertex set V is that of some graph G, the paths in Π are abstract sequences of vertices that do not necessarily correspond to paths in G. The *length* of a path $\pi \in \Pi$, written $|\pi|$, is its number of vertices (hence off by one from the length of the path through some graph). The *degree* of a node $v \in V$, written $\deg(v)$, is the number of paths that contain v (which

may differ significantly from its degree as a node in a graph). The *size* of S, written $\|S\|$, is the quantity

$$\|S\| := \sum_{\pi \in \Pi} |\pi|.$$

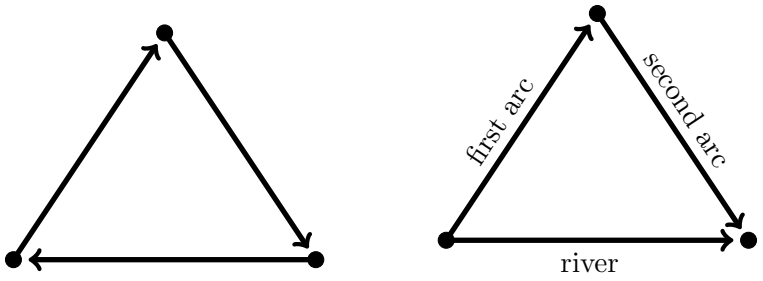


Figure 1: A directed 3-cycle (left) and a 3-bridge (right).

We will later use so-called *Turán-type methods* to bound the size of certain path systems, meaning that we will first establish that S avoids certain subsystems, and then we will bound the maximum possible size of *any* path system that avoids those subsystems. We say that S' is a *subsystem* of S, written $S' \subseteq S$, if it can be obtained by zero or more of the following operations: delete a path from Π, delete a node from V, or delete one instance of a node v from a path $\pi \in \Pi$. We will use two kinds of forbidden subsystems in this paper (see Figure 1):

- A *directed k-cycle* is a path system that has k nodes with a circular ordering $(x_0, x_1, \ldots, x_{k-1}, x_k = x_0)$, and k paths of length two each, which are all paths of the form $(x_i, x_{i+1}), 0 \leq i < k$.
- A *k-bridge* is a path system that has k nodes with a total ordering $(x_1, x_2, \ldots, x_k)$, and k paths of length two each, which are (1) the path (x_1, x_k), called the *river*, and (2) the $k-1$ paths of the form $(x_i, x_{i+1}), 1 \leq i < k$, called the i^{th} *arc*.

We note that k-bridges still count even when they are degenerate; for example, two paths that coincide on two consecutive nodes count as a 2-bridge, and three paths that coincide on three consecutive nodes contain a 3-bridge, etc.

A path system is *acyclic* if it does not contain any directed cycle as a subsystem, or equivalently, if there is a total ordering of the vertices V such that the order of vertices within each path $\pi \in \Pi$ agrees with this ordering. We will frequently consider *ordered path systems*, which are path systems with a total ordering on their path set Π. With this we will sometimes only forbid subsystems with certain ordering constraints, e.g., 3-bridges where the last arc comes before the river in the ordering of Π.

3 Online Reachability Preservers

3.1 Path Growth Algorithms

A key tool in our online upper bounds will be the following two (very similar) path selection algorithms. We use these algorithms to generate a path $\pi(s,t)$ for each demand pair (s,t) as it arrives, and then we add the edges of this path to the current preserver. Our path generation algorithms are greedy, growing paths one edge at a time and locally avoiding new edges if possible.

Input: DAG $G = (V, E)$, current preserver $H \subseteq G$, demand pair (s, t)

Let $\pi \leftarrow (s)$
while *last node of* π *is not* t **do**
  $u \leftarrow$ last node of π
  if *there exists an edge* $(u, v) \in E(H)$ *with* t *reachable from* v **then**
    append any such edge (u, v) to the back of π
  else
    append to the back of π any edge $(u, v) \in E(G)$ with t reachable from v
return π

Algorithm 1: `forwards-growth` path generation

Input: DAG $G = (V, E)$, current preserver $H \subseteq G$, demand pair (s, t)

Let $\pi \leftarrow (t)$
while *first node of* π *is not* s **do**
  $v \leftarrow$ first node of π
  if *there exists an edge* $(u, v) \in E(H)$ *with* u *reachable from* s **then**
    append any such edge (u, v) to the front of π
  else
    append to the front of π any edge of the form $(u, v) \in E(G)$ with u reachable from s
return π

Algorithm 2: `backwards-growth` path generation

The two algorithms are symmetric to each other, and differ only in whether we grow the path from front to back or from back to front.

As we use these algorithms to sequentially generate paths and build our preserver, it will be helpful to track an auxiliary path system $Z = (V, \Pi)$. Each time we add a path $\pi(s, t)$ to H, say that a *new edge* is an edge $e \in \pi(s, t)$ that was not previously in the preserver. We then add a corresponding path π' to Z, whose nodes are

$$\pi' := \begin{cases} \{u \mid \text{there is a new edge } (u, v) \in \pi(s, t)\} \cup \{t\} & \text{if } \texttt{forwards-growth} \text{ is used} \\ \{s\} \cup \{v \mid \text{there is a new edge } (u, v) \in \pi(s, t)\} & \text{if } \texttt{backwards-growth} \text{ is used} \end{cases}$$

and in either case, these nodes are ordered in the path π' the same as their order in $\pi(s, t)$. We will also treat Z as an ordered path system, with the paths in Z ordered by the arrival of the demand pairs that generated each path. The following properties of Z all follow straightforwardly from the construction:

Lemma 10 (Properties of Z)**.**

1. Z is acyclic,
2. $\|Z\| = |E(H)| + p$,
3. Under `forwards-growth`, Z has no bridge in which the first arc comes before the river in the ordering of Π. Under `backwards-growth`, Z has no bridge in which the last arc comes before the river in the ordering of Π.

Proof.

1. Since the input graph $G = (V, E)$ is a DAG, the order of nodes in each path $\pi \in \Pi$ agrees with the topological ordering of the nodes in V, implying that Z is acyclic.

2. Initially, we have $\|Z\| = |E(H)| = 0$. Then, every path π' added to Z corresponds to a path $\pi(s, t)$ that contributes exactly $|\pi'|-1$ new edges to H, so in the end we have $\|Z\| = |E(H)|+p$.

3. We will prove this for `forwards-growth`; the argument for `backwards-growth` is symmetric (up to reversal of direction of the edges of the input graph G). Seeking contradiction, suppose there is a bridge formed by nodes $(x_1, \ldots, x_k)$, arc paths $\pi_1, \ldots, \pi_{k-1}$, and river path π_r, with $\pi_1 <_\Pi \pi_r$. Let $\pi(s_1, t_1), \pi(s_r, t_r)$ be the paths generated by `forwards-growth` corresponding to π_1, π_r respectively. By construction, since $x_1 \in (\pi_1 \cap \pi_r)$, these paths both contribute new edges to H leaving x_1; call the first one $(x_1, y) \in \pi(s_1, t_1)$. Now notice that the arcs witness reachability among all of the node pairs

$$\underbrace{(y, x_2)}_{\text{in } \pi_1}, \underbrace{(x_2, x_3)}_{\text{in } \pi_2}, \ldots, \underbrace{(x_{k-1}, x_k)}_{\text{in } \pi_{k-1}}, \underbrace{(x_k, t_r)}_{\text{in } \pi_k}.$$

   So by transitivity, the node pair (y, t_r) is reachable. When we generate $\pi(s_r, t_r)$ using `forwards-growth`, since we have already added $\pi(s_1, t_1)$ the edge (x_1, y) is already present in H and we have (y, t_r) reachability. So the algorithm will *not* choose to add a new edge leaving x_1 while generating $\pi(s_r, t_r)$, which completes the contradiction. □

3.2 Online Source-Restricted Reachability Preservers

We will prove the following upper bound on source-restricted preservers in the online model:

Theorem 11. In the online model with an n-node input DAG $G = (V, E)$ and p total demand pairs, the final preserver H will have size

$$|E(H)| \leq O\left((np|S|)^{1/2} + n\right)$$

in either of the following two settings:

- S is the set of start nodes used by the given demand pairs P (that is, $P \subseteq S \times V$), and the builder generates paths in each round using the `backwards-growth` algorithm, or
- S is the set of end nodes used by the given demand pairs P (that is, $P \subseteq V \times S$), and the builder generates paths in each round using the `forwards-growth` algorithm.

Notably, neither the `forwards-` nor `backwards-`growth algorithm require the builder to know the number of demand pairs p or any information about the set of source/sink nodes S. We will only prove the latter point in Theorem 11, analyzing `forwards-growth` and assuming $P \subseteq V \times S$. The other point is symmetric.

The proof will work by analyzing the path system Z associated to the online path-adding process; recall its essential properties in Lemma 10. Let $\ell := \|Z\|/p$ be the average path length and let $d := \|Z\|/n$ be the average node degree. If $d \leq O(1)$ then we have $\|Z\| \leq O(n)$ and the theorem holds, so we may assume in the following that d is at least a sufficiently large constant. Imagine that we add the paths from Z to an initially-empty system, one at a time, in the **reverse** of their ordering in Π. We observe:

Lemma 12. There is a path $\pi \in \Pi$ such that when π is added in the above process, it contains at least $\ell/4$ nodes of degree at least $d/4$ each.

Proof. Suppose not. Then, by counting the first $d/4$ times each node appears in a path separate from the remaining times, the total size of Z can be bounded as

$$\begin{aligned} \|Z\| &\leq \frac{nd}{4} + \frac{p\ell}{4} \\ &\leq \frac{\|Z\|}{4} + \frac{\|Z\|}{4} \\ &= \frac{\|Z\|}{2}, \end{aligned}$$

which is a contradiction. □

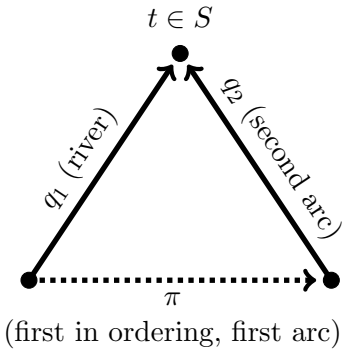


Figure 2: The proof of Lemma 13 works by arguing that no path π may intersect too many paths that come later in the ordering, or else two of those paths q_1, q_2 will share an endpoint in S and thus form a forbidden 3-bridge.

Lemma 13. $\ell d \leq O(|S|)$.

Proof. Suppose for contradiction that $\ell d > 16|S|$. By the previous lemma, there is a path $\pi \in \Pi$ that intersects at least $\ell d/16 > |S|$ other paths in Π, which were added to the system before π (and hence come *later* than π in the ordering of Π). By the Pigeonhole principle, and since the demand pairs satisfy $P \subseteq V \times S$, at least two of these intersecting paths q_1, q_2 end at the same node $t \in S$. Since by Lemma 10 Z does not contain any 2-bridges, q_1, q_2 may not intersect at any other nodes, and so they intersect π at two different nodes. But this implies that π, q_1, q_2 form a 3-bridge in which π is the first arc *and* it precedes both q_1, q_2 in the ordering of Π (see Figure 2). This contradicts Lemma 10, completing the proof. □

We now complete the proof by algebraically rearranging the inequality from the previous lemma. We have:

$$\begin{aligned} \ell d &\leq O(|S|) \\ (p\ell)(nd) &\leq O(|S|pn) \\ \|Z\|^2 &\leq O(|S|pn) \\ \|Z\| &\leq O(|S|pn)^{1/2}. \end{aligned}$$

Since by Lemma 10 we have $\|Z\| \geq |E(H)|$, this implies our desired bound on the size of the output preserver.

3.3 Online Pairwise Reachability Preservers

By following an identical proof strategy to our upper bound in the source-restricted setting (i.e., exploiting forbidden ordered 2- and 3-bridges), it is possible to prove an upper bound of

$$|E(H)| \leq O\left((np)^{2/3} + n\right)$$

(details omitted, since we will show a stronger bound than this). As discussed in [8], this is probably the best upper bound one can show by exploiting *only* the forbidden ordered 2- and 3-bridges from Lemma 10. Nonetheless, we will show a stronger bound, which crucially also exploits the forbidden ordered 4-bridges from Lemma 10.

Recap of [8]. Recent work of Bodwin, Hoppenworth, and Trabelsi [8] on offline reachability preservers introduced a framework for extremal analysis of forbidden 4-bridges, which we will briefly recap here. First, the paper shows:

Lemma 14 (Independence Lemma, c.f. [8], Lemma 38)**.** Let $\beta(n, p, \infty)$ denote the maximum possible size of a path system with n nodes, p paths, and no bridges as subsystems. Then every n-node directed graph and set of p demand pairs has an offline reachability preserver H of size

$$|E(H)| \leq O\left(\beta(n, p, \infty)\right),$$

and this is asymptotically tight.

Thus, it suffices to argue about the extremal size of a bridge-free path system. We remark here that versions of this lemma are perhaps implicit at a low level in prior work, e.g. [1, 19]. It is also an inherently offline lemma, and breaks down completely in the online setting; the reliance on this lemma (or its underlying ideas) is essentially why the known results for offline reachability preservers do not tend to extend to the online setting.

This previous paper then argues as follows. Assume for convenience that all paths have length $\Theta(\ell)$ and all nodes have degree $\Theta(d)$, for some parameters ℓ, d. Recall that the upper bound from forbidden 2- and 3-bridges is $O((np)^{2/3} + n)$, and so our goal is to show that this bound cannot be tight. A few straightforward calculations reveal that, *if* this bound were tight, then for the typical pair of paths π_1, π_2 in the system there will be $\Theta(\ell)$ paths that intersect π_1 and then π_2. However, no pair of these intersecting paths may have crossing intersection points with π_1, π_2 (see Figure 3).

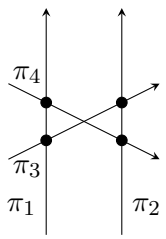


Figure 3: There cannot be two paths π_3, π_4 that both intersect paths π_1, π_2, but where the points of intersection switch places as in this picture, or else they form a 4-bridge (here π_3 is the river).

If there are $\Theta(\ell)$ paths that intersect both π_1, π_2, and yet these intersecting paths cannot cross each other, then the typical intersecting path must "lie flat" in the sense that there is not much of a gap along either π_1 or π_2 to, say, the h nearest intersecting paths (for some parameter h). In order to exploit this, the key strategy in [8] is to sample a random *base path* $\pi_b \in \Pi$, and then analyze the random subsystem S' on the vertex set formed by examining the h adjacent nodes along the paths that intersect π_b (see Figure 4).

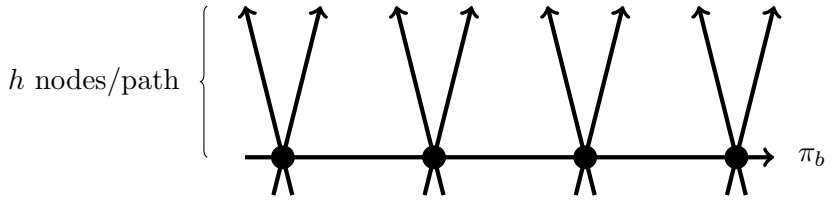


Figure 4: The random subsystem S'. Figure based on [8], Figure 7.

If the intersecting paths do indeed "lie flat," then there will be many such paths within h steps of π_b along its branching paths, and thus we should expect S' to contain many long paths. But we can apply known upper bounds to S' to rule out this possibility. This implies that, in fact, the typical pair of paths π_1, π_2 have $\ll \ell$ paths that intersect both, leading to an improved upper bound.

Our Offline Improvements. An auxiliary result of this paper is an improvement in the bound shown by [8]. We refer back to Theorem 4 for the statement, or Appendix B for the proof.

The source of these improvements is from an improved strategy for controlling the size of the random subsystem S'. One of the two ways in which this part improved is by *recursively* bounding the size of S' (this is executed in Lemma 40). Although this idea is conceptually straightforward, it requires a significant refactoring of the proof to enable it. The technical reason is that [8] bounds the input system Z using an ℓ^1 norm of path lengths (the standard notion of size) but S' using an ℓ^2 norm of path lengths, making it impossible to directly apply the bound on Z recursively to S'. We switch to bounding both using the ℓ^2 norm everywhere, and we only move back to our desired ℓ^1 norm at the very end of the proof. The other new ingredient is an improved counting of the contribution of "short" paths to the size of S', which is executed in Lemma 40.

Our Online Adaptation. It will be slightly more convenient in this exposition to consider the `backwards-growth` strategy for path generation here, although of course by symmetry either strategy works (we use this convention in Appendix A as well). We will analyze the path system Z constructed above, and in particular Lemma 10 states that bridges are forbidden in Z whose *last* arc comes before the river.

Can we exploit forbidden ordered 4-bridges by following the strategy outlined above? Some parts of the method extend easily, with minor tweaks. For example, instead of counting *any* paths π that intersect the typical pair of paths π_1, π_2, we can restrict attention to those that also come after π_1, π_2 in the ordering. Then the "crossing" in Figure 3 is still forbidden (since the river π_3 is assumed to come after the last arc π_2). Relatedly, when we define S', we need to consider only the paths that intersect π_b *and* come before it. Nonetheless, all these definitional adaptations turn out to affect the relevant counting arguments by only constant factors, and so they do not harm the argument. There are many other minor issues that we will not overview, which can be dispatched with a little technical effort.

However, there is one major problem: there is now potential for overlap in the paths that branch off π_b. That is, in the offline setting, we can exploit 3-bridge-freeness to argue that no two paths that branch off π_b also intersect each other, and thus every node in S' (except those on π_b itself) is in exactly one branching path. But in the online setting, this is not so: these intersections form a 3-bridge that may be allowed (see Figure 5).

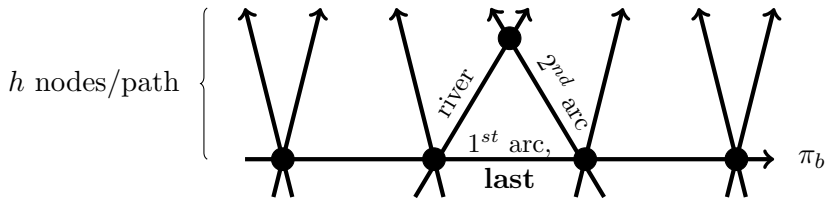


Figure 5: When we generate S' in the online/ordered setting, it is possible for the paths intersecting π_b to intersect each other: the river *could* come before the 2^{nd} arc, which would not violate the conditions of Lemma 10.

Naively, the typical node in S' could be $\Theta(d)$ branching paths, and this complication completely wipes out all the gains from the analysis. Although we cannot rule out the possibility of some nodes having $\Theta(d)$ branching paths, we are able to use a more intricate maneuver to show that there is some threshold $t \ll d$ such that, in expectation, the number of nodes of branching degree $\Omega(x)$ in S' is far less than the trivial bound of $\ell dh/x$ for $x \geq t$. This turns out to be good enough to recover some gains from the method. We unfortunately cannot show this for $t = 1$, which is the fundamental reason why our online bounds are polynomially worse than the corresponding offline bounds.

4 Non-Adaptive Reachability Preservers

We next describe our method to convert our online algorithms to (almost) non-adaptive algorithms. Our algorithm essentially works by simulating a *greedy adversary* in the online model, who repeatedly provides the most costly demand pair in each round, and then we set paths by running our online algorithm against this adversary, halting at the appropriate place. Our proof will thus imply indirectly that this greedy strategy is the most effective one for an adversary in the online model, up to constant factors.

Our algorithms will reference an *extremal function* $f(n, p)$ for online reachability preservers, achieved by a generic path selection algorithm $\pi(s, t \mid G, H)$. That is, we imagine an algorithm that starts with G as any n-node graph, and H as the n-node empty graph. When each demand pair (s, t) arrives, we select the path $\pi(s, t \mid G, H)$ and add all of its edges to H. Then $f(n, p)$ is the largest possible number of edges in H after p rounds of this process.

We use this generic extremal function f, rather than the particular extremal upper bound from Theorem 1, in order to emphasize that this bound and the technical details of the `forwards-` or `backwards-growth` algorithms are not really important in this proof. If a future result improves on Theorem 1, then the new bound will automatically transfer to these results as well. For simplicity we will focus on $f(n, p)$ here, but the proof would generalize readily to extremal functions that incorporate additional parameters beyond n and p. This includes the online source-restricted preservers of Theorem 11, which incorporate $|S|$ as a parameter, although we note that these require the set S to be given on input, so that the we know the set of possible demand pairs and

we can properly simulate the adversary (i.e., search over the proper subset of demand pairs in the condition of the while loop).

Input: n-node directed graph $G = (V, E)$, number of demand pairs p
All reachable node pairs (s, t) in G have "unfinalized" path
$H \leftarrow (V, \emptyset)$
Let $f(n, p)$ be an extremal function for online reachability preservers, achieved by a deterministic path selection algorithm $\pi(s, t \mid G, H)$
while *there is reachable* (s, t) *with* $> f(n, p)/p$ *edges in* $\pi(s, t \mid G, H) \setminus E(H)$ **do**
  finalize path $\pi(s, t \mid G, H)$ for (s, t)
  add edges of $\pi(s, t \mid G, H)$ to H
foreach *remaining unfinalized reachable pair* (s, t) **do**
  finalize path $\pi(s, t \mid G, H)$ for (s, t)

Algorithm 3: `known-p-non-adaptive-rps`

Theorem 15. For all n-node graphs G and sequences P of $|P| =: p$ demand pairs, the paths set by Algorithm 3 satisfy

$$\left| \bigcup_{(s,t) \in P} \pi(s, t) \right| \leq 2f(n, p).$$

Proof. Let Q be the set of demand pairs whose paths are set in the initial while loop, and let H_Q be the subgraph H just after the paths for Q have been set and the while loop terminates. We first note that $|Q| < p$, since otherwise by counting the edges contributed to H, the first p demand pairs in Q create a subgraph H_Q of size $|E(H_Q)| > f(n, p)$ which contradicts the definition of the extremal function f. Since $|Q| < p$, we therefore have

$$|E(H_Q)| \leq f(n, p).$$

Meanwhile, all demand pairs in $P \setminus Q$ have their path set in the final for loop, and by construction there are $\leq f(n, p)/p$ edges outside H_Q in each path. So we have

$$\begin{aligned} \left| \bigcup_{(s,t) \in P} \pi(s, t) \right| &\leq |E(H_Q)| + \sum_{(s,t) \in P \setminus Q} |\pi(s, t) \setminus E(H_Q)| \\ &\leq f(n, p) + p \cdot \left(\frac{f(n, p)}{p} \right) \\ &= 2f(n, p). \end{aligned}$$

□

This theorem implies that, if one uses the precomputed paths from Algorithm 3 to respond to online queries, then the online upper bound of $f(n, p)$ will still apply (up to a factor of 2). The main weakness of Algorithm 3 is that it requires advance knowledge of the parameter p, in order to compute the threshold $f(n, p)/p$ at which we exit the initial while loop. It is tempting, but incorrect, to think this can be generally avoided by setting *all* paths as in the main while loop. Such a strategy would work for a path selection algorithm that happens to satisfy an axiom like *monotonicity*, for which adding edges to H can only decrease the number of new edges in a selected

Input: n-node directed graph $G = (V, E)$
Let $f(n, p)$ be an extremal function for online reachability preservers, achieved by a deterministic path selection algorithm $\pi(s, t \mid G, H)$
Let $p^* := \arg\max_p f(n, p) \leq O(n)$
foreach $q \in \{p^*, 2p^*, 4p^*, 8p^*, \dots\}$ **do**
    Run Algorithm 3 with number-of-paths parameter q
    Denote selected paths by $\pi_q(s, t)$

Algorithm 4: Preprocessing for Index-Sensitive Non-Adaptive Reachability Preservers

Input: demand pair (s, t), index i
`// run Algorithm 4 as preprocessing`
Let q be the least value in $\{p^*, 2p^*, 4p^*, 8p^*, \dots\}$ with $q \geq i$
Return $\pi_q(s, t)$

Algorithm 5: Path Selection for Index-Sensitive Non-Adaptive Reachability Preservers

path $\pi(s, t \mid G, H)$ (it might also be fine to tolerate an approximate version of monotonicity). However, we note that the path selection algorithms (`forwards-` and `backwards-growth`) used in our online upper bounds are **not** monotonic in this way (or even approximately monotonic).

That said, we next describe a wrapper for the algorithm that can avoid the need to know p ahead of time:

Theorem 16. Suppose that the extremal function $f(n, p)$ depends polynomially on its second parameter p in the regime where $p \geq p^*$.[4] Then the online algorithm that runs Algorithm 4 as a preprocessing routine upon receiving G, and which then uses Algorithm 5 to select the path added to the preserver for each i^{th} demand pair (s, t), will construct a reachability preserver H of size $|E(H)| \leq O(f(n, p))$.

Proof. For each possible choice of q, we will add at most q paths selected by π_q to the preserver. By Theorem 15, these paths will have at most $2f(n, q)$ edges in their union. Additionally, letting q^* be the largest choice of q for which we add any corresponding paths, note that we have $q^* \geq p \geq q^*/2$. So we can bound the total number of edges in the preserver as:

$$\begin{aligned} |E(H)| &\leq \sum_{q \in \{p^*, 2p^*, 4p^*, 8p^*, \dots, q^*\}} 2f(n, q) \\ &\leq O(f(n, q^*)) \\ &\leq O(f(n, p)). \end{aligned}$$

Here the second inequality holds because f depends polynomially on its second parameter, and so this sum is asymptotically dominated by its largest term. The third inequality holds because we have $p \geq q^*/2$, and (again since f depends polynomially on its second parameter) this means the values of $f(n, q^*), f(n, p)$ differ by at most a constant factor. □

[4]This phrase "depends polynomially" is intuitive but a bit informal. It is tedious to formalize it, but what we really mean is that this theorem holds for any function f for which the latter two inequalities in the chain hold. This includes all upper bounds shown in this paper.

5 Online Unweighted Directed Steiner Forest Algorithms

We will next apply our extremal bounds for online reachability preservers to the problem of Online UDSF. As a reminder, in this problem we receive an n-node directed graph $G = (V, E)$ on input, and then in each round we receive a new demand pair (s, t) that is reachable in G. We must irrevocably add edges to a reachability preserver H to ensure that (s, t) is reachable in H before the next demand pair is received. We do not know the number of demand pairs p ahead of time. We will denote by OPT the size of the smallest possible (offline) reachability preserver for G, P, where P is the set of all demand pairs received.

5.1 Recap of the Grigorescu-Lin-Quanrud Bound

Our new bound will use the structure and several technical ingredients from the previous state-of-the-art online algorithm by Grigorescu, Lin, and Quanrud [24]. They proved:

Theorem 17 ([24])**.** For online UDSF, there is a randomized polynomial time algorithm with competitive ratio $O(n^{2/3+\varepsilon})$.

Their algorithm carries two parameters, T and τ, which will be set at the end by a balance. In the following, we will say that a demand pair (s, t) is *nontrivial* if it is not already reachable when it arrives, and thus it requires us to add at least one new edge to the preserver. We let p be the total number of nontrivial demand pairs.

- For the first T nontrivial demand pairs that arrive, we use the following result by Chakrabarty et al.:

  Theorem 18 ([21])**.** There is a randomized polynomial-time online algorithm that constructs a preserver of the first T demand pairs of size at most $OPT \cdot O(T^{1/2-\varepsilon})$.

- After the first T nontrivial demand pairs, the remaining demand pairs are further classified using a strategy from previous work on offline DSF [6,16]. Say that a node pair (s, t) is *τ-thin* if the number of vertices that lie along $s \rightsquigarrow t$ paths is at most τ, or *τ-thick* otherwise.

  - In order to handle the thick demand pairs, just after the T^{th} demand pair is processed, we randomly sample a set of $|S| = Cn \log n/\tau$ nodes, where C is a sufficiently large constant. Let us say that a node pair (s, t) is *hit* by S if there exists a node $v \in S$ that lies along an $s \rightsquigarrow t$ path. By standard Chernoff bounds (omitted), with high probability, every thick pair (s, t) is hit by S; in the following we will assume that this high-probability event occurs. We then add an in- and out-tree from each sampled node in S, and so together these trees will contain an $s \rightsquigarrow t$ path. This costs $\widetilde{O}(n^2/\tau)$ edges in total.
  - When each demand pair (s, t) arrives, we first check whether or not it is hit by S. If so, the pair has been satisfied already by our trees and we can do nothing. If not, then (s, t) must be τ-thin. In this case, the algorithm appeals to an LP-rounding algorithm from [6]:

    Theorem 19 ([6])**.** There is a randomized polynomial-time online algorithm that constructs a preserver of all τ-thin pairs of size at most $OPT \cdot \widetilde{O}(\tau)$.

This completes the construction. The last technical ingredient required is the following existential lower bound on OPT:

Lemma 20 ([24])**.** $OPT \geq \Omega(p^{1/2})$.

Proof. Any set of p demand pairs must use at least $p^{1/2}$ total terminal nodes (start or end). In order to preserve connectivity, all terminal nodes must have in- or out-degree at least 1. It follows that any correct solution must have $\Omega(p^{1/2})$ edges. □

Now, assuming that $p \geq T$ we can bound the total competitive ratio as

$$\begin{aligned}
&\frac{OPT \cdot O(T^{1/2+\varepsilon}) + \widetilde{O}\left(\frac{n^2}{\tau}\right) + OPT \cdot \widetilde{O}(t)}{OPT} \\
&= O(T^{1/2+\varepsilon}) + \widetilde{O}(\tau) + \widetilde{O}\left(\frac{n^2}{OPT \cdot \tau}\right) \\
&\leq O(T^{1/2+\varepsilon}) + \widetilde{O}(\tau) + \widetilde{O}\left(\frac{n^2}{p^{1/2} \cdot \tau}\right) \\
&\leq O(T^{1/2+\varepsilon}) + \widetilde{O}(\tau) + \widetilde{O}\left(\frac{n^2}{T^{1/2} \cdot \tau}\right).
\end{aligned}$$

With a parameter balance, one can compute that the optimal setting is (essentially) $T = n^{4/3}$ and $\tau = n^{2/3}$, yielding the claimed competitive ratio of $O(n^{2/3+\varepsilon})$. In the case where $p \leq T$, the bound on competitive ratio is simply $O(T^{1/2+\varepsilon})$, which leads to the same bound.

5.2 Our Adaptation

We improve the competitive ratio from [24]:

Theorem 21. For online DSF with uniform costs, there is a randomized polynomial time algorithm with competitive ratio $O(n^{3/5+\varepsilon})$.

We mostly follow the strategy from [24] outlined previously, but our main change is to the handling of thick pairs. After the first T demand pairs have been processed, we again sample a set S of $|S| = \widetilde{O}(n/\tau)$ nodes, and we note that with high probability this sample hits all τ-thick pairs. However, unlike before, we do not add in- or out-trees from S. Instead, when each new demand pair (s, t) arrives, our strategy is to check whether or not it is hit by S. If so, then we use the two cases of Theorem 11 to add paths for the pairs (s, v) and (v, t), at total cost $O((np|S|)^{1/2} + n)$ over all demand pairs. If (s, t) is not hit by S, then the new demand pair (s, t) must be τ-thin, and we handle it using Theorem 19 like before.

It is intuitive at this point that handling thick pairs with an improved bound would lead to an improved competitive ratio. But unfortunately, it is not so simple: plugging in the improved bound to the previous competitive ratio will not lead to an improvement. The reason is that the previous parameter settings of $T = n^{4/3}, \tau = n^{2/3}$ correspond to the setting where we have $|S| = \widetilde{O}(n^{1/3})$ nodes in our sample and $p = T = n^{4/3}$ demand pairs in total, and in this setting our new bounds roughly tie (up to $\log n$ factors) those obtained from using in- and out-trees. To get around this, we will need a more nuanced version of Lemma 20:

Lemma 22. Let p' be the number of nontrivial demand pairs that are hit by S. Then $OPT \geq \Omega\left(p'/|S|\right)$.

Proof. Similar to Lemma 20, it suffices to argue that the demand pairs use $\Omega\left(p'/|S|\right)$ distinct terminal nodes (start or end), since each terminal node must have (in or out) degree at least 1.

For every demand pair (s,t) that is hit by S, there is a node $v \in S$ for which we add $s \rightsquigarrow v$ and $v \rightsquigarrow t$ paths to the preserver. This must either be the first time we add an $s \rightsquigarrow v$ path, or the first time we add a $v \rightsquigarrow t$ path, since otherwise an $s \rightsquigarrow t$ path in the preserver will already exist and the demand pair will be trivial. Thus the number of start or end terminals that have been paired with v increases by 1. Overall, since there are $|S|$ possible nodes that could hit our demand pairs, we must have at least $p'/|S|$ terminal nodes in total. □

We are now ready to calculate competitive ratio. Assuming that $p \geq T$, this is

$$\begin{aligned}
&O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + O\left(\frac{|S|^{1/2}n^{1/2}p'^{1/2}}{OPT}\right) + O\left(\frac{n}{OPT}\right) \\
&\leq O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + O\left(\frac{|S|^{1/2}n^{1/2}p'^{1/2}}{\max\{p'/|S|, p^{1/2}\}}\right) + O\left(\frac{n}{p^{1/2}}\right) && \text{Lemmas 20, 22} \\
&= O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + O\left(\min\left\{\frac{|S|^{3/2}n^{1/2}}{p'^{1/2}}, \frac{|S|^{1/2}n^{1/2}p'^{1/2}}{p^{1/2}}\right\}\right) + O\left(\frac{n}{p^{1/2}}\right) \\
&= O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + \widetilde{O}\left(\min\left\{\frac{n^2}{p'^{1/2}\tau^{3/2}}, \frac{np'^{1/2}}{p^{1/2}\tau^{1/2}}\right\}\right) + O\left(\frac{n}{p^{1/2}}\right) && |S| = \widetilde{O}(n/\tau) \\
&\leq O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + \widetilde{O}\left(\min\left\{\frac{n^2}{p'^{1/2}\tau^{3/2}}, \frac{np'^{1/2}}{T^{1/2}\tau^{1/2}}\right\}\right) + O\left(\frac{n}{T^{1/2}}\right). && p \geq T
\end{aligned}$$

This bound will be maximized when p' is such that the two terms in the min balance, which occurs when

$$p' = n \cdot \frac{T^{1/2}}{\tau}.$$

Under this setting, we can simplify

$$\leq O\left(T^{1/2+\varepsilon}\right) + \widetilde{O}(\tau) + \widetilde{O}\left(\frac{n^{3/2}}{T^{1/4}\tau}\right) + O\left(\frac{n}{T^{1/2}}\right).$$

Finally, we are ready to choose τ, T to balance these terms. By setting

$$\tau := n^{3/5}, T := n^{6/5},$$

the above expression becomes

$$\begin{aligned}
&O\left(n^{3/5+\varepsilon}\right) + \widetilde{O}(n^{3/5}) + \widetilde{O}\left(\frac{n^{3/2}}{n^{3/10} \cdot n^{3/5}}\right) + O\left(\frac{n}{n^{3/5}}\right) \\
&= O\left(n^{3/5+\varepsilon}\right).
\end{aligned}$$

Finally, as before, in the case where $p \leq T$ the competitive ratio is $O(T^{1/2+\varepsilon})$, giving the same bound.

References

[1] Amir Abboud and Greg Bodwin. Reachability preservers: New extremal bounds and approximation algorithms. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1865–1883. Society for Industrial and Applied Mathematics, 2018.

[2] Noga Alon. Testing subgraphs in large graphs. *Random Structures & Algorithms*, 21(3-4):359–370, 2002.

[3] Noga Alon, Baruch Awerbuch, and Yossi Azar. The online set cover problem. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 100–105, 2003.

[4] Noga Alon, Baruch Awerbuch, Yossi Azar, Niv Buchbinder, and Joseph Naor. A general approach to online network optimization problems. *ACM Transactions on Algorithms (TALG)*, 2(4):640–660, 2006.

[5] Surender Baswana, Keerti Choudhary, and Liam Roditty. Fault tolerant subgraph for single source reachability: generic and optimal. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 509–518. ACM, 2016.

[6] Piotr Berman, Arnab Bhattacharyya, Konstantin Makarychev, Sofya Raskhodnikova, and Grigory Yaroslavtsev. Approximation algorithms for spanner problems and directed steiner forest. *Information and Computation*, 222:93–107, 2013.

[7] Greg Bodwin. New results on linear size distance preservers. *SIAM Journal on Computing*, 50(2):662–673, 2021.

[8] Greg Bodwin, Gary Hoppenworth, and Ohad Trabelsi. Bridge girth: A unifying notion in network design. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 600–648. IEEE, 2023.

[9] Niv Buchbinder and Joseph Naor. Improved bounds for online routing and packing via a primal-dual approach. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 293–304. IEEE, 2006.

[10] Niv Buchbinder and Joseph Naor. Online primal-dual algorithms for covering and packing. *Mathematics of Operations Research*, 34(2):270–286, 2009.

[11] Niv Buchbinder, Joseph Seffi Naor, et al. The design of competitive online algorithms via a primal–dual approach. *Foundations and Trends® in Theoretical Computer Science*, 3(2–3):93–263, 2009.

[12] Diptarka Chakraborty, Kushagra Chatterjee, and Keerti Choudhary. Pairwise Reachability Oracles and Preservers Under Failures. In Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:16, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[13] Diptarka Chakraborty and Keerti Choudhary. New Extremal Bounds for Reachability and Strong-Connectivity Preservers Under Failures. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 25:1–25:20, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[14] Moses Charikar, Chandra Chekuri, To-Yat Cheung, Zuo Dai, Ashish Goel, Sudipto Guha, and Ming Li. Approximation algorithms for directed steiner problems. *Journal of Algorithms*, 33(1):73–91, 1999.

[15] Chandra Chekuri, Guy Even, Anupam Gupta, and Danny Segev. Set connectivity problems in undirected graphs and the directed steiner network problem. *ACM Transactions on Algorithms (TALG)*, 7(2):1–17, 2011.

[16] Eden Chlamtáč, Michael Dinitz, Guy Kortsarz, and Bundit Laekhanukit. Approximating spanners and directed steiner forest: Upper and lower bounds. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 534–553. SIAM, 2017.

[17] Keerti Choudhary. An optimal dual fault tolerant reachability oracle. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 55. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[18] Julia Chuzhoy and Sanjeev Khanna. Polynomial flow-cut gaps and hardness of directed cut problems. *Journal of the ACM (JACM)*, 56(2):1–28, 2009.

[19] Don Coppersmith and Michael Elkin. Sparse sourcewise and pairwise distance preservers. *SIAM Journal on Discrete Mathematics*, 20(2):463–501, 2006.

[20] Yevgeniy Dodis and Sanjeev Khanna. Design networks with bounded pairwise distance. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 750–759, 1999.

[21] Alina Ene, Deeparnab Chakrabarty, Ravishankar Krishnaswamy, and Debmalya Panigrahi. Online buy-at-bulk network design. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 545–562. IEEE, 2015.

[22] Moran Feldman, Guy Kortsarz, and Zeev Nutov. Improved approximation algorithms for directed steiner forest. *Journal of Computer and System Sciences*, 78(1):279–292, 2012.

[23] Michel X Goemans and David P Williamson. A general approximation technique for constrained forest problems. *SIAM Journal on Computing*, 24(2):296–317, 1995.

[24] Elena Grigorescu, Young-San Lin, and Kent Quanrud. Online Directed Spanners and Steiner Forests. In Mary Wootters and Laura Sanità, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021)*, volume 207 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:25, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.

[25] Shang-En Huang and Seth Pettie. Lower Bounds on Sparse Spanners, Emulators, and Diameter-reducing shortcuts. In David Eppstein, editor, *16th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2018)*, volume 101 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:12, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[26] Pawel Winter. Steiner problem in networks: A survey. *Networks*, 17(2):129–167, 1987.

A Online Pairwise Reachability Preservers

We will next prove Theorem 1. Recall from Lemma 10 that, if we construct our online reachability preserver using `forwards-` or `backwards-growth`, then it suffices to bound the size of the associated path system Z as defined in Section 3.1. In particular, under `backwards-growth`, Z will have the following property:

Definition 2 (Half-Bridge-Freeness)**.** A ordered path system is said to be half-k-bridge-free if there are no bridges of size at most k in which the last arc comes before the river.

The focus of our proof will shift to bounding the maximum possible size of *any* half-bridge-free system. That is, let $H(n, p, k)$ denotes the maximum size of a half-k-bridge-free system with n nodes and p paths, and then from Lemma 10 we have

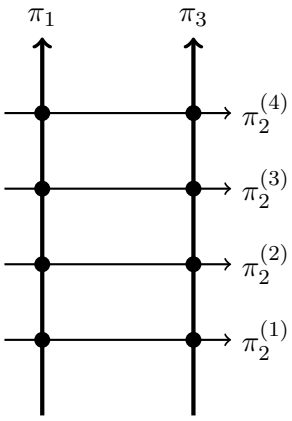
$$\|Z\| \leq H(n, p, \infty) \leq H(n, p, 4).$$

So we may focus on providing an upper bound for $H(n, p, 4)$.

A.1 Proof Overview

The proof gets quite technical in places, so let us start with a higher-level overview of the proof strategy. Let $Z = (V, \Pi)$ be a system with n-nodes, p-paths, and no half 2, 3, or 4 bridges. By the standard *cleaning lemma* (Lemma 23), we may assume that all nodes have degree $\Theta(d)$, and all paths have length $\Theta(\ell)$. For simplicity, we will assume in this overview that all nodes have degree *exactly* d and all paths have length *exactly* ℓ, which will not materially affect the argument.

Recap of Offline Proof from [8]. The previous-best offline upper bound from [8] focuses on a collection of sets $\{R(\pi_1, \pi_3) \mid \pi_1, \pi_3 \in \Pi\}$ which are each the subsets of paths from Π that intersect π_1 and then later intersect π_3:

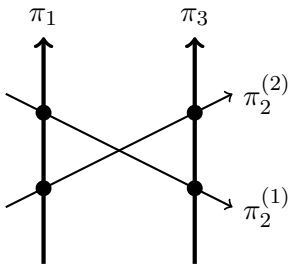


$$R(\pi_1, \pi_3) = \left\{\pi_2^{(1)}, \pi_2^{(2)}, \pi_2^{(3)}, \pi_2^{(4)}\right\}$$

Naively, these sets can have maximum size $|R(\pi_1, \pi_3)| \leq \ell^2$, since there are ℓ^2 ways to choose a node from π_1 and then π_3, and no two paths can use the same pair of intersection points (or else they form a forbidden 2-bridge). Some straightforward algebra from there leads to an initial but very suboptimal bound of

$$\|Z\| \leq O\left(\min\left\{np^{1/2}, n^{1/2}p\right\} + n + p\right).$$

The next improvement comes by observing that we can actually only pack ℓ paths between π_1 and π_3, rather than ℓ^2. This is essentially because the paths in $R(\pi_1, \pi_3)$ cannot cross each other, as in the following picture, or else they will form a forbidden bridge:

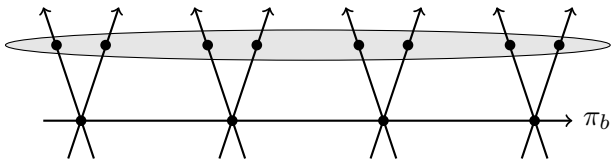


Forbidden:
forms a 4-bridge
($\pi_2^{(2)}$ is the river)

Redoing the algebra with this improved bound $|R(\pi_1, \pi_3)| \leq \ell$ leads to a better bound of

$$\|Z\| \leq O\left((np)^{2/3} + n + p\right).$$

It is indeed possible for *some* sets to have size $|R(\pi_1, \pi_3)| = \ell$, but the next round of improvements works by arguing that the *typical* such set must be slightly smaller. For intuition: suppose towards contradiction that every set has size exactly $|R(\pi_1, \pi_3)| = \ell$. This can occur only if every path $\pi_2 \in R(\pi_1, \pi_3)$ is perfectly *aligned* with π_1 and π_3: that is, if π_2 intersects π_1 on its i^{th} node, then it also intersects π_3 on its i^{th} node. But then - under this assumption of alignment - we can argue that our paths must be unusually concentrated. Consider an arbitrary *base path* π_b, and consider the induced subsystem of nodes that come one step after π_b, along paths that intersect π_b:



consider subsystem
on these nodes

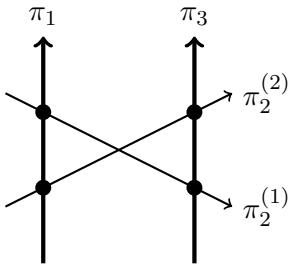
This subsystem will have ℓd nodes, node degrees d, and (due to alignment) the paths in this subsystem will have length ℓ. However, applying the previously-shown bounds to this subsystem reveals that these particular parameters imply a contradiction: the typical length of a path in this induced subsystem must actually be $\ll \ell$, giving contradiction.

Formalizing this intuition gets rather technical. The key moving parts that were not covered in the above sketch are:

- We need to select the base path π_b *randomly* rather than arbitrarily, and then measure the *expected* value of the various statistics of the system.
- We focus on a subsystem formed by h nodes along paths intersecting π_b, rather than just 1 node. (Here h is a new parameter, set by a parameter balance at the end of the proof.)
- The proof involves toggling between bounding the sum of path lengths (which is $\|Z\|$), and bounding the sum of *squared* path lengths. This is necessary because (1) the sum of sizes of $R(\pi_1, \pi_3)$ sets naturally scales with the sum of squared path lengths, rather than the sum of path lengths, and (2) when we focus on our induced subsystem, it may break the cleaning lemma: it is not still guaranteed that all paths have the same length *when restricted to that subsystem.*

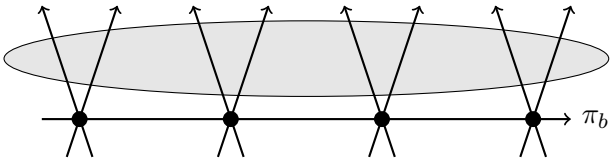
A New Optimization. One of the ways the current proof differs from [8] is in an improved handling of the toggling between sum-of-path-lengths and sum-of-squared-path-lengths. Roughly, instead of switching back and forth between these as the proof proceeds, the proof focuses almost entirely on sum-of-squared-path-lengths. This has several advantages in efficiency; perhaps most notably, it lets us *recursively* apply our bound on the size of Z to control the size of the induced subsystem, rather than only applying the bound $\|Z\| \leq O((np)^{2/3} + n + p)$ once. This is what leads to our improvements in the offline setting. This is also mildly helpful in the online setting, but there are other necessary changes that lead to losses that more than eclipse this improvement.

Changes in the Online Setting. The first change that is required in the online setting is in the definition of the sets $R(\pi_1, \pi_3)$. Instead of including all paths π_2 in these sets that intersect π_1 followed by π_3, we only include such paths π_2 that come after both π_1 and π_3 in the ordering. This restriction only affects the relevant counting by constant factors, and it also suffices to achieve the crucial total-ordering property of the $R(\pi_1, \pi_3)$ sets:



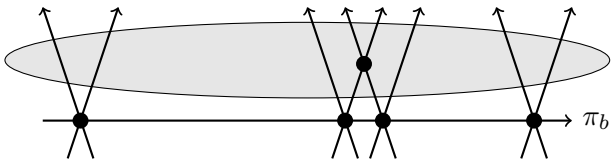
Forbidden:
if $\pi_2^{(1)}$ comes after π_3,
then this forms a
forbidden half-4-bridge
($\pi_2^{(2)}$ is the river)

A more serious problem arises when considering the induced subsystem. In the offline setting, the induced subsystem has ℓdh total nodes. This holds because there are ℓd paths that branch from π_b, and these paths may not intersect each other, or else they form a 3-bridge.



ℓhd nodes

With only half-bridges forbidden, however, it is possible for these branching paths to intersect each other. This will form a 3-bridge, but it will *not* necessarily form a half-3-bridge. This might lead to significantly fewer than ℓdh nodes in the induced subsystem.



possibly fewer than
ℓhd nodes

A lot of our new technical work is to control the amount of overlap that might occur, thus giving a nontrivial lower bound on the number of nodes in the induced subsystem. Roughly, we do this by recursively applying our bounds on the size of a half-bridge-free system in yet another way. Still though, the lower bound on number of nodes is still $\ll ldh$, and this is the fundamental reason why our bounds in the online setting are polynomially worse than the analogous bounds in the offline setting.

A.2 Setup of Formal Proof and Initial Bound

We start with the following standard lemma:

Lemma 23 (Cleaning Lemma – c.f. [8], Lemma 10)**.** There exists a half-k-bridge free system on $\leq n$ vertices, $\leq p$ path whose size is $\Theta(H(n,p,k))$ such that every vertex has degree between $d/4$ and $4d$ and every path has length between $\ell/4$ and 4ℓ where d, ℓ are the average degree and average path length respectively.

Proof. Start with a half-k-bridge free system S with n nodes, p paths, and size $\beta(n,p,k)$. Fix d as the average node degree and ℓ as the average path length (which do not change as we modify the system). Then, perform of the following steps:

- While there exists a node of degree $< d/4$ or a path of degree $< \ell/4$, remove the node or path from the system.
- While there exists a node v of degree $> d$, split it into two nodes $\{v_1, v_2\}$. Replace each instance of v in a path with either v_1 or v_2 in such a way that $\deg(v_1) \in \{deg(v_2), deg(v_2)+1\}$.
- While there exists a path π of length $> \ell$, split it into two node-disjoint subpaths π_1, π_2, with $|\pi_1| \in \{|\pi_2|, |\pi_2|+1\}$.

It is clear from the construction that all remaining nodes have degree in the range $[d/4, d]$, and that all remaining paths have length in the range $[\ell/4, \ell]$, and that the modified system is still half-k-bridge-free. Additionally, by unioning over the nodes and paths, the overall size of the system decreases by $< nd/4 + p\ell/4 = \|S\|/2$ due to deletions, so the size is still $\Theta(\beta(n,p,k))$, and the lemma is satisfied. □

Let ℓ and d denotes the average path length and average node degree respectively, then we have $\|Z\| = nd = p\ell$. For the rest of this section, we assume ℓ and d are at least a sufficiently large constant, otherwise we immediately have $\|Z\| \leq O(n+p)$. By the cleaning lemma, we may assume that every vertex has degree between $d/4$ and $4d$ and every path has length between $\ell/4$ and 4ℓ.

We write $\pi_1 < \pi_2$ if the path π_1 comes before π_2 in the ordering. For nodes $a, b \in \pi_1$, write $a <_{\pi_1} b$ if a comes before b in the path π_1.

Definition 3 (R Sets)**.** Let

$$R = \{(\pi_1, \pi_2, \pi_3) : \pi_1 \cap \pi_2 <_{\pi_2} \pi_2 \cap \pi_3, \pi_3 < \pi_2\}.$$

For any two paths π_1, π_3, we define

$$R(\pi_1, \pi_3) := \{\pi_2 : (\pi_1, \pi_2, \pi_3) \in R\}.$$

Lemma 24 (c.f. [8], Section 3.2.4)**.** For any pair of paths π_1, π_3, there is a total ordering of the elements of $R(\pi_1, \pi_3)$, denoted $<_R$ such that if $\pi_a <_R \pi_b$ then $\pi_a \cap \pi_1 <_{\pi_1} \pi_b \cap \pi_1$ and $\pi_a \cap \pi_3 \leq_{\pi_3} \pi_b \cap \pi_3$. As a corollary, we have $|R(\pi_1, \pi_3)| \leq O(\ell)$.

Proof. We first note that for $\pi_a, \pi_b \in R(\pi_1, \pi_3)$ we have if $\pi_a \neq \pi_b$ then $\pi_a \cap \pi_1 \neq \pi_b \cap \pi_1$. Suppose not, and without loss of generality assume $\pi_a \cap \pi_3 \leq_{\pi_3} \pi_b \cap \pi_3$. If $\pi_a \cap \pi_3 = \pi_b \cap \pi_3$ then π_a, π_b forms a 2-bridge, which is a contradiction. Otherwise, $\pi_a \cap \pi_3 <_{\pi_3} \pi_b \cap \pi_3$, so π_a, π_3, π_b forms a 3 bridge with π_3 being the last arc and π_b is the river and $\pi_3 < \pi_b$, which is also a contradiction.

Thus we can order the elements of $R(\pi_1, \pi_3)$ by $\pi_a <_R \pi_b$ if $\pi_1 \cap \pi_a <_{\pi_1} \pi_1 \cap \pi_b$. It suffices to show that if $\pi_1 \cap \pi_a <_{\pi_1} \pi_1 \cap \pi_b$ then $\pi_a \cap \pi_3 \leq_{\pi_3} \pi_b \cap \pi_3$. Suppose, for contradiction, that there is π_a, π_b such that $\pi_1 \cap \pi_a <_{\pi_1} \pi_1 \cap \pi_b$ but $\pi_a \cap \pi_3 >_{\pi_3} \pi_b \cap \pi_3$. Then we have $\pi_1, \pi_b, \pi_3, \pi_a$ forms a 4-bridge with π_3 being the last arc and π_a is the river, and $\pi_3 < \pi_a$, contradiction. □

Lemma 25. Let $x_1, \ldots, x_n$ be numbers that are at most m and the average is at least a. Then for any $t < a$ the number of numbers that is at least t is at least $\frac{a-t}{m-a} \cdot n$.

Proof. We apply Markov's inequality to the random variable $m - x$ where x is sampled uniformly from $x_1, \ldots, x_n$. This gives

$$\Pr[m - x \geq m - a] \geq \frac{a-t}{m-a},$$

so the number of numbers that is at least t must be at least $\frac{a-t}{m-a} \cdot n$. □

Lemma 26 (Rephrasing of [8], Lemma 23)**.** $|R| \geq \Omega\left(p\ell^2 d^2\right)$

Proof. For a vertex u let

$$f(u) := \sum_{\pi \ni u} |\{y \in \pi : u >_\pi y\}|,$$

e.g. the number of vertices that strictly precede u in some path. Note that

$$\frac{1}{n}\sum_{u \in V} f(u) = \frac{1}{n}\sum_{\pi \in \Pi} \binom{|\pi|}{2} \geq \frac{1}{n} p \binom{\ell/4}{2} \geq \frac{1}{n} \cdot \frac{p\ell^2}{33} = \frac{d\ell}{33}$$

while for each u we have $f(u) \leq 4d \cdot 4\ell = 16d\ell$. Thus by Lemma 25 there is at least $\frac{n}{1054} = \Omega(n)$ vertices u such that $f(u) \geq \frac{d\ell}{66}$. Call such a vertex good.

Now consider any good vertex u. Let d_u be degree of u. For each $\pi \ni u$ define $f(u, \pi)$ to be the number of vertex that strictly precede u in π. Then we have

$$\frac{1}{d_u}\sum_{\pi \ni u} f(u, \pi) = \frac{1}{d_u} f(u) \geq \frac{d\ell}{66 d_u} \geq \frac{\ell}{132}$$

while for each π we have $f(u, \pi) \leq 2\ell$. So by Lemma 25 there is at least $\Omega(d_u) = \Omega(d)$ path $\pi \ni u$ such that $f(u, \pi) \geq \frac{\ell}{200}$. Call such a path good. Let d_g be the number of good paths (through u.)

Now consider any pair π_2, π_3 of good path. Note that $\pi_2 < \pi_3$ or $\pi_3 < \pi_2$ by assumption. Thus there are at least $\binom{d_g}{2} = \Omega(d^2)$ pair of good path $\pi_3 < \pi_2$ for each good vertex u.

Now for a good pair $\pi_3 < \pi_2$, consider all the vertices preceding u in π_2, there are at least $\frac{\ell}{36} = \Omega(\ell)$ such vertices by the definition of good path. Fix such a vertex v. Each such v has $\Omega(d)$ path π_1 passing through them. As we goes through $\Omega(n)$ good vertex u, each with $\Omega(d^2)$ pair of good path $\pi_3 < \pi_2$ passing through u, and each with $\Omega(\ell)$ vertex v such that $v <_{\pi_2} u$, each with $\Omega(d)$ path π_1 passing through v, we have at least $\Omega(nd^2\ell d) = \Omega(p\ell^2 d^2)$ elements in the set R. □

Lemma 27 (Initial bound)**.** We have $H(n, p, 4) \leq O(n + n^{2/3}p^{2/3} + p)$.

Proof. We have

$$\Omega(p\ell^2 d^2) \leq |R| \leq \sum_{\pi_1, \pi_3 \in \Pi} |R(\pi_1, \pi_3)| \leq O(p^2 \ell).$$

Rearranging gives $\|Z\| \leq n^{2/3}p^{2/3}$. □

A.3 Bootstrapping

Our strategy from now on will be the following. Starting with the initial bound above, we will recursively improve the bound. For the recursive improvement step, we sample a path, called the *base path* uniformly at random from Π. For each of the $\Theta(\ell)$ vertices on the base path, consider the $\Theta(d)$ paths that pass through it. Then for each such path consider the h vertices that immediately follow the vertex on the base path, where $h \leq \ell$ is a parameter that will be chosen later (if there are few than h vertices following the base path vertex, consider all of them.) Each such sequence of at most $O(h)$ vertices is called a branching path. Thus at most $O(\ell dh)$ vertices are considered, and consider the subsystem induced by these vertices, which we shall call the h-system induced by the base path. Let Q denotes the set of tuple (π_1, π_3, u, v) such that π_1, π_3 are branching paths, $u \in \pi_1, v \in \pi_3$, $u \notin \pi_b$ (π_b is the base path), and $u <_{\pi_2} v$ for some other path π_2, and $\pi_3 < \pi_2$. Then we will argue that the structural property of the graph force $|Q|$ to be high in expectation, that is, we prove a lower bound of Q in term of ℓ, d, n, p, h. Then, we shall use the current bound for $H(n, p, 4)$, applied appropriately on parts of the h-system, to prove an upper bound for $|Q|$. Comparing the upper and lower bound for $|Q|$ would lead to a better bound for $H(n, p, 4)$, which converges to the final bound claimed.

Let us start with the lower bound for Q.

Lemma 28 (c.f. [8], Lemma 24)**.** If $\ell \geq h \geq \frac{Cp}{\ell d^2}$ for some large enough constant C, we have

$$\mathbb{E}[|Q|] \geq \Omega\left(\frac{h}{\ell p} \sum_{\pi_1, \pi_3 \in \Pi} |R(\pi_1, \pi_3)|^2\right).$$

Proof. For any vertices a and path π such that $a \in \pi$, let $\pi[a]$ denotes the number of vertices in π that is weakly before a. Fix a pair π_1, π_3. Given π_a, π_b in $R(\pi_1, \pi_3)$, we say π_a is close behind π_b if $0 < \pi_1[\pi_a \cap \pi_1] - \pi_1[\pi_b \cap \pi_1] \leq h$ and $0 \leq \pi_3[\pi_a \cap \pi_3] - \pi_3[\pi_a \cap \pi_3] \leq h$. The point of this definition is that, if π_b is pick as the base path, then $(\pi_1, \pi_3, \pi_a \cap \pi_1, \pi_a \cap \pi_3)$ is a tuple in Q, in which case we say π_a is charged to the pair (π_1, π_3).

We first show that if $h \geq 18\frac{\ell}{|R(\pi_1, \pi_3|}$ then the expected number of paths charged to (π_1, π_3) is at least

$$\frac{h}{36\ell p}|R(\pi_1, \pi_3)|^2.$$

For each $0 < i \leq z := |\pi_1| + |\pi_3| + h$, let a_i denotes the number of pair (π_2, j) such that $\pi_2 \in R(\pi_1, \pi_3)$ and $0 < j \leq h$ such that

$$i = \pi_1[\pi_1 \cap \pi_2] + \pi_3[\pi_2 \cap \pi_3] + j.$$

Then we have

$$\sum_{i=1}^{z} a_i = h|R(\pi_1, \pi_3)|$$

since there are $|R(\pi_1, \pi_3)|$ ways to chose π_2 and h ways to choose j. Note that

$$2z \leq 2(4\ell + 4\ell + \ell) = 18\ell \leq h\,|R(\pi_1, \pi_3)|\,.$$

It follows that

$$\begin{aligned}
\sum_{i=1}^{z}\binom{a_i}{2} &= \frac{1}{2}\left(\sum_{i=1}^{z}a_i^2-\sum_{i=1}^{z}a_i\right) \geq \frac{1}{2}\left(\sum_{i=1}^{z}a_i^2-\sum_{i=1}^{z}a_i\right)\\
&\geq \frac{1}{2}\left(\frac{1}{z}\left(\sum_{i=1}^{z}a_i\right)^2-\sum_{i=1}^{z}a_i\right) && \text{Cauchy -Schwarz}\\
&= \frac{1}{2}\left(\frac{h^2|R(\pi_1,\pi_3)|^2}{z}-h|R(\pi_1,\pi_3)|\right)\\
&\geq \frac{h^2|R(\pi_1,\pi_3)|^2}{4z} && \text{since } h|R(\pi_1,\pi_3)|\geq 2z. \quad \square
\end{aligned}$$

Note that $\sum_{i=1}^{z}\binom{a_i}{2}$ is the number of unordered pair of distinct tuple (π_a, j_a) and (π_b, j_b) such that

$$\pi_1[\pi_1\cap\pi_a]+\pi_3[\pi_a\cap\pi_3]+j_a=\pi_1[\pi_1\cap\pi_b]+\pi_3[\pi_b\cap\pi_3]+j_b.$$

Call such a pair of tuples an *aligned pair*. It is clear that if $\pi_a=\pi_b$ then this would implies $j_a=j_b$, which contradicts the fact that these are distinct tuples, so we have $\pi_a\neq\pi_b$. Since we count the number of unordered pair of distinct tuple (π_a, j_a) and (π_b, j_b), we may assume $\pi_b <_R \pi_a$, where $<_R$ is the ordering as in Lemma 24. Note that then we have

$$0<\pi_1[\pi_1\cap\pi_a]-\pi_1[\pi_1\cap\pi_b], 0\leq\pi_3[\pi_3\cap\pi_a]-\pi_3[\pi_3\cap\pi_b]$$

and

$$(\pi_1[\pi_1\cap\pi_a]-\pi_1[\pi_1\cap\pi_b])+(\pi_3[\pi_3\cap\pi_a]-\pi_3[\pi_3\cap\pi_b])=j_b-j_a\leq h. \tag{1}$$

so this implies that π_a is close behind π_b. Conversely, if π_a is close behind π_b then from (1), the number of way to pick corresponding j_b, j_a such that $(\pi_a, j_a), (\pi_b, j_b)$ forms an aligned pair is at most h since we need j_b-j_a to have a specific value. Thus the number of close behind pairs π_a, π_b is at least $\frac{1}{h}$ the number of aligned pairs, which is at least

$$\frac{1}{h}\left(\frac{h^2|R(\pi_1,\pi_3)|^2}{4}z\right)=\frac{h|R(\pi_1,\pi_3|^2}{4z}\geq\frac{h|R(\pi_1,\pi_3)|^2}{36\ell}.$$

Thus the expected number of path charged to (π_1, π_3) when sampling a base path at random is at least

$$\frac{h|R(\pi_1,\pi_3)|^2}{36\ell p}.$$

Now the expected total number of elements in Q is at least

$$\begin{aligned}
\sum_{(\pi_1,\pi_3):|R(\pi_1,\pi_3)|\geq\frac{18\ell}{h}}\frac{h|R(\pi_1,\pi_3)|^2}{36\ell p} &= \sum_{\pi_1,\pi_3}\frac{h|R(\pi_1,\pi_3)|^2}{36\ell p}-\sum_{(\pi_1,\pi_3):|R(\pi_1,\pi_3)|<\frac{18\ell}{h}}\frac{h|R(\pi_1,\pi_3)|^2}{36\ell p}\\
&\geq \sum_{\pi_1,\pi_3}\frac{h|R(\pi_1,\pi_3)|^2}{36\ell p}-p^2\left(\frac{h(18\ell/h)^2}{36\ell p}\right)\\
&\geq \frac{h}{36\ell p}\left(\sum_{\pi_1,\pi_3}|R(\pi_1,\pi_3)|^2-400p^2\ell^2/h^2\right)
\end{aligned}$$

Note that by Cauchy-Schwarz we have

$$\sum_{\pi_1,\pi_3} |R(\pi_1,\pi_3)|^2 \geq \frac{1}{p^2}\left(\sum_{\pi_1,\pi_3} |R(\pi_1,\pi_3)|\right)^2 \geq \Omega\left(\frac{1}{p^2}\left(p\ell^2 d^2\right)^2\right) \geq \Omega\left(\ell^4 d^4\right)$$

and if $h \geq \frac{Cp}{\ell d^2}$ we have $400p^2\ell^2/h^2 \leq \frac{400}{C^2}\ell^4 d^4$. Thus by choosing C to be large enough, we have

$$\sum_{\pi_1,\pi_3} |R(\pi_1,\pi_3)|^2 - 400p^2\ell^2/h^2 \geq \Omega\left(\sum_{\pi_1,\pi_3} |R(\pi_1,\pi_3)|^2\right)$$

and thus

$$\mathbb{E}[|Q|] \geq \Omega\left(\frac{h}{\ell p}\sum_{\pi_1,\pi_3\in\Pi} |R(\pi_1,\pi_3)|^2\right).$$

We now start proving the upper bound for $|Q|$. Our strategy is as follows. Given a vertex u in the h system, let its *branching degree* be the number of branching path passing through u. We will split the vertices into $O(\log n)$ buckets, each with branching degree between x and $2x$ for some x. A Cauchy-Schwarz argument would show that, roughly speaking, at least $\frac{1}{O(\log n)}$ of elements in Q comes from pairs u, v in the same bucket. Note that if u, v is on a same path π_2 and have branching degree $\Theta(x)$, this will contributes x^2 elements to Q. The number of pairs of vertices on the same path is the sum of the square of path length of certain subsystem, so it is helpful to have an upper bound on the sum of the square path length. We would also need upper bounds on the number of vertices of branching degree x or higher, since larger degree vertices contributes more to Q. We first start by giving an upper bound for sum of the square of path length.

Lemma 29. Let $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$ be a decreasing sequence of nonnegative real numbers and $b_1, \ldots, b_n$ is a sequence of nonnegative real numbers such that

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i$$

for any $1 \leq k \leq n$, then we have

$$\sum_{i=1}^{n} a_i^2 \leq \sum_{i=1}^{n} b_i^2.$$

Proof. Without loss of generality, we may assume $b_1 \geq \cdots \geq b_n$, (otherwise we sort the b_i descendingly, it is clear that the assumption $\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i$ still hold with the sorted sequence.) Let $a_{n+1} = b_{n+1} = 0$. We have

$$\begin{aligned}
\sum_{i=1}^{n} b_i^2 - \sum_{i=1}^{n} a_i^2 &= \sum_{i=1}^{n} (b_i - a_i)(b_i + a_i) \\
&= \sum_{i=1}^{n} \left(\sum_{j=1}^{i} (b_j - a_j)\right)(b_i + a_i - b_{i+1} - a_{i+1}) \\
&\geq 0
\end{aligned}$$

where we used summation-by-part in the last equality above. □

Lemma 30 (Continuous version)**.** Let $a_1 \geq a_2 \geq \cdots \geq a_n \geq 0$ be a decreasing sequence of nonnegative real numbers and $f\colon [0,n] \to \mathbb{R}_{\geq 0}$ is a nonnegative function such that

$$\sum_{i=1}^{k} a_i \leq \int_0^k f$$

for any $1 \leq k \leq n$, then we have

$$\sum_{i=1}^{n} a_i^2 \leq \int_0^n f^2.$$

Proof. Let

$$b_i := \int_{i-1}^{i} f$$

then we have

$$\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i$$

for any $1 \leq k \leq n$, so by Lemma 29 we have

$$\sum_{i=1}^{n} a_i^2 \leq \sum_{i=1}^{n} b_i^2 = \sum_{i=1}^{n} \left(\int_{i-1}^{i} f\right)^2 \leq \sum_{i=1}^{n} \int_{i-1}^{i} f^2 = \int_0^n f^2$$

where the inequality $\left(\int_{i-1}^{i} f\right)^2 \leq \int_{i-1}^{i} f^2$ follows from Cauchy-Schwarz (applied with f and the constant function 1.) □

Lemma 31. Let S' be a half-4-bridge free system on at most n_1 vertices and average degree at most d, and max length at most ℓ. Suppose that we have

$$H(n,p,4) \leq O\left(n + n^a p^{2-2a} + n^{2-2b} p^b + p\right)$$

with $\frac{2}{3} \leq b < 0.701$ and $\frac{8}{11} - 0.001 \leq a < \frac{3}{4}$. Let $\|S'\|_2^2$ denotes the sum of the path length squared. Then

$$\|S'\|_2^2 \leq O\left(n_1(\ell + d) + n_1^{3/2} d^{\frac{3-4a}{2-2a}} + n_1^{\frac{1}{b}} d^{2-\frac{1}{b}}\right).$$

Proof. Let $a_1 \geq \cdots \geq a_p$ be the length of the paths in the system. We have $\|S'\| \leq O(n_1 d)$. Then for each $1 \leq k \leq p$ we have

$$\sum_{i=1}^{k} a_i \leq C \min(k\ell, \max(n_1, n_1^a k^{2-2a}, n_1^{2-2b} k^b, k), n_1 d)$$

for some absolute constant C, since the longest k path forms a half-4-bridge free system. Thus let f be the derivative of the above function with respect to k, which, through some straightforward

calculus (omitted), turns out to be

$$f(x) = \begin{cases} C\ell, & 0 \le x \le \min\left(\frac{n_1}{\ell}, \frac{n_1^{\frac{a}{2a-1}}}{\ell^{\frac{1}{2a-1}}}\right), \\ 0, & \min\left(\frac{n_1}{\ell}, \frac{n_1^{\frac{a}{2a-1}}}{\ell^{\frac{1}{2a-1}}}\right) \le x \le \min\left(\sqrt{n_1}, \frac{n_1^{\frac{a}{2a-1}}}{\ell^{\frac{1}{2a-1}}}\right), \\ (2-2a)Cn_1^a x^{1-2a}, & \min\left(\sqrt{n_1}, \frac{n_1^{\frac{a}{2a-1}}}{\ell^{\frac{1}{2a-1}}}\right) \le x \le \min\left(n_1^{\frac{2b+a-2}{2a+b-2}}, n_1^{1/2} d^{\frac{1}{2-2a}}\right), \\ bCn_1^{2-2b}x^{b-1}, & \min\left(n_1^{\frac{2b+a-2}{2a+b-2}}, n_1^{1/2} d^{\frac{1}{2-2a}}\right) \le x \le \min\left(n_1^{2-\frac{1}{b}} d^{\frac{1}{b}}, n_1^{1/2} d^{\frac{1}{2-2a}}, n_1^2\right), \\ 1, & \min\left(n_1^{2-\frac{1}{b}} d^{\frac{1}{b}}, n_1^{1/2} d^{\frac{1}{2-2a}}, n_1^2\right) \le x \le \min\left(n_1^{2-\frac{1}{b}} d^{\frac{1}{b}}, n_1^{1/2} d^{\frac{1}{2-2a}}, n_1 d, p\right), \\ 0, & \min\left(n_1^{2-\frac{1}{b}} d^{\frac{1}{b}}, n_1^{1/2} d^{\frac{1}{2-2a}}, n_1 d, p\right) \le x \le p, \end{cases}$$

we have

$$\sum_{i=1}^{k} a_i \le \int_0^k f.$$

Thus

$$\sum_{i=1}^{p} a_i^2 \le \int_0^p f^2 \le O\left(n_1(\ell + d) + n_1^{3/2} d^{\frac{3-4a}{2-2a}} + n_1^{\frac{1}{b}} d^{2-\frac{1}{b}}\right),$$

where we omitted some straightforward calculus calculations. □

We now start proving bounds on the number of vertices with branching degree $\Omega(x)$.

Lemma 32. Fix a vertex v. Consider h previous vertices in the each of the $\Theta(d)$ path passing through v. Suppose that we have

$$H(n, p, 4) \le O(n + n^a p^{2-2a} + n^{2-2b} p^b + p)$$

with $\frac{2}{3} \le b < 0.701$ and $\frac{8}{11} - 0.001 \le a < \frac{3}{4}$. Consider the induced system on those $O(dh)$ vertices. Then the number of path of length at least $\Omega(x)$ in the subsystem is at most

$$O\left(\min\{\frac{dh}{x} + \frac{(dh)^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{(dh)^2}{x^{\frac{1}{1-b}}}, \frac{d^2 h}{x}\}\right).$$

Proof. Let p_x be the number of paths. Note that there are $O(dh)$ with degree $O(d)$ so we have $xp_x \le (dh)d$ so

$$p_x \le O\left(\frac{d^2 h}{x}\right).$$

If x is at most a constant we are done since this term is smaller in the min. Otherwise we have

$$p_x x \le O\left(dh + (dh)^a p_x^{2-2a} + (dh)^{2-2b} p_x^b\right).$$

Rearranging algebraically gives

$$p_x \le O\left(\frac{dh}{x} + \frac{(dh)^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{(dh)^2}{x^{\frac{1}{1-b}}}\right).$$ □

Lemma 33. Let n_x denotes the number of vertices in the subsystem that has branching degree $\Omega(x)$, not including those on the base path itself. Suppose that we have

$$H(n,p,4) \leq O\left(n + n^a p^{2-2a} + n^{2-2b}p^b + p\right)$$

with $\frac{2}{3} \leq b < 0.701$ and $\frac{8}{11} - 0.001 \leq a < \frac{3}{4}$. Then we have

$$\mathbb{E}[n_x] \leq O\left(\frac{\ell h}{x} + \frac{\ell d^{\frac{1-a}{2a-1}} h^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{\ell d h^2}{x^{\frac{1}{1-b}}}\right),$$

and $n_x \leq O(\frac{\ell dh}{x})$ deterministically.

Proof. There are $O(\ell d)$ branching paths, so the number of vertices on these path couting with multiplicity is at most $O(\ell dh)$. Thus the number of vertices that is on $\Omega(x)$ of the branching paths is at most $O\left(\frac{\ell dh}{x}\right)$ deterministically. By Lemma 32, there are at most

$$O\left(\frac{ndh}{x} + \frac{n(dh)^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{n(dh)^2}{x^{\frac{1}{1-b}}}\right)$$

pairs of vertex v and path π_b such that v is in at least $\Omega(x)$ of the branching path if π_b was pick as a base path, and $v \notin \pi_b$. Thus the expected number of vertices which lies in at least $\Omega(x)$ branching path when a random base path is chosen is at most

$$O\left(\frac{ndh}{px} + \frac{n(dh)^{\frac{a}{2a-1}}}{px^{\frac{1}{2a-1}}} + \frac{n(dh)^2}{px^{\frac{1}{1-b}}}\right) = O\left(\frac{\ell h}{x} + \frac{\ell d^{\frac{1-a}{2a-1}} h^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{\ell d h^2}{x^{\frac{1}{1-b}}}\right). \qquad \square$$

A.4 Completing the Proof

The previous lemma relates the value of $H(n,p,4)$ to n_x, which in turn relates to upper bounds on Q. Since an upper bound on Q implies an upper bound on $H(n,p,4)$ in turn, it is intuitive that this will imply *some* upper bound on $H(n,p,4)$. It requires some straightforward but tedious algebra to realize this upper bound.

We will next start combining the upper bound on the sum of square of path length, and on n_x, for upper bounds on $|Q|$. Note that, n_x only counts vertices that are not on the base path, but an element (π_1, π_3, u, v) could potentially have $v \in \pi_b$ (the base path.) Thus we have to deal with elements (π_1, π_3, u, v) of Q with $v \in \pi_b$ separately. Define

$$Q_1 := \{(\pi_1, \pi_3, u, v) \in Q : v \notin \pi_b\}$$

and

$$Q_2 := \{(\pi_1, \pi_3, u, v) \in Q : v \in \pi_b\}.$$

Lemma 34. Suppose that we have

$$H(n,p,4) \leq O\left(n + n^a p^{2-2a} + n^{2-2b}p^b + p\right)$$

with $\frac{2}{3} \leq b < 0.701$ and $\frac{8}{11} - 0.001 \leq a < \frac{3}{4}$. Suppose $h \leq \min(\ell, d)$. Then we have

$$\mathbb{E}[|Q_1|] \leq \tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}}\right)$$

Proof. Firstly, we consider the tuples (π_1, π_3, u, v) such that $v \notin \pi_b$, let the number of such tuple be Q_1. For each path π and x, let Q_π denotes the contribution of π to Q_1, which is the number of tuple (π_1, π_3, u, v) such that $u <_\pi v$ and π_1, π_3 are branching paths passing through u and v, and $u, v \notin \pi_b$, and $\pi_3 < \pi_2$. Let π_x denotes the number of vertices on π with branching degree between x and $2x$, not including vertices in the base graph if any. Then we have

$$\begin{aligned}
Q_1 = \sum_\pi Q_\pi &\le \sum_\pi \left(\sum_x x\pi_x\right)^2 \\
&\le O(\log n) \sum_\pi \sum_x (x\pi_x)^2 && \text{Cauchy-Schwarz} \\
&\le O(\log n) \sum_x x^2 \sum_\pi \pi_x^2 \\
&\le O(\log n) \sum_x x^2 \left(n_x \ell + n_x^{3/2} d^{\frac{3-4a}{2-2a}} + n_x^{\frac{1}{b}} d^{2-\frac{1}{b}}\right) && \text{Lemma 31}
\end{aligned}$$

where in all of the above, the sum over x runs over all powers of 2 from $x = \Theta(1)$ up to $x = \Theta(d)$. Thus

$$\mathbb{E}[Q_1] \le \tilde{O}\left(\sum_x x^2 \left(\mathbb{E}[n_x]\ell + \mathbb{E}[n_x^{3/2}] d^{\frac{3-4a}{2-2a}} + \mathbb{E}[n_x^{\frac{1}{b}}] d^{2-\frac{1}{b}}\right)\right).$$

Since the above sum has $O(\log n)$ term, it suffices to show that for any x we have

$$x^2 \left(\mathbb{E}[n_x](\ell + d) + \mathbb{E}[n_x^{3/2}] d^{\frac{3-4a}{2-2a}} + \mathbb{E}[n_x^{\frac{1}{b}}] d^{2-\frac{1}{b}}\right) \le O\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}}\right)$$

We consider the possible values of x.

- **Case 1:** $x \le O(h^{\frac{1-b}{b}})$**.**

  Note that $x = h^{\frac{1-b}{b}}$ is the threshold at which $\frac{\ell dh}{x} = \frac{\ell dh^2}{x^{\frac{1-b}{b}}}$. We have $n_x \le O\left(\frac{\ell dh}{x}\right)$ deterministically, so we have

$$\begin{aligned}
&x^2 \left(\mathbb{E}[n_x](\ell + d) + \mathbb{E}[n_x^{3/2}] d^{\frac{3-4a}{2-2a}} + \mathbb{E}[n_x^{\frac{1}{b}}] d^{2-\frac{1}{b}}\right) \\
&\le x^2 \left(\left(\frac{\ell dh}{x}\right)(\ell + d) + \left(\frac{\ell dh}{x}\right)^{3/2} d^{\frac{3-4a}{2-2a}} + \left(\frac{\ell dh}{x}\right)^{\frac{1}{b}} d^{2-\frac{1}{b}}\right) \\
&\le O\left((\ell + d)\ell dh x + (\ell dh)^{3/2} x^{1/2} d^{\frac{3-4a}{2-2a}} + (\ell dh)^{\frac{1}{b}} x^{2-\frac{1}{b}} d^{2-\frac{1}{b}}\right) \\
&\le O\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}}\right) && \text{since } x \le h^{\frac{1-b}{b}}.
\end{aligned}$$

- **Case 2:** $\Omega\left(h^{\frac{1-b}{b}}\right) \le x \le O\left((dh)^{\frac{(3a-2)(1-b)}{2a+b-2}}\right)$**.**

  In this case, $\mathbb{E}[n_x] \le O\left(\frac{\ell dh^2}{x^{\frac{1}{1-b}}}\right)$. We still have $n_x \le O\left(\frac{\ell dh}{x}\right)$ deterministically. Hence

$$\mathbb{E}[n_x^t] \le O\left(\mathbb{E}[n_x]\left(\frac{\ell dh}{x}\right)^{t-1}\right) \le O\left(\frac{\ell dh^2}{x^{\frac{1}{1-b}}}\left(\frac{\ell dh}{x}\right)^{t-1}\right) = O\left(\frac{(\ell d)^t h^{t+1}}{x^{\frac{b}{1-b}+t}}\right)$$

for any constant $t \geq 1$. Thus we have

$$\begin{aligned}
&x^2 \left(\mathbb{E}[n_x](\ell + d) + \mathbb{E}[n_x^{3/2}] d^{\frac{3-4a}{2-2a}} + \mathbb{E}[n_x^{\frac{1}{b}}] d^{2-\frac{1}{b}} \right) \\
&\leq x^2 \left(\left(\frac{\ell d h^2}{x^{\frac{1}{1-b}}} \right) (\ell + d) + \left(\frac{(\ell d)^{3/2} h^{5/2}}{x^{\frac{b}{1-b}+\frac{3}{2}}} \right) d^{\frac{3-4a}{2-2a}} + \left(\frac{(\ell d)^{\frac{1}{b}} h^{\frac{b+1}{b}}}{x^{\frac{b}{1-b}+\frac{1}{b}}} \right) d^{2-\frac{1}{b}} \right) \\
&\leq O\left((\ell + d) \ell d h^2 x^{-\frac{2b-1}{1-b}} + (\ell d)^{3/2} h^{5/2} x^{-\frac{3b-1}{2(1-b)}} d^{\frac{3-4a}{2-2a}} + (\ell d)^{\frac{1}{b}} h^{\frac{b+1}{b}} x^{-\frac{3b^2-3b+1}{b(1-b)}} d^{2-\frac{1}{b}} \right) \\
&\leq O\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}} \right) \qquad \text{since } x \geq h^{\frac{1-b}{b}}.
\end{aligned}$$

Remark. It is not a coincidence that the bound for these first two cases is the same. In the first case, we obtained an upper bound which turns out to be an increasing function in x: namely, the function

$$f_1(x) := (\ell + d) \ell d h x + (\ell d h)^{3/2} x^{1/2} d^{\frac{3-4a}{2-2a}} + (\ell d h)^{\frac{1}{b}} x^{2-\frac{1}{b}} d^{2-\frac{1}{b}}.$$

Thus the upper bound is obtained from plugging in the largest value of x, which is $h^{\frac{1-b}{b}}$. Meanwhile, in the second case, we obtained an upper bound which turns out to be an decreasing function in x: namely, the function

$$f_2(x) := (\ell + d) \ell d h^2 x^{-\frac{2b-1}{1-b}} + (\ell d)^{3/2} h^{5/2} x^{-\frac{3b-1}{2(1-b)}} d^{\frac{3-4a}{2-2a}} + (\ell d)^{\frac{1}{b}} h^{\frac{b+1}{b}} x^{-\frac{3b^2-3b+1}{b(1-b)}} d^{2-\frac{1}{b}}.$$

Thus the upper bound is when we plug in the smallest value of x which is $h^{\frac{1-b}{b}}$. Notice that these two f_1, f_2 functions are obtained by plugging in appropriate upper bound for $\mathbb{E}[n_x], \mathbb{E}[n_x^{3/2}], \mathbb{E}[n_x^{\frac{1}{b}}]$. Since $x = h^{\frac{1-b}{b}}$ is the threshold value at which $\frac{\ell d h}{x} = \frac{\ell d h^2}{x^{\frac{1-b}{b}}}$, the appropriate upper bound for $\mathbb{E}[n_x], \mathbb{E}[n_x^{3/2}], \mathbb{E}[n_x^{\frac{1}{b}}]$ in these two cases would be the same when $x = h^{\frac{1-b}{b}}$, which means that $f_1(x) = f_2(x)$ when $x = h^{\frac{1-b}{b}}$, and thus the two upper bounds agree.

- **Case 3:** $\Omega\left((dh)^{\frac{(3a-2)(1-b)}{2a+b-2}} \right) \leq x \leq O\left((dh)^{\frac{1}{2}} \right)$.

  In this case,

$$\mathbb{E}[n_x] \leq O\left(\frac{\ell d^{\frac{1-a}{2a-1}} h^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} \right).$$

  We still have $n_x \leq O\left(\frac{\ell d h}{x}\right)$ deterministically. Hence

$$\mathbb{E}[n_x^t] \leq O\left(\mathbb{E}[n_x] \left(\frac{\ell d h}{x} \right)^{t-1} \right) \leq O\left(\left(\frac{\ell d^{\frac{1-a}{2a-1}} h^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} \right) \left(\frac{\ell d h}{x} \right)^{t-1} \right)$$

  We can then give an upper bound for the value of

$$x^2 \left(\mathbb{E}[n_x] \ell + \mathbb{E}[n_x^{3/2}] d^{\frac{3-4a}{2-2a}} + \mathbb{E}[n_x^{\frac{1}{b}}] d^{2-\frac{1}{b}} \right)$$

by plugging in the upper bound for $\mathbb{E}[n_x^t]$ as done in the previous case, and thus would obtain an upper bound as some function $f_3(x)$. To simplify calculations, we will just compute the power of x in the terms here. Note that $E[n_x^t]$ is bounded by a term whose power of x is

$$-\frac{1}{2a-1}-(t-1)$$

Then the power of x in the first, second and third term respectively would be

$$2-\frac{1}{2a-1}-(t-1)$$

for $t=1,\frac{3}{2},\frac{1}{b}$ respectively. Note that

$$2-\frac{1}{2a-1}-(t-1)\leq 2-\frac{1}{2a-1}<0$$

since $a<\frac{3}{4}$ and $t\geq 1$, so the function f_3 is decreasing in x. So we have $f_3(x)$ is largest when

$$x=\Theta\left((dh)^{\frac{(3a-2)(1-b)}{2a+b-2}}\right).$$

By similar reasoning to last case, we have that $f_3(x)=f_2(x)$ when

$$x=\Theta\left((dh)^{\frac{(3a-2)(1-b)}{2a+b-2}}\right).$$

And we proved

$$f_2(x)\leq f_2(h^{\frac{1-b}{b}})\leq O\left(\ell^2 dh^{\frac{1}{b}}+\ell^{\frac{3}{2}}d^{\frac{6-7a}{2-2a}}h^{\frac{2b+1}{2b}}+\ell^{\frac{1}{b}}d^2h^{\frac{4b-2b^2-1}{b^2}}\right)$$

so we have

$$f_3(x)\leq O\left(\ell^2 dh^{\frac{1}{b}}+\ell^{\frac{3}{2}}d^{\frac{6-7a}{2-2a}}h^{\frac{2b+1}{2b}}+\ell^{\frac{1}{b}}d^2h^{\frac{4b-2b^2-1}{b^2}}\right).$$

- **Case 4:** $\Omega\left((dh)^{\frac{1}{2}}\right)\leq x\leq O(d)$**.**

  In this case we have $\mathbb{E}[n_x]\leq O(\frac{\ell h}{x})$, and $n_x\leq O\left(\frac{\ell dh}{x}\right)$. Thus

$$\mathbb{E}[n_x^t]\leq O\left(\left(\frac{\ell h}{x}\right)\left(\frac{\ell dh}{x}\right)^{t-1}\right)=O\left(\ell^t d^{t-1}h^t x^{-t}\right).$$

  Thus

$$\begin{aligned}
&x^2\left(\mathbb{E}[n_x](\ell+d)+\mathbb{E}[n_x^{3/2}]d^{\frac{3-4a}{2-2a}}+\mathbb{E}[n_x^{\frac{1}{b}}]d^{2-\frac{1}{b}}\right)\\
&\leq O\left((\ell+d)\ell hx+\ell^{3/2}d^{1/2+\frac{3-4a}{2-2a}}h^{3/2}x^{1/2}+\ell^{\frac{1}{b}}dh^{\frac{1}{b}}x^{2-\frac{1}{b}}\right)\\
&\leq O\left((\ell+d)\ell dh+\ell^{3/2}d^{\frac{5-6a}{2-2a}}h^{3/2}+\ell^{\frac{1}{b}}d^{3-\frac{1}{b}}h^{\frac{1}{b}}\right) && \text{since } x\leq d\\
&\leq O\left(\ell^2 dh^{\frac{1}{b}}+\ell^{\frac{3}{2}}d^{\frac{6-7a}{2-2a}}h^{\frac{2b+1}{2b}}+\ell^{\frac{1}{b}}d^2h^{\frac{4b-2b^2-1}{b^2}}\right). && \square
\end{aligned}$$

Lemma 35. Suppose that we have $H(n,p,4) \leq O(n + n^a p^{2-2a} + n^{2-2b}p^b + p)$ with $\frac{2}{3} \leq b < 0.701$ and $\frac{8}{11} - 0.001 \leq a < \frac{3}{4}$. Suppose $h \leq \min(\ell, d)$. Then we have

$$\mathbb{E}[|Q_2|] \leq \tilde{O}\left(\ell^{3/2}d^2h^{5/4}\right) + O(\ell^2 d^2)$$

Proof. Let Q_{2x} denotes the set of tuple in Q_2 such that u has branching degree between x and $2x$

Now, fix a vertex u, we show an upper bound on the number of pairs (π_b, v) such that the h-system with base path π_b has $\Theta(x)$ branching path passing through u, and $v \in \pi_b$, and v follows u in some path - call such a pair *important*. Let p_x denotes the number of paths with at least x vertices in the $O(dh)$ vertices that precedes u no further than h away in some path. Consider the system T induced by taking the paths to be these p_x path, and taking the vertices to be the union of the $O(d\ell)$ vertices that follow u in some path and the $O(dh)$ vertices that precedes u no further than h away in some path. Note that the number of important pairs (π_b, v) is no more than $\|T\|$. Note that T is source-restricted into the $O(dh)$ vertices preceding u. This is because if the source of a path π_b was a vertex v that follows u in some path π instead of preceding u, let v' be one of the $\Theta(x)$ vertex on π_b that precede u in some path π', then π_b, π', π forms a 3-cycle. Also, T has at most p_x paths and at most $O(d\ell)$ vertices. Thus by the bound on the path system in the proof of Theorem 11 we have

$$\|T\| \leq O\left(\ell d + ((\ell d)(dh)p_x)^{1/2}\right) = O\left(\ell d + \ell^{1/2}dh^{1/2}p_x^{1/2}\right)$$

Recall from Lemma 32 that we have

$$p_x \leq O\left(\min\left\{\frac{dh}{x} + \frac{(dh)^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{(dh)^2}{x^{\frac{1}{1-b}}}, \frac{d^2h}{x}\right\}\right)$$

and we now use a crude bound

$$p_x \leq O\left(\min\left\{\frac{dh}{x} + \frac{(dh)^{\frac{a}{2a-1}}}{x^{\frac{1}{2a-1}}} + \frac{(dh)^2}{x^{\frac{1}{1-b}}}, \frac{d^2h}{x}\right\}\right) \leq O\left(\min\left\{\frac{dh}{x} + \frac{(dh)^2}{x^3}, \frac{d^2h}{x}\right\}\right)$$

Thus

$$\|T\| \leq O\left(\ell d + \min\left\{\ell^{1/2}d^{3/2}hx^{-1/2} + \ell^{1/2}d^2h^{3/2}x^{-3/2}, \ell^{1/2}d^2hx^{-1/2}\right\}\right)$$

We then have that the number of triples (u, π_b, v) such that u has branching degree $\Theta(x)$ when π_b is selected as the base path and $v \in \pi_b$ and v follows u in some path is at most n times the above bound, and the note that such a triple (u, π_b, v) contributes $\Theta(xd)$ to the size of Q_{2x} with probability $\frac{1}{p}$. Thus we have

$$\begin{aligned}\mathbb{E}[|Q_{2x}|] &\leq O\left(\frac{nxd}{p}\left(\ell d + \min\left\{\ell^{1/2}d^{3/2}hx^{-1/2} + \ell^{1/2}d^2h^{3/2}x^{-3/2}, \ell^{1/2}d^2hx^{-1/2}\right\}\right)\right) \\ &\leq O\left(\ell^2 dx + \min\left\{\ell^{3/2}d^{3/2}hx^{1/2} + \ell^{3/2}d^2h^{3/2}x^{-1/2}, \ell^{3/2}d^2hx^{1/2}\right\}\right).\end{aligned}$$

Note that

$$\min\left\{\ell^{3/2}d^{3/2}hx^{1/2} + \ell^{3/2}d^2h^{3/2}x^{-1/2}, \ell^{3/2}d^2hx^{1/2}\right\} \leq \ell^{3/2}d^2h^{5/4}$$

for all $x \leq d$, by some omitted casework. Thus we have

$$\mathbb{E}[|Q_2|] \leq \sum_x O(\ell^2 dx) + \sum_x O\left(\ell^{3/2}d^2h^{5/4}\right) = O\left(\ell^2 d^2\right) + \tilde{O}\left(\ell^{3/2}d^2h^{5/4}\right).$$

where the sum runs over all powers of 2, and the $\sum_x O(\ell^2 dx)$ term is a geometric series that is dominated by the largest term, which is when $x = \Theta(d)$. □

Lemma 36. Suppose that we have $H(n,p,4) \le O(n + n^a p^{2-2a} + n^{2-2b}p^b + p)$ with $\frac{2}{3} \le b < 0.701$ and $\frac{8}{11} - 0.001 \le a < \frac{3}{4}$. Suppose $h \le \min(\ell, d)$. Then we have

$$\mathbb{E}[|Q|] \le \tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}} + \ell^{3/2} d^2 h^{5/4}\right) + O(\ell^2 d^2).$$

Proof. We have

$$\begin{aligned}
&\mathbb{E}[|Q|] = \mathbb{E}[|Q_1|] + \mathbb{E}[|Q_2|] \\
&\le \tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}}\right) + O(\ell^2 d^2) + \tilde{O}\left(\ell^{3/2} d^2 h^{5/4}\right) \\
&= \tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}} + \ell^{3/2} d^2 h^{5/4}\right) + O(\ell^2 d^2). \qquad \square
\end{aligned}$$

Lemma 37. Suppose that we have

$$H(n,p,4) \le O(n + n^a p^{2-2a} + n^{2-2b}p^b + p)$$

with $\frac{2}{3} \le b < 0.701$ and $\frac{8}{11} - 0.001 \le a < \frac{3}{4}$. Then

$$H(n,p,4) \le \tilde{O}\left(n^{\frac{2+b}{3+b}} p^{\frac{2}{3+b}} + n^{\frac{8b-4b^2-2}{11b-4b^2-3}} p^{\frac{7b-2b^2-2}{11b-4b^2-3}} + \ell^{3/2} d^2 h^{5/4}\right) + O(n+p).$$

Proof. We set $h = \frac{Cp}{\ell d^2}$ where C is a large enough constant. We can assume $h \le \min(\ell, d)$, since otherwise through some simple calculations (omitted) we immediately get $\|Z\| \le O\left(n^{3/4}p^{1/2} + n^{1/2}p^{3/4}\right)$ which is better than the above bound. Then we have

$$\begin{aligned}
&\tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}} + \ell^{3/2} d^2 h^{5/4}\right) + O(\ell^2 d^2) \\
&\ge \mathbb{E}[Q] && \text{Lemma 36} \\
&\ge \Omega\left(\frac{h}{\ell p} \sum_{\pi_1, \pi_3 \in \Pi} |R(\pi_1, \pi_3)|^2\right) && \text{Lemma 28} \\
&\ge \Omega\left(\frac{h}{\ell p^3} \left(\sum_{\pi_1, \pi_3 \in \Pi} |R(\pi_1, \pi_3)|\right)^2\right) && \text{Cauchy-Schwarz} \\
&\ge \Omega\left(\frac{h}{\ell p^3} \left(p\ell^2 d^2\right)^2\right) && \text{Lemma 26} \\
&= \Omega(\ell^2 d^2).
\end{aligned}$$

By choosing C to be large enough, the constant inside $\Omega(\ell^2 d^2)$ is larger than the constant inside $O(\ell^2 d^2)$, so we must have

$$\tilde{O}\left(\ell^2 d h^{\frac{1}{b}} + \ell^{\frac{3}{2}} d^{\frac{6-7a}{2-2a}} h^{\frac{2b+1}{2b}} + \ell^{3/2} d^2 h^{5/4} + \ell^{\frac{1}{b}} d^2 h^{\frac{4b-2b^2-1}{b^2}} + \ell^{3/2} d^2 h^{5/4}\right) \ge \Omega\left(\ell^2 d^2\right).$$

Rearranging, and applying the identity $\|Z\| = nd = p\ell$, we get

$$\|Z\| \le \tilde{O}\left(n^{\frac{2+b}{3+b}} p^{\frac{2}{3+b}} + n^{\frac{8b-4b^2-2}{11b-4b^2-3}} p^{\frac{7b-2b^2-2}{11b-4b^2-3}}\right).$$

(Note: this bound removes some terms that do not end up dominating the sum.) $\square$

Theorem 38. Let $\alpha \approx 0.7009$ be a root of $4x^3 - 13x^2 + 10x - 2$. Then

$$H(n,p,4) \leq O\left(n + n^{\frac{2+\alpha}{3+\alpha}+o(1)}p^{\frac{2}{3+\alpha}} + n^{2-2\alpha+o(1)}p^{\alpha} + p\right).$$

Consider the sequences $a_0 = \frac{8}{11}, b_0 = \frac{2}{3}$, $a_{i+1} = g(b_i)$ and $b_{i+1} = f(b_i)$ where

$$g(x) = \frac{2+x}{3+x}, f(x) := \frac{7x - x^2 - 2}{11x - 4x^2 - 3}$$

Note that $\frac{8}{11} \leq a_0 < \frac{3}{4}$ and $\frac{2}{3} \leq b_0 \leq \alpha$. Since f is increasing and $f(x) > x$ for $\frac{2}{3} \leq b_0 \leq \alpha$ and $f(\alpha) = \alpha$, we have that $\frac{2}{3} < b_i < \alpha$ for all $i \geq 1$ and b_i is an increasing sequence that converges to α. It follows that $\frac{8}{11} \leq a_i < \frac{3}{4}$ for all a_i. Lemma 27 shows that we have

$$H(n,p,4) \leq O\left(n + n^{a_0}p^{2-2a_0} + n^{2-2b_0}p^{b_0} + p\right)$$

and so Lemma 37 shows that we have

$$H(n,p,4) \leq \tilde{O}\left(n^{a_1}p^{2-2a_1} + n^{2-2b_1}p^{b_1}\right) + O(n+p)$$

so for any $\varepsilon > 0$ we have

$$H(n,p,4) \leq O\left(n^{a_1-\varepsilon}p^{2-2a_1+2\varepsilon} + n^{2-2b_1+2\varepsilon}p^{b_1-\varepsilon} + n + p\right)$$

We will prove by induction on i that for any $\varepsilon > 0$ we have

$$H(n,p,4) \leq O\left(n^{a_i-\varepsilon}p^{2-2a_i+2\varepsilon} + n^{2-2b_i+2\varepsilon}p^{b_i-\varepsilon} + n + p\right)$$

Above we showed the claim is true for $i = 1$. Suppose the claim is true for i, we show it's true for $i+1$. Fix $\varepsilon > 0$. Note that $b_{i+1} = f(b_i)$ and f is continuous so there is $\delta > 0$ such that $f(x) > b_{i+1} - \varepsilon$ for $b_i - \delta < x < b_i$ and g is continuous so there is δ_1 such that $g(x) > a_{i+1} - \varepsilon$ for $b_i - \delta_1 < x < b_i$. Take $\delta' = \frac{1}{2}\min\{\delta, \delta_1, 0.001, b_i - \frac{2}{3}\}$ we have

$$H(n,p,4) \leq O\left(n^{a_i-\delta'}p^{2-2a_i+2\delta'} + n^{2-2b_i+2\delta'}p^{b_i-\delta'} + n + p\right)$$

where $\frac{8}{11} - 0.001 < a_i - \delta' < \frac{3}{4}$ and $\frac{2}{3} < b_i - \delta' < \alpha$. Thus by Lemma 37 we have

$$H(n,p,4) \leq O\left(n^{a'_{i+1}}p^{2-2a'_{i+1}} + n^{2-2b'_{i+1}}p^{b'_{i+1}} + n + p\right)$$

where $a'_{i+1} = g(b_i - \delta') > a_{i+1} - \varepsilon$ and $b'_{i+1} = f(b_i - \delta') > b_{i+1} - \varepsilon$, and so the claim is true for $i+1$. Finally, it suffices to show that for any $\varepsilon > 0$ we have

$$H(n,p,4) \leq O\left(n + n^{\frac{2+\alpha}{3+\alpha}+\varepsilon}p^{\frac{2}{3+\alpha}-2\varepsilon} + n^{2-2\alpha+2\varepsilon}p^{\alpha-\varepsilon} + p\right).$$

Fix $\varepsilon > 0$. Note that b_i converges to α so a_i converges to $\frac{2+\alpha}{3+\alpha}$ so there is i such that $a_i > \frac{2+\alpha}{3+\alpha} + \varepsilon/2$ and $b_i > \alpha - \varepsilon/2$. Then by the above claim we have

$$\begin{aligned} H(n,p,4) &\leq O\left(n^{a_i-\varepsilon/2}p^{2-2a_i+\varepsilon} + n^{2-2b_i+\varepsilon}p^{b_i-\varepsilon/2} + n + p\right) \\ &\leq O\left(n + n^{\frac{2+\alpha}{3+\alpha}-\varepsilon}p^{\frac{2}{3+\alpha}+2\varepsilon} + n^{2-2\alpha+2\varepsilon}p^{\alpha-\varepsilon} + p\right) \end{aligned}$$

as desired.

B Offline Pairwise Reachability Preservers

We next prove Theorem 4. See Appendix A.1 for a high-level view of the changes from [8] implicit in the following proof.

As in [8], we use $\beta(n,p,4)$ to denote the maximum size of a path system that is 4-bridge-free, where we say a path system is k-bridge-free if it has no bridge of size at most k. Furthermore, as shown in [8] (via the Independence Lemma, overviewed in Section 3.3 that in order to prove Theorem 4, it suffices to prove that

$$\beta(n,p,4) \leq O\left(n + n^{3/4}p^{1/2} + n^{2-\sqrt{2}+o(1)}p^{\frac{1}{\sqrt{2}}} + p\right).$$

We will use the same strategy as for bounding $H(n,p,4)$. First note that a similar cleaning lemma hold for bridge-free system.

Lemma 39 (Cleaning Lemma for bridge-free systems – c.f. [8], Lemma 10)**.** There exists a k-bridge free system on $\leq n$ vertices, $\leq p$ path whose size is $\Theta(H(n,p,k))$ such that every vertex has degree between $d/4$ and $4d$ and every path has length between $\ell/4$ and 4ℓ where d, ℓ are the average degree and average path length respectively.

The proof is almost identical to Lemma 23, except that we argue that the modified system remains bridge-free instead of half-bridge free, so we omit the proof.

We shall follow the same strategy with bounding $H(n,p,4)$. That is, we also use recursion and select the random h-system. Note that Lemma 24, Lemma 26, Lemma 27 and Lemma 28 holds in this setting as well, since it holds in a more general setting. In fact the total ordering of Lemma 24 has the following property: If $\pi_a <_R \pi_b$ then $\pi_a \cap \pi_1 <_{\pi_1} \pi_b \cap \pi_1$ and $\pi_a \cap \pi_3 <_{\pi_3} \pi_b \cap \pi_3$. This is because if $\pi_a \cap \pi_3 = \pi_b \cap \pi_3$, then π_1, π_b, π_a will form a 3 bridge.

As a result, when bounding $|Q|$, we no longer need to deal with the case an element (π_1, π_3, u, v) of Q has v in the base path. Also note that due to 3-bridge free, all nodes have branching degree 1.

We have the following version of Lemma 31.

Lemma 40. Let S' be a 4-bridge free system on at most $n_1 = \Theta(\ell d h)$ vertices and average degree at most d, and max length at most ℓ. Suppose that we have

$$\beta(n,p,4) \leq O\left(n + n^{3/4}p^{2-2a} + n^{2-2b}p^b + p\right)$$

with $\frac{2}{3} \leq b < \frac{3}{4}$. Then

$$\|S'\|_2^2 \leq O\left(n_1\ell + n_1^{3/2}\log(n_1) + n_1^{\frac{1}{b}}d^{2-\frac{1}{b}}\right)$$

Proof. The proof is similar to Lemma 31, so we will sketch it here. Let $a_1 \geq \cdots \geq a_p$ be the length of the paths in the system, and so we have

$$\sum_{i=1}^{k} a_i \leq C\min(k\ell, \max(n_1, n_1^{3/4}k^{1/2}, n_1^{2-2b}k^b, k), n_1 d)$$

for some absolute constant C. Letting f be the derivative of the above function with respect to k,

i.e

$$f(x)=\begin{cases} C\ell, & 0\le x\le \min\left(\frac{n_1}{\ell},\frac{n_1^{3/2}}{\ell^2}\right),\\ 0, & \min\left(\frac{n_1}{\ell},\frac{n_1^{3/2}}{\ell^2}\right)x\le \min\left(\sqrt{n_1},\frac{n_1^{3/2}}{\ell^2}\right),\\ \frac{1}{2}Cn_1^{3/4}x^{-1/2}, & \min\left(\sqrt{n_1},\frac{n_1^{3/2}}{\ell^2}\right)\le x\le \min\left(n_1^{\frac{8b-5}{4b-2}},n_1^{1/2}d^2\right),\\ bCn_1^{2-2b}x^{b-1}, & \min\left(n_1^{\frac{8b-5}{4b-2}},n_1^{1/2}d^2\right)\le x\le \min\left(n_1^{2-\frac{1}{b}}d^{\frac{1}{b}},n_1^{1/2}d^2\right),\\ 0, & \min\left(n_1^{2-\frac{1}{b}}d^{\frac{1}{b}},n_1^{1/2}d^2\right)\le x\le p, \end{cases}$$

we have

$$\sum_{i=1}^{k}a_i\le\int_0^k f$$

. Thus we have

$$\begin{aligned}\sum_{i=1}^{p}a_i^2&\le\int_0^p f^2\\&\le O\left(n_1\ell+n_1^{3/2}\log n_1+n_1^{\frac{1}{b}}d^{2-\frac{1}{b}}\right).\end{aligned}$$

□

Note that in this case we directly have an upper bound for $|Q|$.

Lemma 41. Suppose that we have

$$\beta(n,p,4)\le O\left(n+n^{3/4}p^{2-2a}+n^{2-2b}p^b+p\right)$$

with $\frac{2}{3}\le b<\frac{3}{4}$. Then we have

$$|Q|\le O\left(\ell^2dh+(\ell dh)^{3/2}\log n+\ell^{\frac{1}{b}}d^2h^{\frac{1}{b}}\right)$$

Proof. Each pair (u,v) in the same path in the subsystem contributes only one element to Q, since they only have branching degree 1. Thus $|Q|$ is less than the sum of square of path length of the randomly sampled h-system, which has at most $n_1=\Theta(\ell dh)$ vertices. Plugging in $n_1=\Theta(\ell dh)$ we have

$$|Q|\le O\left(\ell^2dh+(\ell dh)^{3/2}\log n+\ell^{\frac{1}{b}}d^2h^{\frac{1}{b}}\right).$$

□

Lemma 42. Suppose that we have

$$\beta(n,p,4)\le O\left(n+n^{3/4}p^{2-2a}+n^{2-2b}p^b+p\right)$$

with $\frac{2}{3}\le b<\frac{3}{4}$. Then we have

$$\beta(n,p,4)\le O\left(n+n^{3/4}p^{1/2}+n^{\frac{2}{b+1}}p^{\frac{2b+1}{2b+2}}+p\right).$$

Proof. Similar to the non-adaptive setting, we set $h=\frac{Cp}{\ell d^2}$ where C is a large enough constant. We can assume $h\le\min(\ell,d)$, since otherwise by some straightforward calculations, we immediately get

$$\|Z\|\le O\left(n^{1/2}p^{3/4}+n^{3/4}p^{1/2}\right)$$

which is better than the above bound. Then we have

$$
\begin{aligned}
&O\left(\ell^2 dh + (\ell dh)^{3/2}\log n + \ell^{\frac{1}{b}} d^2 h^{\frac{1}{b}}\right) \\
&\geq \mathbb{E}[|Q|] && \text{Lemma 40} \\
&\geq \Omega\left(\frac{h}{\ell p}\sum_{\pi_1,\pi_3\in\Pi} |R(\pi_1,\pi_3)|^2\right) && \text{Lemma 28} \\
&\geq \Omega\left(\frac{h}{\ell p^3}\left(\sum_{\pi_1,\pi_3\in\Pi} |R(\pi_1,\pi_3)|\right)^2\right) && \text{Cauchy-Schwarz} \\
&\geq \Omega\left(\frac{h}{\ell p^3}\left(p\ell^2 d^2\right)^2\right) && \text{Lemma 26} \\
&= \Omega(\ell^2 d^2).
\end{aligned}
$$

Rearranging, and using the identity $\|Z\| = nd = p\ell$, we get

$$\|Z\| \leq O\left(n + n^{3/4}p^{1/2} + n^{\frac{2}{b+1}}p^{\frac{2b+1}{2b+2}} + p\right)$$

(noting again that some terms would end up never dominating the sum and thus have been removed). □

Theorem 43. We have $\beta(n,p,4) \leq O\left(n + n^{3/4}p^{1/2} + n^{2-\sqrt{2}+o(1)}p^{\frac{1}{\sqrt{2}}} + p\right)$.

Proof. Note that by Lemma 27 we have $\beta(n,p,4) \leq O\left(n + n^{2/3}p^{2/3} + p\right)$. Let $b_0 = \frac{2}{3}$ and $b_{i+1} = \frac{2b_i+1}{2b_i+2}$. Then we have

$$\beta(n,p,4) \leq O\left(n + n^{3/4}p^{1/2+}n^{2-2b_0}p^{b_0} + p\right)$$

and using Lemma 42, by induction we have

$$\beta(n,p,4) \leq O\left(n + n^{3/4}p^{1/2+}n^{2-2b_i}p^{b_i} + p\right)$$

for all i. Since b_i converges to $\frac{1}{\sqrt{2}}$ we have $\beta(n,p,4) \leq O\left(n + n^{3/4}p^{1/2} + n^{2-\sqrt{2}+o(1)}p^{\frac{1}{\sqrt{2}}} + p\right)$. □