# Having Hope in Hops: New Spanners, Preservers and Lower Bounds for Hopsets

Shimon Kogan Weizmann Institute shimon.kogan@weizmann.ac.il Merav Parter \* Weizmann Institute merav.parter@weizmann.ac.il

#### Abstract

Hopsets and spanners are fundamental graph structures, playing a key role in shortest path computation, distributed communication, and more. A (near-exact) *hopset* for a given graph *G* is a (small) subset of weighted edges *H* that when *added* to the graph *G* reduces the number of hops (edges) of near-exact shortest paths. Spanners and distance preservers, on the other hand, ask for *removing* many edges from the graph while approximately preserving shortest path distances.

We provide a general reduction scheme from graph hopsets to the known metric compression schemes of spanners, emulators and distance preservers. Consequently, we get new and improved upper bound constructions for the latter, as well as, new lower bound results for hopsets. Our main results include:

- For *n*-vertex **directed weighted** graphs, one can provide  $(1 + \epsilon)$ -approximate distance preservers<sup>1</sup> for *p* pairs in  $V \times V$  with  $\tilde{O}_{\epsilon}(n \cdot p^{2/5} + (np)^{2/3})$  edges. For  $p \ge n^{5/4}$ , this matches the state-of-the art bounds for *reachability* preservers by [Abboud and Bodwin, SODA 2018] and the lower bound for exact-distance preservers by [Bodwin, SODA 2016].
- For *n*-vertex undirected weighted graphs, one can provide (1 + ε) distance preserves with Õ<sub>ε</sub>(n<sup>1+o(1)</sup> + p · n<sup>o(1)</sup>) edges. So far, such bounds could be obtained only for *unweighted* graphs. Consequently, we also get improved sourcewise spanners [Roditty, Thorup and Zwick, ICALP 2005] and spanners with slack [Chan, Dinitz and Gupta, ESA 2006].
- Exact hopsets of linear size admit a worst-case hopbound of  $\beta = \Omega(n^{1/3})$ . This holds even for **undirected weighted** graphs, improving upon the  $\Omega(n^{1/6})$  lower bound by [Huang and Pettie, SIAM J. Discret. Math 2021]. Interestingly this matches the recent diameter bound achieved for linear directed shortcuts.

More conceptually, our work makes a significant progress on the tantalizing open problem concerning the formal connection between hopsets and spanners, e.g., as posed by Elkin and Neiman [Bull. EATCS 2020].

<sup>\*</sup>This project is funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 949083).

<sup>&</sup>lt;sup>1</sup>I.e., subgraphs that preserve the pairwise distances up to a multiplicative stretch of  $(1 + \epsilon)$ .

# Contents

1	1 Introduction					
	1.1 Hopsets vs. Spanners, Emulators and Distance Preservers	1				
	1.2 New Results	3				
	1.3 Technical Overview	7				
	1.4 Preliminaries	10				
2	The Key Reduction: Hopsets $\rightarrow$ Missing Spanners					
3	New Distance Preservers for Directed Graphs					
4	New Spanners and Preservers for Undirected Graphs					
	4.1 Hopsets $\rightarrow$ (Unweighted) Near-Additive Spanners	18				
	4.2 Hopsets $\rightarrow$ Weighted Preservers and Spanners	19				
5	New Hopset Lower Bounds					
Α	Missing Proofs					

### 1 Introduction

### 1.1 Hopsets vs. Spanners, Emulators and Distance Preservers

This paper is concerned with establishing a formal connection between graph hopsets and the well known graphical compression schemes of spanners, emulators and distance preservers. Our proposed reduction provides a new (and black-box) approach for computing distance preserving structures as well as lower bounds for hopsets. The collection of graph structures considered in this paper has been intensively studied over the last two decades [PS89, ADD]90, EP04, BKMP05, Woo06, Pet09, Che13, HP19, EN19] due to their algorithmic centrality in the context of e.g., shortest path computation, routing, synchronization, distributed communication and beyond. While there has been evidence for algorithmic connections between these structures, no rigorous reductions have been known before. We start by providing a quick presentation of the studied graph structures.

**Hopset.** Graph hopsets, introduced by Cohen [Coh00] (and informally by [UY91, KS97, SS99]) augment the input (possibly weighted and directed) graph G = (V, E) with additional weighted edges in  $V \times V$  to reduce the hopbound of approximate shortest paths. The primary objective in this context is in optimizing the tradeoff between the size of H, the stretch parameter  $\epsilon$  and the hopbound  $\beta$ . The *bounded-hop distance* dist<sup>( $\beta$ )</sup><sub>G'</sub>(u, v) stands for the length of the shortest path between u and v in G' that has at most  $\beta$  edges (hops).

**Definition 1.1** (( $\beta$ ,  $\varepsilon$ )-Hopsets). Given a (possibly weighted and directed) graph G = (V, E), positive integer  $\beta$  and  $\epsilon \in (0, 1)$ , a subset  $H \subseteq V \times V$  of weighted edges is called a ( $\beta$ ,  $\epsilon$ )-hopset if for every  $u, v \in V$ ,  $\operatorname{dist}_G(u, v) \leq \operatorname{dist}_{G \cup H}^{(\beta)}(u, v) \leq (1 + \epsilon)\operatorname{dist}_G(u, v)$ . The edges of the hopset H are weighted by the length of the shortest path connecting their endpoints in G.

Due to their importance in the context of shortest path computation, most notably in the parallel [LJS19, EN19, Fin20, CFR20, EM21], distributed [EN16, FN18] and dynamic [HKN14] settings, hopsets have attracted a lot of activity, from combinatorial and algorithmic perspectives. The state-of-the-art bounds for *undirected* ( $\beta$ ,  $\varepsilon$ )-hopsets are obtained by employing the Thorup-Zwick algorithm as shown independently by Huang and Pettie [HP19] and Elkin and Neiman [EN19]. This algorithm has the remarkable property of providing universal ( $\beta$ ,  $\varepsilon$ )-hopsets with  $O(n^{1+1/(2^{k+1}-1)})$  edges, for  $\beta = O(k/\epsilon)^k$ , for all values of  $\epsilon$  simultaneously. For constant values of k and sufficiently small values of  $\epsilon$  (as a function of k), Abboud, Bodwin and Pettie [ABP18] showed that this tradeoff is almost tight. Our understanding of *directed* hopsets is considerably less complete. In one of the earliest papers on this topic, Ullman and Yannakakis [UY91] described a sampling-based approach to provide exact hopsets with hopbound  $\beta$  and  $O((n/\beta)^2)$  edges. Until very recently, this was the only known upper bound, even if one settles for an arbitrarily large stretch (i.e., merely preserving reachability). A recent work of Kogan and Parter [KP22] provided an improved tradeoff, which for example, provides directed  $(\beta, \varepsilon)$ -hopsets with  $\beta = O(n^{2/5})$ and linear number of edges, improving upon the state-of-the-art hopbound of  $\sqrt{n}$ . The current hopbound lower bound for linear hopsets is  $\Omega(n^{1/6})$  by Huang and Pettie [HP21], which holds even when settling for reachability.

**Spanners and emulators.** Graph spanners introduced by Peleg and Schäffer [PS89] are sparse *subgraphs* that preserve shortest path distances up to a small stretch. In contrast to hopsets, those structures are defined only for *undirected* graphs, as no sparsificiation is possible in the worst-case for directed *n*-vertex graphs with  $\Omega(n^2)$  edges.

**Definition 1.2** (( $\alpha$ ,  $\beta$ )-Spanner). Given an undirected graph G = (V, E), a subgraph  $G^* \subseteq G$  is called a ( $\alpha$ ,  $\beta$ )-*spanner* if dist<sub>*G*\*</sub>(u, v)  $\leq \alpha \cdot \text{dist}_G(u, v) + \beta$  for every  $u, v \in V$ .

An  $(\alpha, \beta)$ -emulator is a weighted set of edges in  $V \times V$  (i.e., not necessarily a *G*-subgraph) that provides the same stretch guarantees as  $(\alpha, \beta)$ -spanners. The four decade old history of spanners has initiated with the earlier works of [PS89, ADDJ90] on multiplicative spanners (where  $\beta = 0$ ). Althöfer et al. [ADDJ90] showed that a simple greedy procedure yeilds a (2k - 1)multiplicative spanner with  $O(n^{1+1/k})$  edges. This tradeoff is believed to be optimal by the Erdős girth conjecture [Erd65]. Constructions of spanners with additive stretch of +2,4,6 and  $\tilde{O}(n^{3/2}), \tilde{O}(n^{7/5}), \tilde{O}(n^{4/3})$  edges, respectively, are known by [ACIM99],[Che13],[BKMP05]. Abboud and Bodwin [AB17] demonstrated that no  $n^{4/3-o(1)}$ -size additive spanners exist, even for small *polynomial* stretch.

Elkin and Peleg [EP04] have demonstrated that the  $n^{4/3-o(1)}$  barrier can be bypassed when allowing a small *multiplicative* error; these  $(1 + \epsilon, \beta)$ -spanners are denoted as *near-additive* spanners. For every integer  $k \ge 1$  and  $\epsilon \in (0, 1)$ , [EP04] obtained  $(1 + \epsilon, \beta)$  spanners with  $O(\beta \cdot n^{1+1/k})$ edges and  $\beta = O(\log k/\epsilon)^{\log k}$ . Thorup and Zwick [TZ06] presented an alternative construction for near-additive emulators, which can be converted to near-additive spanners.

The Mystery of the Hopests–Spanners Connection. Currently, near-additive  $(1 + \epsilon, \beta)$  spanners and near-exact  $(\beta, \epsilon)$ -hopsets share almost the same tradeoff functions between their size,  $\beta$  and  $\epsilon$ . A collection of recent work noted that not only the combinatorial bounds obtained for these structures are similar but so are also their proof techniques: The Thorup-Zwick algorithm [TZ06], for example, can be slightly adapted to provide near-additive spanners, emulators and hopsets, [HP19, EN19]. Abboud, Bodwin and Pettie [ABP18] showed that the TZ bounds are nearly tight for enulators and hopsets (but still quite far from optimal for spanners). This phenomenon is quite extra-ordinary if one takes into account their fundamental differences:

- The meaning of their β parameter is very different; β is the additive *error* term for spanners and the bound on the number of hops for hopsets.
- While (1 + ε, β) spanners apply only for undirected, unweighted graphs<sup>2</sup>, there are near-exact hopsets for undirected, weighted with the same tradfeoff as given for the unweighted setting. In addition, sub-quadratic hopset constructions can also be provided for directed, weighted graphs. (The latter is impossible, in the worst case, for spanners).
- For near-additive spanners it is sufficient (and necessary) to restrict attention to **nearby** pairs (i.e., pairs at distance  $O(\beta/\epsilon)$  in *G*). Hopset constructions for unweighted graphs, on the other hand, are concerned with **distant** pairs (e.g., at distance  $\Omega(\beta)$ ).
- Spanners are trivial when the input graph is sparse, but this is not necessarily true for hopsets. For example, for a linear-size graph *G* and parameters  $\epsilon = 0, \beta = o(n^{1/5})$ , the output  $(1 + \epsilon, \beta)$  spanner is simply *G*. In contrast, the output  $(\beta, \epsilon)$ -hopset for *G* might require (in the worst case) a super-linear number of edges.

This quite mysterious connection between hopsets and spanners has been a subject for a thorough research [EN16, EN17, ABP18, HP19, EN19, BP20, SN21]. In their inspiring and comprehensive survey on this topic [EN20], Elkin and Neiman conclude by asking:

<sup>&</sup>lt;sup>2</sup>For weighted graphs, the  $\beta$  parameter must depend on the edge weights, see [EGN22].

[[EN20]] There is a striking similarity not just between the results concerning near-additive spanners for unweighted graphs and near-exact hopsets for weighted ones, but also between the techniques used to construct them and to analyze these constructions. A very interesting open problem is to explain the relationship between near-additive spanners and near exact hopsets rigorously, i.e., by providing a reduction between these two objects.

**Distance Preservers.** An additional very active line of graph sparsification research considers the fundamental property of pairwise shortest path distances. For a given graph G = (V, E) and a subset of p demand pairs  $P \subset V \times V$ , a distance preservers is a sparse subgraph  $G^* \subseteq G$  satisfying that dist<sub>G</sub>\*(u, v) = dist<sub>G</sub>(u, v) for every  $u, v \in P$ . Distance preservers, introduced by Coppersmith and Elkin [CE06] have played an important role in the constructions of spanners [Pet09, EFN17], distance oracles [EP16], as well as in establishing lower bounds for spanners [AB16] and hopsets [HP21, AB18]. The related notion of pairwise spanners allows one to approximately preserve the pairwise distances, up to a small stretch. Pairwise preservers. There is a long line of works on exact preservers [BCE03, CE06, BW16, Bod17, Bod21] and additive pairwise spanners [CGK13, KV13, Kav15, KV15, AB16] from an upper bound and lower bound perspectives. In their seminal work, Coppersmith and Elkin [CE06] provided a family of lower bounds for exact preservers of  $\Omega(n^{2d/(d^2+1)} \cdot p^{d(d-1)/d^2+1})$  edges for all integers  $d \ge 2$ . Hence, exact preservers might be very dense in the worst-case.

Near-exact preservers are known to be quite sparse but only for unweighted undirected graphs. Using a folklore construction (see e.g.,  $[ABS^+20a]$ )  $(1 + \epsilon, \beta)$  spanners can be converted into near-exact preservers with  $O(n + n^{o(1)}p)$  edges. No analogous result is known for weighted graphs (as near-additive spanners exist only in the unweighted detting). In this paper, we fill in this missing gap by providing an alternative reduction from (near-exact) hopsets to (near-exact) preservers. Unlike spanners, the TZ hopset have the same quality in the weighted case, which allows us to compute sparse (near-exact) weighted preservers that almost match their unweighted counterparts. The latter has immediate applications to additional useful notions of spanners, such as sourcewise spanners [RTZ05, Par14] and spanners with slack [CDG06]. We next present our main contribution in more details.

### 1.2 New Results

Throughout, the notation  $\widehat{O}(.)$  (resp.,  $\widetilde{O}(.)$ ) hides  $n^{o(1)}$  (resp. poly(log n)) factors. For clarity of presentation, we assume that the edge weights of the weighted graph considered are polynomial.

**From Hopset to Distance Preservers.** We translate constructions of near-exact hopsets into near-exact preservers in a black-box manner. Importantly, we do not adapt existing hopset algorithms into preserver algorithms, but rather only convert their *output* hopsets into near-exact preservers. While the reduction scheme is general, in the context of directed graphs, it is most insightful to consider the family of algorithms that compute for every  $\beta \le n^b$  (for some given parameter  $b \in [0, 1)$ ) and any  $\epsilon \in (0, 1)$ , a ( $\beta, \epsilon$ ) (directed) hopset with  $\widetilde{O}(n^2/\beta^a)$  edges, for some parameter a > 1. The intuition behind this function is the following. First, the size bound converges to  $n^2$  as  $\beta$  approaches to 1 (which indeed makes sense as for  $\beta = 1$ , the hopset collides with the transitive closure of the graphs). Second, this function captures all existing hopset constructions for directed graphs: In the folklore algorithm of [UY91], a = 2 and b = 0. In the algorithm of [KP22], we have a = 3 and b = 1/4.

**Theorem 1.1 (From Directed Hopset Hierarchy to Preservers).** Let  $0 \le b < \min(1, 1/(a-1))$  and a > 1 be fixed parameters. Then, given any algorithm for computing  $(\beta, \epsilon)$  hopsets for n-vertex directed (possibly weighted) graphs with  $\tilde{O}_{\epsilon}(n^2/\beta^a)$  edges for every  $\beta \le n^b$ , the following holds: For any n-vertex directed (possibly weighted) graph G and a set P of p demand pairs, one can compute an  $(1 + \epsilon')$ -distance preserver  $G^* \subseteq G$  where

$$\epsilon' = O(\epsilon \log \log n) \text{ and } |G^*| = \widetilde{O}_{\epsilon'}(n \cdot p^{1-1/k} + n^{2/a} \cdot p^{1-1/a}) \text{ for } k = \frac{2-a \cdot b}{1-b}.$$

More specifically, Theorem 1.1 is based on translating a *collection* (or *hierarchy*) of  $(\beta_i, \epsilon)$  hopsets, for  $\beta_1 > \beta_2 > ... > \beta_\ell$  where  $\ell = O(\log \log n)$ , into near-exact preservers. The sparsity of the preservers is a function of the size of the given hopsets and their  $\beta$  parameter. Note that Lemma 1.1 can be safely applied even for  $\epsilon > 1/\log \log n$  (e.g., in our applications to reachability preservers). Plugging the bounds of  $(\beta, \epsilon)$  hopsets from [KP22] (i.e., a = 3 and b = 1/4) to Theorem 1.1 yields the first construction of near-exact hopsets for weighted and directed graphs:

**Theorem 1.2** ((1 +  $\epsilon$ ) **Directed Preservers**). *For every n-vertex* weighted directed graph G = (V, E) and p pairs  $P \subset V \times V$ , one can compute a  $(1 + \epsilon)$ -preserver for P with  $\tilde{O}(np^{2/5} + (np)^{2/3})$  edges.

So-far, only *exact* preservers have been known for this setting with  $O(n\sqrt{p})$  edges for unweighted graphs [CE06], and  $O(\min\{n\sqrt{p}, n^{2/3}p\})$  edges for weighted graphs [CE06, BW16, Bod21]. Interestingly, for  $p \ge n^{5/4}$ , our bounds match the state-of-the-art results for (the weaker notion of) reachability preservers by Abboud and Bodwin [AB18]. Moreover, it also matches the lower-bound for exact preservers by Bodwin [Bod21]. Theorem 1.1 is quite general and can also be shown to recover other known results in the literature but in a *black-box* manner. For example, applying this theorem with the *diameter-reducing shortcut*<sup>3</sup> algorithm of [KP22] for which a = 3 and b = 1/3, provides *reachability preservers* with  $\tilde{O}(np^{1/3} + (np)^{2/3})$  edges. Hence, matching the known bound of [AB18] for  $p \ge n$ . Plugging in this theorem the known bounds for *exact* hopsets, for which a = 2 and b = 0, recovers the Coppersmith-Elkin bounds of  $O(n \cdot \sqrt{p})$  edges for *exact* preservers [CE06].

We also provide a similar reduction for *undirected* graphs for which considerably sparse hopsets exist. By using the state-of-the-art bounds for ( $\beta$ ,  $\epsilon$ ) hopsets of [HP19, EN19] we provide the first near-exact preservers for weighted undirected graphs:

**Theorem 1.3**  $((1 + \epsilon)$  **Undirected Weighted Preservers).** For every *n*-vertex weighted undirected graph G = (V, E) and a subset of *p* pairs *P*, one can compute a  $(1 + \epsilon)$ -preserver for *P* with  $n^{o(1)} \cdot (n + p)$  edges.

This almost matches the state-of-the-art bounds of  $O(n + p \cdot n^{o(1)})$  edges known for the unweighted case. Note that while the unweighted case follows easily from near-additive spanners, our weighted preservers cannot be obtained by using the weighted-analogue of near-additive spanners<sup>4</sup> (e.g., as given by [EGN22]). Instead, Lemma 1.3 is obtained by using our hierarchy of hopset constructions. This hopset-based construction enjoys the fact there are near-exact hopsets that meet the TZ tradeoff even for weighted graphs (in contrast to near-additive spanners). See Table 1.2 for a more detailed comparison with the known bounds.

<sup>&</sup>lt;sup>3</sup>The latter can viewed as  $(\beta, n)$  hopsets as it preservers merely reachability.

<sup>&</sup>lt;sup>4</sup>This is because the additive term in those spanners depends linearly on the edge weight.

		Approximation	Upper Bound
	Unweighted	1	$O(\min\{n+n^{1/2}p, np^{1/3}+n^{2/3}p^{2/3}\}) \text{ [CE06, BW16]}$
		$(1+\epsilon)$	$O_{\epsilon}(n+p\cdot n^{o(1)})$ [ABS <sup>+</sup> 20a]
Undirected	directed Weighted	1	$O(\min\{n+n^{1/2}p,n\sqrt{p}\})$ [CE06]
		$(1+\epsilon)$	New: $\widetilde{O}_{\epsilon}((n+p)\cdot n^{o(1)})$
	Unweighted	1	$O\left(\min\{n\sqrt{p}, n^{2/3}p + n\}\right)$ [CE06, Bod21]
		n (reachability)	$O(n + (np)^{2/3})$ [AB18]
Directed	rected	1	$O\left(\min\{n\sqrt{p}, n^{2/3}p + n\}\right)$ [CE06, Bod21]
	weighted	$(1+\epsilon)$	<b>New:</b> $\widetilde{O}_{\epsilon} \left( np^{2/5} + (np)^{2/3} \right)$

Table 1: The state-of-the-art bounds for pairwise preservers with p pairs. New results provided in this paper are shown in purple. We also obtain reachability preservers (by our black-box reduction) that matches the bound of [AB18] for every  $p \ge n$ .

**From Hopset to Emulators and Spanners.** While the reduction from hopsets to preservers (of Thm. 1.1) is based on converting a hopset *hierarchy* into the output preservers, we observe that the reduction to emulators for unweighted graphs is significantly simpler and requires only a single hopset<sup>5</sup>. Specifically, we show that augmenting a near-exact hopset with a multiplicative spanner provides a near-additive emulator! Formally, we have:

**Observation 1.4** (Hopsets + Multiplicative Spanners = Emulator). For a given n-vertex unweighted graph G = (V, E), let  $H^*$  be some  $(\beta, \varepsilon)$  hopset for G and let  $G^*$  be a t-multiplicative spanner for G. Then,  $H^* \cup G^*$  is a  $(1 + \varepsilon, \beta \cdot t)$  emulator.

By using a standard reduction from emulators to near-additive spanners, we get:

**Observation 1.5 (Near-Exact Hopsets**  $\rightarrow$  **Near-Additive Spanner).** For an unweighted *n*-vertex graph G = (V, E), let  $H^*$  be some  $(\beta, \varepsilon)$  hopset for G, and let  $G^*$  be a *t*-multiplicative spanner for G. Then, one can compute an  $(1 + 2\varepsilon, \beta \cdot t)$  spanner  $\widehat{G}$  with  $O(t|H^*| \cdot \beta/\varepsilon + |G^*|)$  edges.

While the near-additive emulators and spanners obtained in this black-box manner have slightly suboptimal bounds (see our **Concluding Remark.**), we find this reduction to be useful and arguably powerful for two main reasons. (i) It makes a progress on the open problem raised by Elkin and Neiman [EN20] by providing a near-optimal reduction from hopsets to spanners; reducing the hopset-type  $\beta$  to the spanner-type  $\beta$ . (ii) This reduction serves us later on to provide improved lower bound for *hopsets*. More broadly, it allows one to translate metric-compression lower bounds into hopset lower bounds.

While the above reduction is limited to unweighted graphs, we use a more delicate reduction for *weighted* graphs. This is done again by converting a hierarchy of hopsets (rather than a single hopset) into a spanner. Using the best-on-shelf undirected hopsets (e.g., of [EN19, HP19]) we get:

<sup>&</sup>lt;sup>5</sup>Hopsets and emulators are indeed expected to be "closer" as they are both allowed to use non-*G* edges.

**Theorem 1.6 (New Spanners for Weighted Graphs).** *Given an n-vertex weighted graph G, the hopsets provided by the algorithms of* [*EN19, HP19*] *can be converted into the following subgraphs:* 

- 1.  $(1 + \epsilon, O(k/\epsilon)^k \cdot W_{\max} \cdot \log n)$ -spanner  $G' \subseteq G$  with  $\widehat{O}(n^{1+1/(2^{k+1}-1)})$  edges where  $W_{\max}$  is the largest edge weight<sup>6</sup>.
- 2. For every integer k and  $S \subseteq V$ , a S-sourcewise  $(4k 1 + \epsilon)$ -spanner  $G' \subseteq G$  with  $\widehat{O}(n + |S|^{1+1/k})$ .
- 3. For any  $\epsilon \in (0,1)$  and integer  $k \ge 1$ , an  $\epsilon$ -slack (12k 1 + o(1))-spanner  $G' \subseteq G$  with  $\widehat{O}(n + (1/\epsilon)^{1+1/k})$  edges.

Spanner constructions with weighted additive stretch have been recently presented by [EGN22, ABS<sup>+</sup>20b, EGN21, ABS<sup>+</sup>21]. An *S*-sourcewise *t*-spanner for a subset of sources  $S \subseteq V$  provides *t*-multiplicative stretch for all pairs in  $S \times V$  [RTZ05, Par14]. The state-of-the-art bounds for sourcewise spanners are given by Elkin, Filtser and Neiman [EFN17] that provided *S*-sourcewise (4k - 1)-spanners with  $O(n + \sqrt{n}|S|^{1+1/k})$  edges. The notion of  $\epsilon$ -slack *t*-spanners introduced by Chan, Dinitz and Gupta [CDG06] guarantees a *t*-multiplicative stretch for all but  $\epsilon$  fraction of the pairs. Our bounds should then be compared with [CDG06] which provide  $\epsilon$ -slack (12k - 1 + o(1))-spanners with  $O(n + (1/\epsilon)^{1+1/k}\sqrt{n})$  edges. In the lack of near-exact preservers for weighted graphs, the prior constructions of sourcewise spanners and spanners with slack simply used the exact preservers of Elkin and Coppersmith [CE06] with  $O(n + \sqrt{n}p)$  edges. Our constructions use instead the near-exact preservers of Theorem 1.3 which have only  $\hat{O}(n + p)$  edges, hence sparser for  $p > n^{1/2+o(1)}$ .

**New Hopset Lower Bounds.** Currently all existing lower bound constructions for hopsets are based on packing multiple pairs of vertices whose shortest paths are unique, edge disjoint and long. These extremal graphs have been studied by Alon [Alo02], Hesse [Hes03] and Coppersmith and Elkin [CE06]. Interestingly, the same lower bound techniques (up to small adaptations) are shared by spanners, emulators and hopsets [Hes03, HP21, ABP18, KLX22]. Due to the fundamental differences between hopsets and metric-compression structures, prior work had to repeat the arguments for hopsets while introducing adaptions. We obtain the following lower bounds in black-box manner:

### **Theorem 1.7** (New Hopset Lower Bounds). For undirected *n*-vertex graphs, we have:

- *Exact Hopsets*: For weighted graphs, exact hopsets of linear size admit a worst-case hopbound of  $\beta = \Omega(n^{1/3})$ . For undirected unweighted graphs the lower bound becomes  $\beta = \Omega(n^{1/5})$ .
- Hopsets with Sublinear Error: Any hopset with  $f(d) \leq d + O(kd^{1-1/k}) + \widetilde{O}(1)$  and  $\beta = O(kd^{1-1/k}) + \widetilde{O}(1)$  requires in the worst case  $\Omega(n^{1+1/(2^k-1)-o(1)})$  edges.

Our hopbound lower bound for exact hopsets improves upon the the state-of-the-art  $\Omega(n^{1/6})$  lower bound by Huang and Pettie [HP21] which holds for reachability. No stronger bounds were known for exact hopsets. The lower bounds for hopsets with sublinear error nearly match the emulator size bounds by Abboud, Bodwin and Pettie [ABP18].

**From Reachability Preservers to** *d***-Shortcuts Lower Bounds.** Our reduction scheme of Theorem 1.1 also allows us to translate lower bounds for reachability *preservers* [AB18] into a lower bound on the diameter (hopbound) of *d*-shortcuts (i.e., ( $\beta = d, \varepsilon = n$ ) hopsets) of linear size. We show:

<sup>&</sup>lt;sup>6</sup>See Lemma 4.6 for the more precise statement.

**Theorem 1.8.** If there is a lower bound of  $\Omega(n^{\alpha} \cdot p^{\gamma})$  edges, where  $\alpha + \gamma \ge 1 + o(1)$ , for reachability preservers with p-pairs for n-vertex graphs, then there exists an n-vertex directed graph for which any linear sized d-shortcut must satisfy that  $d \ge n^{(\alpha+\gamma-1-o(1))}$ .

Plugging in this theorem the state-of-the-art lower bound for reachability preservers of [AB18] yields the state-of-the-art diameter lower bound of [HP21]. Any improved lower bound result for reachability preservers would immediately lead to improve diameter bound for shortcuts.

#### 1.3 Technical Overview

Our key result is a reduction from a hierarchy of hopsets to preservers and spanners in weighted and directed graphs. The reduction works by converting the given hopset hierarchy into an intermediate graph structure that we call *bounded-missing spanners*. This notion of spanners, which to the best of our knowledge has not been introduced before, serves as a key transitional junction on the road from hopsets to distance preserving subgraphs.

**New Graph Notion: Bounded-Missing Spanners.** For a given (possibly weighted and directed) graph *G*, a *bounded-missing spanner* is a subgraph of *G* that contains all but a bounded number of (missing) edges of some approximate *u-v* shortest-path for every  $u, v \in V$ . For integer parameters *r* and *t*, a subgraph *G'* of *G* is an *r*-missing *t*-spanner if for every  $u, v \in V$ , there is some *u-v* path  $P_{u,v} \subseteq G$  of length at most  $t \cdot \text{dist}_G(u, v)$  such that  $|P_{u,v} \setminus G'| \leq r$ ; i.e., for every  $u, v \in V$  there is a *t*-approximate shortest path in the spanner that misses at most *r* edges. We call such a *u-v* path  $P_{u,v}$  an *r*-missing *t*-approximate path. Note that for r = 0, we get the standard definition of *t*-spanners. Our proof technique is based on the following steps. First, we show how to convert a given hierarchy of near-exact hopsets into an *r*-missing *t*-spanner  $G' \subseteq G$ . This also provides a polynomial time algorithm for computing the desired *t*-approximate *r*-missing shortest paths in *G* w.r.t *G'*. Second, we show that for a wide class of hopset algorithms (namely, those that admit natural size vs. hopbound tradeoff functions) one can carefully select the hopset hierarchy in a way that optimizes the tradeoff between *r*, *t* and the size of the output spanner. Lastly, we show how to convert these missing spanners into approximate distance preservers and near-additive spanners. This last part is the most immediate. We next elaborate mainly on that two first parts.

The Key Sparsification Lemma. Our main lemma shows that for any given hierarchy of  $(\beta_i, \epsilon)$ hopsets  $H_0, H_1, H_2, \ldots, H_\ell$ , for a decreasing sequence  $n = \beta_0 > \beta_1 > \beta_2 > \ldots > \beta_\ell$ , one can
compute a *r*-missing *t*-spanner *G* with  $r = \beta_\ell$  and  $t = (1 + \epsilon)^\ell$  and

$$|G'| = \sum_{i=1}^{\ell} |H_i| \cdot \beta_{i-1} \text{ edges }.$$
 (1.1)

Note that the subset of  $H_i$  graphs are *not* subgraphs of G, while the missing spanner G' is a Gsubgraph. The lemma is shown by applying an inductive argument with the following intuition. Suppose that in the given graph G, it holds that  $\operatorname{dist}_{G}^{(r)}(u,v) \leq t \cdot \operatorname{dist}_{G}(u,v)$  for every  $u, v \in V$ . In such a case,  $G' = \emptyset$  is a valid *r*-missing *t*-spanner for G. Hence, computing missing spanners is trivial with this assumption. Recall that reducing the number of hops of approximate shortest paths is precisely the goal of hopsets! Consider next the  $(\beta_1, \epsilon)$  hopset  $H_1$ , namely, the first hopset in the hierarchy. By the hopset definition, for any  $u, v \in V$  there is some  $(1 + \epsilon)$ -approximate shortest path  $P_{u,v} \subseteq G \cup H$  of at most  $\beta_1$  edges. To compute the missing spanner G', each (x, y)edge in  $H_1$  is then replaced by some *x*-*y* shortest path in *G* (which has at most  $\beta_0 = n$  edges). This explains the first summand of Eq. (1.1). In step  $i \geq 1$  of the induction given the current missing spanner  $G'_{i-1}$ , it is shown that for each edge  $(x, y) \in H_i$  there is already some  $(1 + \epsilon)^{i-1}$ approximate *x*-*y* shortest path  $P_{x,y}$  in *G* such that  $|P_{x,y} \setminus G'_{i-1}| \leq \beta_{i-1}$ . Hence, to include  $P_{x,y}$  in *G'*, it is sufficient to add at most  $\beta_{i-1}$  edges for each  $(x, y) \in H_i$ . This explains the  $i^{th}$  summand of Eq. (1.1). As each inductive step can be shown to increase the shortest-path approximation by  $(1 + \epsilon)^{\ell}$ -spanners.

A-priori, it is unclear if a *nice* hierarchy of hopsets, for which Eq. (1.1) yields a *sparse* output spanner *G'*, even *exists*. Interestingly, we show that it does. This is mainly due to the fact that there are near-exact hopsets of *sublinear* size and with a  $o(\sqrt{n})$  hopbound (e.g., of [KP22]). While prior work on hopsets mostly focused on *linear* size hopsets, in our context it is the sublinear regime that plays the critical role in the computation of preserving subgraphs.

**Bounded-Missing Spanners**  $\rightarrow$  **Approximate Pairwise Preservers and Spanners.** Consider an *r*-missing *t*-spanner  $G' \subseteq G$ , and suppose that for every  $u, v \in V$  there is a polynomial time algorithm for computing the *r*-missing *t*-approximate *u*-*v* path  $P_{u,v}$  in *G*. It then easy to see that one can obtain an approximate preserver for a given *p* pairs  $P \subset V \times V$  by augmenting G' with the missing edges of  $P_{u,v} \setminus G'$  for every  $u, v \in P$ . This adds at most  $p \cdot r$  edges. Taking  $t = (1 + \epsilon)$  provides a near-exact preservers. Our near-exact preservers find further applications in the context of sourcewise spanners [RTZ05, Par14, EFN17] and spanners with slack [CDG06]. We also show that by augmenting a bounded-missing spanner with a standard multiplicative spanner, one can obtain a near-additive spanner (see Lemma 2.4).

A High Level Intuition for Theorem 1.1, Reachability Preservers as a Case Study. To get a clean intuition into this general reduction, consider<sup>7</sup> the most basic notion of "reachability" hopsets, known as *diameter-reducing shortcuts* [Tho92]. For a given directed graph *G* with a transitive closure TC(G), a graph  $H \subseteq TC(G)$  is a *d*-shortcut if for every  $(u, v) \in TC(G)$ , *H* contains a *d*-hop directed path from *u* to  $v^8$ . Theorem 1.1 provides a recipe for translating a hierarchy of *d*-shortcuts into *reachability preservers*, namely, subgraphs that preserve reachability between a given set *P* of *p* demand pairs [ABP18]. We first define the hopset hierarchy (or a *d*-shortcut hierarchy), and then show how to translate it into a preserver with  $\tilde{O}(np^{1/3} + (np)^{2/3})$  edges. To define the hopset hierarchy, we use (in a black-box manner) the state-of-the-art bounds for *d*-shortcut with  $\tilde{O}(n^2/d^3)$  edges, and (ii) for every  $d > n^{1/3}$ , there is a *d*-shortcut with  $\tilde{O}((n/d)^{3/2})$  edges (Lemma 3.1 shows how to derive (ii) from (i)<sup>9</sup>). Hence, we use Theorem 1.1 with a = 3 and b = 1/3. The proof has two parts, depending on the number of pairs *p*.

Assume first that  $p \le n$ , in which case the dominating term in the desired size of the output preserver is  $np^{1/3}$ . We define a hopset hierarchy  $H_0, H_1, \ldots, H_\ell$  for  $\ell = O(\log \log n)$  where each  $H_i$  is a  $(\beta_i, \epsilon)$  hopset for an exponentially decaying sequence  $n = \beta_0 > \beta_1 \ldots > \beta_\ell = n^{1/3}$ . Hence, all hopsets in this hierarchy are of *sublinear* size. We set the final value of  $\beta_\ell$  to the value *d* for which the following equality holds:  $(\mathbf{n}/\mathbf{d})^{3/2} \cdot \mathbf{d} = \mathbf{p} \cdot \mathbf{d}$ . The reasoning behind this equation is as follows. The left term corresponds to the size<sup>10</sup> of the *d*-missing *n*-spanner *G'* (see Eq. (1.1)). The right term corresponds to the number of edges needed to be added to *G'* in order preserve

<sup>&</sup>lt;sup>7</sup>The main applicability of this theorem is in translating near-exact hopsets to preservers. However, it is easier to convey the ideas on shortcuts, which have a simpler size vs. diameter tradeoff function.

<sup>&</sup>lt;sup>8</sup>A *d*-shortcut is analogous to ( $\beta = d, n$ ) hopsets, but usually one uses the term hopsets in the context of approximating shortest paths.

<sup>&</sup>lt;sup>9</sup>This was shown in [KP22] for a particular function, and Lemma 3.1 considers a general function.

<sup>&</sup>lt;sup>10</sup>The series of Eq. 1.1 converges at first point for which  $\beta_i = \beta_{i+1}$ .

the reachability for the given *p* pairs (i.e., adding to *G*', the *d* missing edges for each pair). Solving this equation for *d* provides the desired size bound of  $O(np^{1/3})$ .

Next, assume that p > n, and thus the size of the output preserver is dominated by  $\tilde{O}((np)^{2/3})$ . The hopset hierarchy in this case consists of  $2\ell$  hopsets. The first  $\ell$  hopsets in the hierarchy are  $(\beta_i, \epsilon)$  hopsets for  $\beta \le n^{1/3}$  (hence of *sublinear* size, except for the last one). The second set of hopsets are *superlinear* with  $\beta_i > n^{1/3}$ . We set the final value of  $\beta_\ell > n^{1/3}$  to the value *d* for which the following equality holds:  $\mathbf{n}^2/\mathbf{d}^3 \cdot \mathbf{d} = \mathbf{p} \cdot \mathbf{d}$ , which follows the same logic as explained above, with the only distinction that we use the super-linear regime of the *d*-shortcut function. Solving for *d*, provides the desired bound of  $\tilde{O}((np)^{2/3})$ . An important conclusion from these relations (also discussed next and in Sec. 5) is that lower bounds for reachability preservers would immediately provide lower bounds for *diameter-reducing shortcuts* (see Theorem 1.8).

The near-exact weighted and directed preservers of Theorem 1.2 are obtained using the output  $(\beta, \epsilon)$ -hopsets of [KP22] (for which we plug a = 3 and  $\beta = 1/4$  in Theorem 1.1, see Thm. 1.12). **Graph Compression Lower Bounds**  $\rightarrow$  **Hopset Lower Bounds.** Finally, we enjoy the complementary aspect of our reduction to provide lower bounds for hopsets. These lower bounds are obtained in a black-box manner from existing lower bound results for preservers, spanners and emulators. Our automatic translation provides two new meaningful results. By translating Bodwin's lower bound result for exact preservers [Bod21] we get an improved lower bound on the hopbound  $\beta$  for linear-sized hopsets already for undirected graphs. So-far, the only known lower bound for this setting was  $\beta = \Omega(n^{1/6})$  by Huang and Pettie (which holds for the weaker notion of diameter-reducing shortcut). We show that Bodwin's lower bounds [Bod21] directly improve the hopbound lower bound to  $\Omega(n^{1/3})$  and  $\Omega(n^{1/5})$  for weighted (reps., unweighted) graphs. In addition, by translating the lower bound results for emulators of Abboud, Bodwin and Pettie [ABP18], we get new lower bounds for hopsets with sublinear stretch and hopbound functions.

A summary of our results and their connections is illustrated in Fig. 1.

A Concluding Remark. In this work, we delved into the mysterious connection between hopsets and distance preserving subgraphs by providing a general reduction from the former to the latter. We note that a reverse reduction is somewhat less plausible, as the subgraph problem is trivial for sparse graphs, while hopsets are not. The main two benefits of hopsets (in comparison to spanners) that we enjoy of in our constructions are: (i) there are sublinear hopsets with  $o(\sqrt{n})$  hopbounds, and (ii) there are efficient near-exact hopsets for weighted and directed graphs. Note that while most (if not all) of the known algorithmic applications of hopsets (e.g., for shortest-path computation) require linear size, our work put the spotlight on hopsets of *sublinear* size.

Finally, we note that our reductions from near-exact hopsets to near-additive emulators and spanners are slightly suboptimal in an "unavoidable" manner: emulators and spanner do not "account" for edges incident to low-degree vertices (as one can simply add all these edges to the structure)<sup>11</sup>. Hopsets however do account for these edges in their hopbound measure<sup>12</sup>. While one can slightly adapt the hopset definition (e.g., counting only edges incident to high-degree vertices in the hopbound measure) to make these notions closer, this work seeks for a *black-box connection* with no adaptations! Perhaps surprisingly this on its own provides a collection of improved subgraph constructions and lower bounds. It would be interesting to study the "minimal" adaptations for the hopset definition to provide a (hopefully tight) reverse reduction.

 $<sup>\</sup>overline{{}^{11}\text{E.g., for a desired size bound of } O(n^{1+1/k})}$  edges, one can add all edges incident to vertices with degree at most  $n^{1/k}$  to the output subgraph.

<sup>&</sup>lt;sup>12</sup>This leads to the tradeoff differences between emulators and hopsets [HP19] in the TZ algorithm [TZ06].



Figure 1: High-level description of the main results and their connections.

### 1.4 Preliminaries

We use  $O_{\epsilon,W}(.)$  to hide polynomially factor in  $1/\epsilon$  and logarithmic factors in the maximum edge weight. For an *n*-vertex directed graph *G*, let TC(G) denote the transitive closure of *G*, and let  $TC_W(G)$  denote the transitive closure of *G* weighted by the corresponding shortest path distances. For a possibly weighted graph *G*, let dist<sub>*G*</sub>(*u*, *v*) at the *weight* of the shortest path from *u* to *v*. Let len(*Q*) be the *length* of a path *Q*, measured by the sum of its weighted edges. Let |Q| be the number of edges on this path. For unweighted graphs, len(*Q*) = |Q|. For any vertex pair  $u, v \in V$ , define dist<sup>(\beta)</sup><sub>*G*</sub>(*u*, *v*) to be the minimum length *u*-*v* path with at most  $\beta$  edges (hops). If there is no such path, then dist<sup>(\beta)</sup><sub>*G*</sub>(*u*, *v*) =  $\infty$ . We assume w.l.o.g that every edge (*u*, *v*)  $\in$  *G* is a shortest path between *u* and *v* in *G*. When considering a weighted graph G = (V, E, W) we may omit the weight function *W*, when it is not explicitly used.

**Definition 1.3** (Pairwise Spanners). For a given (possibly weighted) graph G = (V, E) and a subset of pairs P, a subgraph  $G^* \subseteq G$  is an  $(\alpha, \beta)$  P-spanner if  $\operatorname{dist}_{G^*}(u, v) \leq \alpha \cdot \operatorname{dist}_G(u, v) + \beta$ . When  $P \subseteq S \times V$  (resp.,  $P \subseteq S \times S$ ) for  $S \subseteq V$ ,  $G^*$  is denoted as an  $(\alpha, \beta)$  sourcewise spanner (resp., subsetwise spanner).

**Lemma 1.9** (Lemma 5 of [Zwi02, CC20]). Given a directed n-vertex graph G with integer weights<sup>13</sup> in

<sup>&</sup>lt;sup>13</sup>One can also handle rational weights with  $O(\log n)$  precision by multiplying each weight by the lowest common denominator of all the edge weights. The running time is polynomial, which is sufficient for our purposes.

[1, *M*], there is an algorithm  $APSP^{\leq R}$  that compute the *R*-hop distances and paths in time  $O(M \cdot R \cdot n^{\omega})$  time. The output of the algorithm is given by  $\mathcal{P}$  that consists of all *u*-*v R*-hop shortest paths  $P_{u,v}$  for every  $(u, v) \in TC(G)$ . I.e., for every  $P_{u,v} \in \mathcal{P}$ , it holds that  $len(P_{u,v}) = dist_G^{(R)}(u, v)$ .

**Lemma 1.10.** [*ADDJ90*] For every *n*-vertex (possibly weighted) graph G and a given integer  $k \ge 1$ , one can compute a (2k - 1)-spanner  $H \subseteq G$  with  $|H| \le n^{1+1/k}$  edges.

**Inequality 1.1.** For every  $\epsilon \in (0,1)$  and any positive integer t, it holds  $(1 + \epsilon/(2t))^t \leq (1 + \epsilon)$ .

**Diameter-Reducing Shortcuts, Reachability Preservers and Directed Hopsets.** For a directed graph *G*, a *d*-shortcut  $H \subseteq TC(G)$  satisfies that the directed diameter of  $H \cup G$  is at most *d*. I.e., for every  $(u, v) \in TC(G)$ , the graph  $G \cup H$  contains a *u*-*v* path with at most *d* edges. Note that a *d*-shortcut is simply a  $(\beta = d, n)$  hopset. A subgraph  $G' \subseteq G$  is a *reachability preserver* for a pair set  $P \subseteq V$  if it contains a *u*-*v* (directed) path for every  $(u, v) \in TC(G) \cap P$ . The following results are given in [KP22] for *d*-shortcuts and near-exact directed hopset:

**Theorem 1.11.** [*Theorem 1.1 in* [*KP22*]] For every *n*-vertex graph *G*, there is an  $\tilde{O}(n^{\omega})$ -time<sup>14</sup> randomized algorithm for computing a *d*-shortcut set of cardinality:

$$S(n,d) = \begin{cases} \widetilde{O}(n^2/d^3), & \text{for } d \le n^{1/3}; \\ \widetilde{O}((n/d)^{3/2}), & \text{for } d > n^{1/3}, \end{cases}$$

**Theorem 1.12.** [*Theorem 1.4 in* [KP22]] For every *n*-vertex graph *G* with integer weights [1, M],  $\epsilon \in (0, 1)$  and  $\beta \in \mathbb{N}_{\geq 1}$ , there is an  $\widetilde{O}(n^3 \cdot \text{poly} \log(nM))$ -time randomized algorithm for computing  $(\epsilon, \beta)$  hopsets whose number of edges is bounded by:

$$H(n,\beta,\epsilon) = \begin{cases} \widetilde{O}_{W,\epsilon}\left(\left(\frac{n^2}{\beta^3}\right)\right), & \text{for } \beta \le n^{1/4}; \\ \widetilde{O}_{W,\epsilon}\left(\left(\frac{n}{\beta}\right)^{5/3}\right), & \text{for } \beta > n^{1/4} \end{cases}$$

## **2** The Key Reduction: Hopsets → Missing Spanners

Let G = (V, E) be an *n*-vertex (possibly weighted and directed) graph.

**Definition 2.1** (*r*-missing *t*-spanner). A subset of edges  $G' \subseteq G$  is an *r*-missing *t*-spanner for *G* if for every  $(u, v) \in TC(G)$ , there is a *u*-*v* path  $P_{u,v} \subseteq G$  satisfying:

$$\operatorname{len}(P_{u,v}) \le t \cdot \operatorname{dist}_G(u,v) \text{ and } |P_{u,v} \setminus G'| \le r.$$

$$(2.1)$$

For r = 0, a 0-missing *t*-spanner is simply a *t*-spanner.

Our key lemma shows that one can transform a family of  $(\beta_i, \epsilon)$ -hopsets for a sequences  $\beta_1 \ge \beta_2 \ge \ldots \ge \beta_\ell$  into an *r*-missing *t*-spanner for  $r = \beta_\ell$  and  $t = (1 + \epsilon)^\ell$ . This lemma is general enough to later on provide black-box constructions of near-exact preservers (even in the directed weighted setting), spanners, as well as, reachability preservers (for which  $\epsilon = n$ ).

<sup>&</sup>lt;sup>14</sup>Where  $\omega$  is the optimal matrix multiplication constant.

**Lemma 2.1.** [From Hopsets to r-Missing t-Spanners] Let  $G = (V, E, \omega)$  be an n-vertex (possibly directed) graph and a set  $P \subseteq V \times V$  of demand pairs. Then given a sequence of  $\ell$  values  $n = \beta_0 \ge \beta_1 \ge \beta_2 \ge \ldots \ge \beta_\ell$  and a a collection of  $(\beta_i, \epsilon)$ -hopsets  $H_i$  for every  $i \in \{1, \ldots, \ell\}$ , one can compute an r-missing t-spanner G' for  $r = \beta_\ell$  and  $t = (1 + \epsilon)^\ell$  of cardinality  $|G'| \le \sum_{i=1}^\ell |H_i| \cdot \beta_{i-1}$  (where |G'| denotes the number of edges in the spanner G').

We start by describing an algorithm for computing the missing spanner G' given the collection of hopsets  $H_1, \ldots, H_\ell$ . Alg. HopsetToMissingSpanner builds the missing spanner G' in  $\ell$  consecutive steps. Let  $G'_0, H_0 = \emptyset$ . Then the output of each step  $i \ge 1$  is a subgraph  $G'_i$  computed as follows. Let  $\mathcal{P}'_i$  be collection of all  $V \times V$  ( $\beta_{i-1}$ )-bounded hop shortest paths in the graph  $G \cup H_{i-1}$  (where for each pair of vertices we take exactly one path arbitrarily). These paths can be computed by applying Alg.  $APSP^{(\beta_{i-1})}(G \cup H_{i-1})$  of Lemma 1.9. Then, for every edge  $(u, v) \in H_i$  the algorithm adds to  $G'_i$  the edges  $P_{u,v} \cap G$  where  $P_{u,v}$  is the  $\beta_{i-1}$ -hop u-v path in  $\mathcal{P}'_i$ . Note that as  $P_{u,v}$  is computed in the augmented graph  $G \cup H_{i-1}$ , the algorithm adds to  $G'_i$  only its G-edges. The output missing spanner is the output of the  $\ell^{th}$  step,  $G'_\ell$ .

### Algorithm HopsetToMissingSpanner

**Input:** An *n*-vertex  $G = (V, E, \omega)$ , hopset hierarchy of  $(\beta_i, 1 + \epsilon)$ -hopsets  $H_i, i \in \{1, ..., \ell\}$ . **Output:** A *r*-missing *t*-spanner  $G'_{\ell} \subseteq G$  for  $r = \beta_{\ell}$  and  $t = (1 + \epsilon)^{\ell}$ .

1. Set  $G', G'_0 = \emptyset$ .

2. For i = 1 to  $\ell$  do:

- (a)  $G'_i \leftarrow G'_{i-1}$ .
- (b)  $\mathcal{P}'_i \leftarrow \mathsf{APSP}^{(\beta_{i-1})}(G \cup H_{i-1})$  where  $\mathcal{P}'_i = \{P^i_{u,v} \mid (u,v) \in TC(G)\}.$
- (c) For each edge  $(u, v) \in H_i \setminus \bigcup_{j \le i-1} H_j$  do:

• 
$$G'_i \leftarrow G'_i \cup (P^i_{u,v} \cap E(G)).$$

3. Return  $G'_{\ell}$ .

We now turn to analyze the algorithm and start with the size analysis.

**Lemma 2.2.**  $|G'_{\ell}| \leq \sum_{i=1}^{\ell} |H_i| \cdot \beta_{i-1}$ .

*Proof.* We claim that in iteration *i*, the algorithm adds to  $G'_i$  at most  $|H_i| \cdot \beta_{i-1}$  edges to  $G'_i$ . This holds since for every edge  $e = (u, v) \in H_i \setminus \bigcup_{j \le i-1} H_j$ , the algorithm computes a *u*-*v* path of at most  $\beta_{i-1}$  hops, therefore adding at most  $|H_i| \cdot \beta_{i-1}$  edges.

We next show the more delicate argument that  $G'_{\ell}$  is indeed an  $\beta_{\ell}$ -missing  $(1 + \epsilon)^{\ell}$ -spanner. For ease of presentation, let  $H_{\ell+1} = TC(G)$  and  $G'_{\ell+1} = G$ . We show by induction on  $i \in \{1, \ldots, \ell+1\}$ , that  $G'_{i-1}$  is a (partial)  $\beta_{i-1}$ -missing  $(1 + \epsilon)^{i-1}$ -spanner of the subgraph  $G'_i$  but only w.r.t to u, v pairs such that  $(u, v) \in H_i$ . For  $\ell + 1$ , since  $G'_{\ell+1} = G$  and  $H_{\ell+1} = TC(G)$ , we get that  $G'_{\ell}$  is indeed a  $\beta_{\ell}$ -missing  $(1 + \epsilon)^{\ell}$ -spanner.

**Lemma 2.3.** For every  $i \in \{1, ..., \ell + 1\}$ , there is a polynomial time algorithm that given  $G'_i$  returns a collection of paths  $Q_i = \{Q^i_{u,v} \subseteq G'_i \mid (u,v) \in H_i\}$  in  $G'_i$ , such that every path  $Q^i_{u,v} \in Q_i$  satisfies:

- 1.  $\operatorname{len}(Q_{u,v}^i) \leq (1+\epsilon)^{i-1} \cdot \operatorname{dist}_G(u,v).$
- 2.  $|Q_{u,v}^i \setminus G_{i-1}'| \le \beta_{i-1}$ .

In other words,  $G'_{i-1}$  satisfies the  $\beta_{i-1}$ -missing  $(1 + \epsilon)^{i-1}$ -spanner properties for every pair  $(u, v) \in H_i$ .

*Proof.* For ease of notation, let  $\mathcal{P}'_{\ell+1} \leftarrow \mathsf{APSP}^{(\beta_\ell)}(G \cup H_\ell)$  where  $\mathcal{P}'_{\ell+1} = \{P^{\ell+1}_{u,v} \mid (u,v) \in TC(G)\}$ . We prove the lemma by induction on  $i \in \{1, \dots, \ell+1\}$ . The base case of i = 1 holds by taking  $Q^1_{u,v}$  as a shortest path in *G*. Since  $\beta_0 = n$ , we have that  $Q^1_{u,v} = P^1_{u,v} \in \mathcal{P}'_1$  and indeed  $Q^1_{u,v} \subseteq G'_1$ . Property (2) is immediate as  $\beta_0 = n$ .

Assume that the claim holds up to (i - 1), consider  $i \in \{2, ..., \ell\}$  and an edge  $(u, v) \in H_i$ . Our goal is to show that  $G'_i$  contains a path  $Q^i_{u,v}$  that satisfies the two desired properties. We define the path  $Q^i_{u,v}$  as follows. Let  $P^i_{u,v} \in \mathcal{P}'_i$  be the *u*-*v* shortest path with (at most)  $\beta_{i-1}$  hops in  $G \cup H_{i-1}$  (obtained in Step (2b)). Since  $H_{i-1}$  is an  $(\beta_{i-1}, \epsilon)$ -hopset, it holds that:

$$\operatorname{len}(P_{u,v}^{i}) = \operatorname{dist}_{G \cup H_{i-1}}^{(\beta_{i-1})}(u,v) \le (1+\epsilon)\operatorname{dist}_{G}(u,v) .$$

$$(2.2)$$

Let  $P_{u,v}^i = [u = x_1, x_2, ..., x_q = v]$  and denote  $e_j = (x_j, x_{j+1})$  for every  $j \in \{1, ..., q-1\}$ . For every  $e_j$ , define a  $x_j$ - $x_{j+1}$  path  $P'_j$ , as follows. For  $e_j \in E(G)$ , let  $P'_j = e_j$ . Otherwise, it holds that  $e_j \in H_{i-1}$ , and  $P'_j$  is defined by taking  $x_j$ - $x_{j+1}$  path  $Q_{x_j,x_{j+1}}^{i-1} \subseteq G'_{i-1}$  that belongs to the set  $Q_{i-1}$ . Since  $e_j \in H_{i-1}$ , the path  $Q_{x_j,x_{j+1}}^{i-1}$  is well defined by the induction assumption. Let  $Q_{u,v}^i = P'_1 \circ P'_2 \circ \ldots \circ P'_q$ . See Fig. 2 for an illustration of the definition of  $Q_{u,v}^i$ . We now show that  $Q_{u,v}^i \subseteq G'_i$ . Since  $G'_{\ell+1} = G$ , this holds for  $i = \ell + 1$  and it is sufficient to consider  $i \in \{2, ..., \ell\}$ . By the induction assumption for i - 1,  $G'_{i-1}$  contains the path  $Q_{x,y}^{i-1}$  for every  $(x, y) \in H_{i-1}$ . Since at the end of step  $i \leq \ell$ , the algorithm adds to  $G'_i$ , the edges in  $P_{u,v} \cap E(G)$ , and since  $G'_{i-1} \subseteq G'$ , we conclude that  $Q_{u,v}^i \subseteq G'_i$ . We next bound the length of  $Q_{u,v}^i$ .

$$\ln(Q_{u,v}^{i}) \leq \sum_{j=1}^{q} \ln(P_{j}') \leq (1+\epsilon)^{i-2} \cdot \sum_{j=1}^{q} \operatorname{dist}_{G}(x_{j}, x_{j+1}) = (1+\epsilon)^{i-2} \cdot \operatorname{len}(P_{u,v})$$
  
=  $(1+\epsilon)^{i-2} \cdot \operatorname{dist}_{G \cup H_{i-1}}^{(\beta_{i-1})}(u, v) \leq (1+\epsilon)^{i-1} \cdot \operatorname{dist}_{G}(u, v),$  (by Eq. (2.2))

where the second inequality follows by the induction assumption for i - 1. This satisfies property (1). It remains to show property (2). Recall that  $Q_{u,v}^i = P'_1 \circ P'_2 \circ \ldots \circ P'_q$  where  $q \leq \beta_{i-1}$ . By the induction assumption for i - 1, each  $P'_j$  is either contained in  $G'_{i-1}$  or else corresponds to a *G*-edge. Therefore,  $|Q_{u,v}^i \setminus G'_{i-1}| \leq \beta_{i-1}$ , and the induction step holds.

Lemma 2.1 follows by applying Lemma 2.3 for  $i = \ell + 1$  with  $G'_{\ell+1} = G$  and  $H_{\ell+1} = TC(G)$ .

We conclude this section by showing how missing spanners can be augmented to provide preservers and near-additive spanners.

**Lemma 2.4.** Given an r-missing t-spanner G' for a possibly directed and weighted graph along with a polynomial time algorithm for computing the r-missing t-approximate paths in G, one can compute a t-approximate preserver<sup>15</sup> for a given p pairs  $P \subseteq V \times V$  of size  $|G'| + p \cdot r$ .

In addition, for undirected and unweighted graphs, one can compute a  $(\alpha = t, \beta = r \cdot (2k - 1))$  spanner with  $|G'| + O(n^{1+1/k})$  edges.

<sup>&</sup>lt;sup>15</sup>Known also as *t* pairwise spanners



Figure 2: An illustration for the inductive argument of Lemma 2.3. Shown is the  $\beta_{i-1}$ -hop u-v shortest path  $P_{u,v}^i \subseteq G \cup H_{i-1}$  for some hopset edge  $(u,v) \in H_i$ . By the induction assumption, for each  $(x,y) \in H_{i-1}$  the subgraph  $G'_{i-1}$  contains a  $(1+\epsilon)^{i-1}$ -approximate shortest path denoted as  $Q_{x,y}^{i-1}$ . The final path  $Q_{u,v}^i$  is obtained by replacing each edge  $(x,y) \in P_{u,v} \setminus G$  by the path  $Q_{x,y}^{i-1}$ . This results in a  $(1+\epsilon)^i$ -approximate u-v shortest path which is fully contained in  $G'_i$ , and at most  $\beta_{i-1}$  of its edges are missing from  $G'_{i-1}$ .

*Proof.* For every pair  $u, v \in P$ , let  $P_{u,v}$  be the *r*-missing *t*-approximate *u*-*v* path in *G*. The output preserver  $G^*$  is given by  $G^* = G' \cup \bigcup_{(u,v) \in P} P_{u,v} \setminus G'$ . The size and correctness follow by Def. 2.1.

Now consider the transformation to a near-additive spanner  $G^*$ . Let  $G^* = G' \cup G''$  where G'' is a (2k - 1) multiplicative spanners (obtained by Lemma 1.10). The size bound is immediate. We consider the stretch argument for some  $u, v \in V \times V$ . Let  $P_{u,v}$  be the *r*-missing *t*-approximate *u-v* path in *G*. Then for each missing edge  $(x, y) \in P_{u,v} \setminus G'$ , the multiplicative spanner G'' provides a (2k - 1)-length path. Altogether, we have  $dist_{G^*}(u, v) \leq |P_{u,v} \cap G'| + (2k - 1) \cdot |P_{u,v} \setminus G'| \leq t \cdot dist_G(u, v) + r \cdot (2k - 1)$  as required. The lemma follows.

# 3 New Distance Preservers for Directed Graphs

In this section, we prove Theorem 1.1 and its concrete applications to directed weighted preservers, Theorem 1.2. Recall the definition of pairwise spanners (also denoted as approximate distance preservers), see Def. 1.3. Exact preservers are  $(\alpha, \beta)$  *P*-spanners for  $\alpha = 1, \beta = 0$  (namely, preserve the exact distances). In near-exact preservers, the pairwise distances are preserved up to a multiplicative factor of  $\alpha = (1 + \epsilon)$ , for  $\epsilon \in (0, 1)$ . To provide the reduction in its most general form, we consider  $(\beta, \epsilon)$  hopsets with a smooth tradeoff function of the following form. There is a threshold hopbound  $\beta^* = n^b$  such that for every  $\beta \leq \beta^*$  the  $(\beta, \epsilon)$ -hopset has  $\widetilde{O}(n^2/\beta^a)$  edges for some  $a \geq 1$ . In the following,  $\epsilon$  can be any arbitrary number (not necessarily in (0, 1)). This is important in order to also capture reachability<sup>16</sup> (where  $\epsilon = n$ ). We start by showing that given a hopset algorithm for the superlinear regime, one can derive the sublinear regime. Missing proofs are deferred to Appendix.

**Lemma 3.1** (From Superlinear Hopsets to Sublinear Hopsets). Let  $0 \le b < \min(1, 1/(a-1))$ and a > 1 be fixed parameters. Then, given an algorithm  $\mathcal{A}$  for computing  $(\beta, \epsilon)$  hopsets for *n*-vertex directed (possibly weighted) graphs with  $\widetilde{O}_{W,\epsilon}(n^2/\beta^a)$  edges for  $\beta = n^b$ , there exists an algorithm  $\mathcal{A}'$  for computing  $(\beta, \epsilon)$  hopsets for  $\beta > n^b$  with  $\widetilde{O}_{W,\epsilon}((n/\beta)^k)$  edges, where  $k = \frac{2-ab}{1-b}$ .

<sup>&</sup>lt;sup>16</sup>I.e., translating shortcuts (reachability hopsets) into reachability preservers.

*Proof.* The algorithm works in a similar manner to the sublinear hopsets provided in [KP22]. On a high level, the algorithm samples a set of landmarks vertices *L* and computes a graph N = (L, E') where edges in *E'* are added between landmarks whose shortest path has a sufficiently small number of hops. Alg. A' applies Alg. A to compute a ( $\beta^* = n^b, \epsilon$ )-hopset for the graph *N*. This is shown to also provide the desire hopset for *G* by setting the parameter carefully as follows. Set

$$q = \left(\frac{\beta}{(n \cdot \log n)^b}\right)^{1/(b-1)}$$
 and  $p = \Theta(q \cdot \log n)$ .

Let L = V[p] be a random sample of  $\Theta(n \cdot p)$  vertices of G, obtained by sampling each vertex independently into L with probability of p. For a vertex pair u, v, let h(u, v) be the minimal number of hops over all the u-v shortest path in G. I.e., the minimal number i satisfying that  $\operatorname{dist}_{G}^{(i)}(u,v) = \operatorname{dist}_{G}(u,v)$ . Consider a net graph N = (L, E') defined by the edges  $E' = \{(u,v) \mid u,v \in L, h(u,v) \leq q^{-1}\}$ . Let  $\beta' = |L|^b = \Theta((p \cdot n)^b)$ . Then, let H be an  $(\beta', \epsilon)$ -hopset for N where  $\beta' = |L|^b = \Theta((p \cdot n)^b)$ . The hopset H is computed by applying Algorithm  $\mathcal{A}$ . The output  $(\beta, \epsilon)$ -hopset for G is then given by H.

We next claim that w.h.p. *H* is indeed a  $(\beta = O(\beta' \cdot q^{-1}), \epsilon)$ -hopset for *G*. Consider a *u*-*v* shortest path  $P_{u,v}$  in *G* with  $\lceil 10\beta' \cdot q^{-1} \rceil$  hops, and let *u'* and *v'* be the closest sampled vertices in *L* to *u*, *v* respectively on  $P_{u,v}$ . Then, w.h.p.,  $h(u, u'), h(v', v) \leq q^{-1}/2$ . Let  $u' = x_0, x_1, \ldots, x_k = v'$  be the set of sampled vertices in *L* ordered based on their appearance on  $P_{u,v}$ . W.h.p., we have that  $(x_i, x_{i+1}) \in E(N)$  for every  $i \in \{0, \ldots, k-1\}$ . Therefore, the path  $P' = [x_0, x_1, \ldots, x_k] \subseteq N$ . Since *H* is a  $(\beta', \epsilon)$ -hopset for *N*, it holds that:

$$\operatorname{dist}_{N\cup H}^{(\beta')}(u',v') \leq (1+\epsilon)\operatorname{len}(P') \ .$$

As each edge in *N* corresponds to a *G*-path of  $O(q^{-1})$  hops, we get that  $G \cup H$  contains a *u*-*v* path of length at most  $(1 + \epsilon) \operatorname{len}(P)$  and with at most  $\beta = O(\beta' \cdot q^{-1})$  hops.

We now bound the size of *H*. By the guarantees of Alg. A, we have that:

$$|H| = \widetilde{O}_{W,\epsilon} \left( \frac{|L|^2}{(\beta')^a} \right)$$
  

$$= \widetilde{O}_{W,\epsilon} \left( |L|^{2-a \cdot b} \right) \qquad \text{as } \beta' = |L|^b$$
  

$$= \widetilde{O}_{W,\epsilon} \left( (p \cdot n)^{2-a \cdot b} \right) \qquad \text{as } |L| = \Theta(n \cdot p)$$
  

$$= \widetilde{O}_{W,\epsilon} \left( \left( \frac{n}{\beta} \right)^{\frac{2-a \cdot b}{1-b}} \right) \qquad \text{as } p = q \cdot \log n \text{ and } q = \left( \frac{\beta}{(n \cdot \log n)^b} \right)^{1/(b-1)}$$
  

$$= \widetilde{O}_{W,\epsilon} \left( \left( \frac{n}{\beta} \right)^k \right) \qquad \text{as } k = \frac{2-a \cdot b}{1-b}$$

We are now ready to prove Theorem 1.1.

*Proof of Theorem* 1.1. The proof considers two cases depending on the number of pairs *p*. For a *small* value of *p* the preservers are obtained by using a sequence of  $\ell = O(\log \log n)$  (sublinear)

hopsets  $H_1, \ldots, H_\ell$  obtained by applying Alg.  $\mathcal{A}'$  of Lemma 3.1 and using Lemma 2.1. For a *large* value of p the construction is more delicate, it generates first a sequence of  $2\ell$  hopsets  $H_1, \ldots, H_{2\ell}$ . The first half of this set, namely,  $H_1, \ldots, H_\ell$  have sublinear number of edges, obtained by applying Alg.  $\mathcal{A}'$ . The second half,  $H_{\ell+1}, \ldots, H_{2\ell}$  have superlinear number of edges, obtained by applying the hopset algorithm  $\mathcal{A}$  for the superlinear regime. We now describe the construction of the hopsets in details.

**Case 1:**  $p \le n^{2-a \cdot b}$ . We start with some preliminaries.

- Set  $D = \frac{n}{p^{1/k}}$  and notice that  $D \ge n^b$ . Hence  $D = n^{\alpha}$  for  $\alpha \ge b$ .
- Set  $\ell = \lceil \log \log n \rceil$  and  $\epsilon = \epsilon' / (2\ell)$ .
- The hopbound sequence  $\beta_1, \ldots, \beta_\ell$  is defined by  $\beta_i = n^{b_i}$  where  $b_i = (1 \alpha)(\frac{1}{k})^i + \alpha$ .

Note that  $\beta_{\ell} = O(n^{\alpha}) = O(D)$ . By Alg.  $\mathcal{A}'$  of Lemma 3.1, for every  $i \in \{1, \dots, \ell\}$ , one can compute a  $(\beta_i, \epsilon)$ -hopset  $H_i$  for G of cardinality  $|H_i| = \widetilde{O}_{W,\epsilon} ((n/\beta_i)^k)$ .

By applying Lemma 2.1 on the graph *G* and these  $\ell$  hopsets  $H_1, \ldots, H_\ell$ , we get an *r*-missing *t*-spanner  $G' \subseteq G$  for  $r = \beta_\ell$  and  $t = (1 + \epsilon)^\ell$  where

$$|G'| \le \sum_{i=1}^{\ell} |H_i| \cdot \beta_{i-1} = \widetilde{O}_{W,\epsilon} \left( \left(\frac{n}{D}\right)^k \cdot D \right) = \widetilde{O}_{W,\epsilon'} \left( \left(\frac{n}{D}\right)^k \cdot D \right) , \qquad (3.1)$$

where the first equality follows from the fact that by the definition of  $\beta_i$  we have for all  $1 \le i \le \ell$ ,  $(n/(\beta_i))^k \cdot \beta_{i-1} = (n/D)^k \cdot D$ , and the last inequality follows as  $\epsilon = \epsilon'/(2\ell)$ . By the definition of *G*', we have that for every  $(u, v) \in TC(G)$ , there is a *u*-*v* path  $P_{u,v} \subseteq G$  that satisfies the following:

(Q1) 
$$\operatorname{len}(P_{u,v}) \leq (1+\epsilon)^{\ell} \cdot \operatorname{dist}_G(u,v) \leq (1+\epsilon') \cdot \operatorname{dist}_G(u,v)$$
 (by Inequality (1.1)).

(Q2) 
$$|P_{u,v} \setminus G'| \leq \beta_{\ell} = O(D).$$

In addition, there is an polynomial time algorithm for computing these  $P_{u,v}$  paths. The final distance preserver  $G^*$  is given by  $G^* = G' \cup \{P_{u,v} \setminus G' \mid u, v \in P\}$ . By (Q1), we have that  $|G^*| \leq |G'| + O(Dp)$ . Using Eq. (3.1), the proof for Case 1 follows.

**Case 2:**  $p \ge n^{2-a \cdot b}$ . We start with some preliminaries.

- Set  $D_1 = n^b$  and  $D_2 = \left(\frac{n^2}{p}\right)^{1/a}$  and notice that  $D_2 \le n^b$ . Hence  $D_2 = n^{\alpha}$  for  $\alpha \le b$ .
- Set  $t = 2 \cdot \lceil \log \log n \rceil$  and  $\epsilon = \epsilon' / (2t)$ .
- For each  $0 \le i \le \ell/2$  let  $\beta_i = n^{b_i}$  where  $b_i = (1-b)(\frac{1}{k})^i + b$ .
- For each  $\ell/2 < i \le \ell$  let  $\beta_i = n^{b_i}$  where  $b_i = (b \alpha)(\frac{1}{a})^{i-\ell/2} + \alpha$ .

Note that  $\beta_{\ell/2} = O(n^b) = O(D_1)$  and  $\beta_{\ell} = O(n^{\alpha}) = O(D_2)$ .

The sequence of first  $\ell/2$  hopsets  $H_1, \ldots, H_{\ell/2}$  are obtained by computing a  $(\beta_i, \epsilon)$  hopset  $H_i$  using Alg.  $\mathcal{A}'$ . The remaining  $\ell/2$  hopsets are obtained by computing a  $(\beta_i, \epsilon)$  hopset  $H_i$  using Alg.  $\mathcal{A}$  for every  $i \in \{\ell_2 + 1, \ldots, \ell\}$ . By the size guarantees of Algorithm  $\mathcal{A}'$ , we have that  $H_i$  is a  $(\beta_i, \epsilon)$  hopset with  $|H_i| = \widetilde{O}_{W,\epsilon} \left( \left( \frac{n}{\beta_i} \right)^k \right)$  edges for every  $i \in \{1, \ldots, \ell/2$ . Similarly, by Algorithm

 $\mathcal{A}$  we have that  $|H_i| = \widetilde{O}_{W,\epsilon} \left(\frac{n^2}{\beta_i^a}\right)$  for every  $i \in \{\ell/2, \ldots, \ell\}$ . By applying Lemma 2.1 on the graph G with the  $\ell$  hopsets  $H_i$  for  $1 \leq i \leq \ell$ , we get an r-missing t-spanner G' where  $r = O(D_2)$  and  $t = (1 + \epsilon)^{\ell}$  and of cardinality:

$$|G'| \leq \sum_{i=1}^{r} |H_i(\beta_i, \epsilon)| \cdot \beta_{i-1}$$
  
=  $\widetilde{O}_{W,\epsilon} \left( \left( \frac{n}{D_1} \right)^k \cdot D_1 + \left( \frac{n^2}{D_2^a} \right) \cdot D_2 \right)$   
=  $\widetilde{O}_{W,\epsilon'} \left( \left( \frac{n}{D_1} \right)^k \cdot D_1 + \left( \frac{n^2}{D_2^a} \right) \cdot D_2 \right),$  as  $\epsilon = \epsilon'/(2t)$  (3.3)

where Identity (3.2) follows from the fact that by the definition of  $\beta_i$  we have the following:

1. For all 
$$1 \le i \le \ell/2$$
:  $\left(\frac{n}{\beta_i}\right)^k \cdot \beta_{i-1} = \left(\frac{n}{D_1}\right)^k \cdot D_1$   
2. For all  $\ell/2 < i \le \ell$ :  $\left(\frac{n^2}{\beta_i^a}\right) \cdot \beta_{i-1} = \left(\frac{n^2}{D_2^a}\right) \cdot D_2$ .

For every  $u, v \in V$ , let  $P_{u,v} \subseteq G$  be the *u*-*v* paths satisfying the properties for *G'*. Again, the output preserve  $G^*$  is obtained by taking *G'* and adding to  $G^*$  the edges  $P_{u,v} \setminus G'$  for every  $(u, v) \in P$ . We have that  $|G^*| \leq |G'| + O(D_2p)$ . The size bound follows by combining with Eq. (3.3).

Weighted and Directed Preservers (Theorem 1.2). Theorem 1.2 follows by using the algorithms for  $(\beta, \epsilon)$  hopsets of [KP22]described in Theorem 1.12. For a = 3 and b = 1/4, their algorithm computes  $(\beta, \epsilon)$  hopsets for *n*-vertex directed weighted graphs with  $\widetilde{O}_{W,\epsilon}(n^2/\beta^a)$  edges for  $\beta \le n^b$ . Theorem 1.2 follows by plugging these parameters in Theorem 1.1. We note that for  $p \ge n^{5/4}$  this matches the bounds of reachability preservers<sup>17</sup> of [AB18]. It should also be compared against the exact preservers with  $O(\min\{n\sqrt{p}, n^{2/3}p + n\})$  edges, by [CE06, BW16].

Additional Implications. Finally, we illustrate to the applicability of our reduction in additional settings for which we recover the bounds known in the literature. Let p denote the number of demand pairs throughout.

- In [KP22] it is proven that for parameters *a* = 3 and *b* = 1/3, there is an algorithm *A* for computing (β, *n*) hopsets for *n*-vertex directed unweighted graphs with Õ(n<sup>2</sup>/β<sup>a</sup>) edges for β ≤ n<sup>b</sup> (Theorem 1.11). Plugging this in Theorem 1.1 provides reachability preservers with Õ(n · p<sup>1/3</sup> + (n · p)<sup>2/3</sup>) edges. For p ≥ n, this matches the state-of-the-art bounds by [AB18].
- 2. Plugging parameters a = 2 and b = 0 in Theorem 1.1 provides exact preservers with  $\tilde{O}(np^{1/2})$  edges for directed weighted graph. This matches the Coppersmith-Elkin [CE06] bound (see Table 1.2). In this case, the given hopset algorithm simply outputs H = G. Note that interestingly applying Lemma 3.1 with the parameters a = 2 and b = 0 provides the known folklore result (see [UY91]) that there exist an  $(\beta, 0)$ -hopset of size  $\tilde{O}(n^2/\beta^2)$  for all  $1 \le \beta \le n$ .

<sup>&</sup>lt;sup>17</sup>I.e., *n*-approximate distance preservers.

# 4 New Spanners and Preservers for Undirected Graphs

In this section we employ our approach to obtain new constructions of spanners and distance preservers in undirected (but possibly weighted) graphs. Subsec. 4.1 considers the simpler unweighted setting. We show that in this setting a single (rather than a hierarchy) hopset can be converted into a near-additive spanner. Then in Subsec. 4.2 we address the more challenging weighted setting for which (pure) near-additive spanners do not exist. We show that given a hierarchy of hopsets one can compute missing spanners which can then be converted into near-exact preservers for weighted graphs (recall that previously such preservers were known only for unweighted graphs). The missing spanners can also be used to provide a weighted variant of near-additive spanners that have been addressed recently in the literature [EGN22, ABS<sup>+</sup>20b, EGN21, ABS<sup>+</sup>21].

### 4.1 Hopsets $\rightarrow$ (Unweighted) Near-Additive Spanners

We start with the following lemma that shows how to convert a near-exact hopset into a nearadditive emulator. This reduction has the benefit of preserving the universality of the structure: given a universal hopset, the output emulator is *universal* as well.

**Lemma 4.1.** [Hopsets  $\rightarrow$  Emulators] Given an **unweighted** *n*-vertex graph G = (V, E),  $a(\beta, \epsilon)$ -hopset H, for every integer  $k \ge 1$ , there is  $(1 + \epsilon, (2k - 1) \cdot \beta)$  emulator H' with  $O(|H| + n^{1+1/k})$  edges.

*Proof.* The output emulator H' is given by taking the union of H and a (2k - 1)-spanner for G. The size bound follows by Lemma 1.10, we next bound the stretch for a fixed pair  $u, v \in V$ .

First observe that a  $(\beta, \epsilon)$ -hopset H is an  $\beta$ -missing  $(1 + \epsilon)$ -spanner of the graph  $G \cup H$ . This holds as by the hopset definition, there is a u-v path  $P_{u,v}$  in  $G \cup H$  satisfying that (i) len $(P) \leq (1 + \epsilon)$ dist<sub>*G*</sub>(u, v) and (ii)  $|P| \leq \beta$ . By adding a (2k - 1)-spanner to H (of Lemma 1.10), we get that for every edge  $(x, y) \in P \cap G$ , it holds that dist<sub>*H'*</sub> $(x, y) \leq (2k - 1)$ . Since all edges of H are in the emulator, and as P consists of at most  $\beta$  edges, we get that

$$\begin{aligned} \operatorname{dist}_{H'}(u,v) &\leq & \operatorname{len}(P \cap H) + (2k-1)\operatorname{len}(P \cap G) \leq \operatorname{len}(P) + (2k-1)|P \cap G| \\ &\leq & (1+\epsilon)\operatorname{dist}_G(u,v) + (2k-1)\beta . \end{aligned}$$

The lemma follows.

By using the standard reduction from near-additive emulators to near-additive spanner (see e.g., [TZ06]), we also have the following, which makes a progress on the open problem raised by Elkin and Neiman [EN20] and proves Observation 1.5.

**Corollary 4.2.** Given an unweighted n-vertex graph G = (V, E),  $a (\beta, \epsilon)$ -hopset H and an integer  $k \ge 1$ , one can compute  $a (1 + 2\epsilon, (2k - 1)\beta)$  spanner  $G' \subseteq G$  with  $O(k \cdot (\beta/\epsilon) \cdot |H| + n^{1+1/k})$  edges.

*Proof.* Let H' = (V, E', W) be the  $(1 + \epsilon, (2k - 1)\beta)$ -emulator obtained in Lemma 4.1. The output spanner *G'* is obtained by replacing every non-*G* edge  $(x, y) \in H'$  of weight at most  $10k\beta/\epsilon$ , with its corresponding *x*-*y* shortest path in *G*. Formally, for every  $(x, y) \in H' \setminus G$ , let  $P_{x,y}$  be an *x*-*y* shortest path in *G*. Then, the output spanner *G'* is given by:

$$G' = (H' \cap G) \cup \bigcup_{(x,y) \in H', W(x,y) \le 10k\beta/\epsilon} P_{x,y}.$$

Since we replace  $H' \setminus G = H$  edges by *G*-paths of length at most  $10k\beta/\epsilon$ , the size bound follows. We now provide the stretch argument. Let  $\beta' = (2k - 1)\beta/\epsilon$ .

We start by showing the stretch guarantee for every  $u, v \in V$  pair satisfying that  $\operatorname{dist}_G(u, v) \leq 2\beta'$ . Fix such a u, v pair and let  $P'_{u,v}$  be the u-v shortest path in H'. Since H' is a  $(1 + \epsilon, (2k - 1)\beta)$ -emulator, we have that  $\operatorname{len}(P'_{u,v}) \leq (1 + \epsilon)\operatorname{dist}_G(u, v) + (2k - 1)\beta < 10k\beta/\epsilon$ . Since G' contains the path  $P_{x,y}$  for every  $(x, y) \in P'_{u,v}$ , we have that  $\operatorname{dist}_{G'}(x, y) = \operatorname{dist}_{H'}(x, y)$  for every  $(x, y) \in P'_{u,v}$ . Consequently,  $\operatorname{dist}_{G'}(u, v) = \operatorname{dist}_{H'}(u, v) \leq (1 + \epsilon)\operatorname{dist}_G(u, v) + (2k - 1)\beta$ .

We now turn to consider an arbitrary pair  $u', v' \in V$  of distance at least  $2\beta'$  in G. Let  $P_{u',v'}$  be the u'-v' shortest path in G. Partition the path  $P_{u',v'}$  into consecutive segments of length  $[\beta', 2\beta']$ . Let these segments be given by  $P_{u',v'} = P_1 \circ P_2 \circ \ldots \circ P_\ell$  and let  $(u_i, v_i)$  be the endpoints of the segment  $P_i$ . Since dist<sub>*G*</sub> $(u_i, v_i) \leq 2\beta'$ , we have that

$$\operatorname{dist}_{H'}(u_i, v_i) \leq (1+\epsilon)\operatorname{dist}_G(u_i, v_i) + (2k-1)\beta \leq (1+2\epsilon)\operatorname{dist}_G(u_i, v_i)$$

where the last inequality follows as  $\epsilon \cdot \text{dist}_G(u_i, v_i) \ge (2k - 1)\beta$ . Consequently, we have that  $\text{dist}_{H'}(u', v') \le (1 + 2\epsilon) \text{dist}_G(u', v')$ .

#### **4.2** Hopsets → Weighted Preservers and Spanners

In this section, we show the applications of our approach for weighted undirected graphs and prove Theorem 1.6. Similarly to the directed setting of Sec. 3, we first translate a hierarchy of (mostly sublinear-sizes) hopsets into missing spanners. The latter can then be converted to near-exact preservers and other variants of spanners. We start by stating the state-of-the-art hopset bounds, show how to translate them (in a black-box manner) into hopsets of *sublinear* size.

**Theorem 4.3** (Undirected Hopsets). [*EN19*, *HP19*] For any weighted graph G = (V, E) on n vertices, and any integer  $k \ge 1$ , there exists an  $(\beta, \epsilon)$ -hopset H for any  $0 < \epsilon < 1$  with  $\beta = O(k/\epsilon)^k$  and  $|H| = O(n^{1+1/(2^{k+1}-1)})$  edges. Setting  $k = \Theta(\log \log n)$ , provides linear size hopset with hopbound  $\beta = O(\log \log n/\epsilon)^{O(\log \log n)}$ .

Henceforth, we assume w.o.l.g that  $k \leq \log \log n - 1$ . In Appendix A, we show:

**Lemma 4.4** (Sublinear Undirected Hopsets). For any weighted graph G = (V, E) on n vertices, integers  $k, D \ge 1$ , there exists an  $(\beta, \epsilon)$ -hopset H for any  $0 < \epsilon < 1$  with  $\beta = O(k/\epsilon)^k \cdot D$  and  $|H| = O((n \log n/D)^{1+1/(2^{k+1}-1)})$  edges.

*Proof.* Let *L* be a random sample of  $O(n \log n/D)$  vertices, obtained by sampling each vertex independently into *L* with probability of  $p = \Theta(\log n/D)$ . Let h(u, v) be the minimal number of hops over all *u*-*v* shortest path in *G*. Consider a net graph N = (L, E') where

$$E' = \{(u, v) \mid u, v \in L, h(u, v) \le D\}.$$

Then, let *H* be an  $(\beta', \epsilon)$ -hopset for *N* taken from Theorem 4.3. We claim that w.h.p. *H* is a  $(\beta = O(\beta' \cdot D), \epsilon)$ -hopset for *G*. Consider a *u*-*v* shortest path  $P_{u,v}$  in *G* of length  $10\beta' \cdot D$ . Let *u'* and *v'* be the closest sampled vertices in *L* to *u*, *v* respectively on the path  $P_{u,v}$ .

Then, w.h.p.,  $h(u, u'), h(v', v) \le D/2$ . Let  $u' = x_0, x_1, \dots, x_k = v'$  be the set of sampled vertices in *L* ordered based on their appearance on  $P_{u,v}$ . W.h.p., we have that  $(x_i, x_{i+1}) \in E(N)$  for every

 $i \in \{0, ..., k-1\}$ . Therefore, the path  $P' = [x_0, x_1, ..., x_k] \subseteq E(N)$ . Since *H* is a  $(\beta', \epsilon)$ -hopset for *N*, it holds that:

$$\operatorname{dist}_{N\cup H}^{(\beta')}(u',v') \leq (1+\epsilon)\operatorname{len}(P')$$
.

As each edge in *N* corresponds to a *G*-path of *D* hops, we get that  $G \cup H$  contains a *u*-*v* path of length at most  $(1 + \epsilon)$ len(P) with at most  $\beta' = O(\beta D)$  hops.

The Key Step: Hopsets  $\rightarrow$  Missing Spanners (Weighted, Undirected). We are now ready to state the main sparsification lemma that computes missing spanners given an undirected hopset hierarchy. In contrast to the directed setting, the size vs. hopbound tradeoff for undirected hopsets is almost tight (up to  $n^{o(1)}$  factors). Therefore, we show how translate the *output* hopsets obtained by the state-of-the-art algorithms into missing spanners<sup>18</sup>.

**Lemma 4.5** (Hopsets  $\rightarrow$  Missing Spanners). Given  $(\beta = O(k/\epsilon)^k, \epsilon)$  hopsets from Lemma 4.4 and Theorem 4.3, one can compute an  $\beta$ -missing  $(1 + \epsilon')$ -spanner  $G' \subseteq G$  with  $|G'| = \widetilde{O}(n^{1+1/(2^{k+1}-1)} \cdot (c \cdot k/\epsilon)^{2k})$  edges, where  $\epsilon' = O(\epsilon \cdot 2^k \cdot (\log \log n - k)) = O(\epsilon \cdot \log n)$  and some constant c > 1.

*Proof.* Let  $\beta_0 = n$  and for every  $i \in \{1, \dots, \ell = \Theta(2^k \cdot (\log \log n - k))\}$  we set

$$\beta_{i} = \max\left\{ \left(\frac{1}{2}\right) \cdot \beta_{i-1}^{\left(1-2^{-k-1}\right)}, \Theta((k/\epsilon)^{k}) \right\}$$
(4.1)

We will prove now that by the definition of  $\beta_i$  we have that  $\beta_{\ell} = \Theta((k/\epsilon)^k)$ . Assume towards a contradiction that  $\beta_i = (\frac{1}{2}) \cdot \beta_{i-1}^{(1-2^{-k-1})}$  for all  $1 \le i \le \ell$ . Set  $c = 1 - 2^{-k-1}$ . Hence as  $\beta_0 = n$ , we have that for  $1 \le i \le \ell$  we have the following.

$$\beta_{i} = \left(\frac{1}{2}\right)^{\frac{1-c^{i}}{1-c}} \cdot n^{c^{i}} = \left(\frac{1}{2}\right)^{2^{k+1} \cdot (1-c^{i})} \cdot n^{c^{i}}$$
(4.2)

Hence, we have

$$\begin{split} \beta_{\ell} &= \left(\frac{1}{2}\right)^{2^{k+1} \cdot (1-c^{\ell})} \cdot n^{c^{\ell}} \\ &\leq \left(\frac{1}{2}\right)^{2^{k}} \cdot n^{c^{\ell}} \leq \left(\frac{1}{2}\right)^{2^{k}} \cdot n^{(\frac{1}{e})^{\Theta(\log\log n-k)}} \leq \left(\frac{1}{2}\right)^{2^{k}} \cdot n^{(\frac{1}{2})^{\log\log n-k}} \leq \left(\frac{1}{2}\right)^{2^{k}} \cdot 2^{2^{k}} = 1, \end{split}$$

where the first inequality follows as  $1 - c^{\ell} \ge 1 - \frac{1}{e} > 1/2$ . This contradicts the definition of Eq. (4.1), concluding that  $\beta_{\ell} = \Theta((k/\epsilon)^k)$ .

Let  $H_i$  be a  $(\beta_i, \epsilon)$ -hopset for every  $i \in \{1, ..., \ell\}$ , and let G' be the output *r*-missing *t*-spanner obtained by applying Lemma 2.1 given  $\{H_i\}$ . We then have that  $r = (k/\epsilon)^k$  and  $t = (1 + O(\epsilon \cdot 2^k \cdot (\log \log n - k)))$  (by Inequality 1.1). The cardinality of G' is bounded by  $|G'| \leq \sum_{i=1}^{\ell} |H_i| \cdot \beta_{i-1}$ . Using the bounds of Lemma 4.4 and the fact that for  $\beta_i = \omega((k/\epsilon)^k)$  we have  $(2\beta_i)^{1+1/(2^{k+1}-1)} = \beta_{i-1}$ , we get that:

$$|H_i| \cdot \beta_{i-1} = (n \log n / \beta_i)^{1+1/(2^{k+1}-1)} \cdot \beta_{i-1} \cdot (k/\epsilon)^{2k} = \widetilde{O}((n \log n)^{1+1/(2^{k+1}-1)} \cdot (k/\epsilon)^{2k}).$$
  
Therefore,  $|G'| = \widetilde{O}(n^{1+1/(2^{k+1}-1)} \cdot (k/\epsilon)^{2k}).$ 

<sup>&</sup>lt;sup>18</sup>In Sec. 3 we considered a general tradeoff function with the purpose that future algorithms that improves the Kogan and Parter bounds [KP22] could immediately derive improved preservers.

We next provide a collection of applications of Lemma 4.5 to new constructions of preservers and spanners in weighted undirected graphs.

Weighted Near-Additive Spanners and Near-Exact Preservers. The next lemma provides a transformation from near-exact hopsets to a variant of near-additive spanners adapted to weighted graphs with maximum edge weight  $W_{\text{max}}$ . The transformation is based on using missing spanners as an intermediate structure. By using a similar argument to Lemma 2.4, we have the following which proves Theorem 1.6(1):

**Lemma 4.6** (Near-Exact Weighted Hopsets  $\rightarrow$  Near-Additive Weighted Spanners). *Given* ( $\beta = O(k/\epsilon)^k, \epsilon$ ) hopsets from Lemma 4.4 and Theorem 4.3 for  $k \in \{1, ..., \log \log n\}$  and a weighted graph G = (V, E) with maximum edge weight  $W_{max}$ , one can compute  $(1 + \epsilon', \beta')$ -spanner  $G' \subseteq G$  with  $\epsilon' = O(\epsilon \cdot 2^k \cdot (\log \log n - k)), \beta' = O(\beta \cdot 2^k \cdot W_{max})$  and  $|G'| = \widetilde{O}(n^{1+1/(2^{k+1}-1)} \cdot (ck/\epsilon)^{2k})$  edges, for some constant c > 1.

*Proof.* Let  $G'' \subseteq G$  be the  $\beta$ -missing  $t = (1 + \epsilon')$ -spanner obtained by applying Theorem 4.5 given the family of hopsets. The output near-additive spanner G' is given by  $G' = G'' \cup G^*$  where  $G^* \subseteq G$  is an  $(2^{k+2} - 3)$ -multiplicative spanner for G obtained by Lemma 1.10.

Since  $|G^*| = O(n^{1+1/(2^{k+1}-1)})$ , the size bound holds. We next consider the stretch guarantee. Consider a pair  $u, v \in V$ . By the definition of the partial spanner G'', the graph G contains a u-v path  $P_{u,v}$  such that (i)  $\operatorname{len}(P_{u,v}) \leq (1+\epsilon) \cdot \operatorname{dist}_G(u,v)$  and (ii)  $|P_{u,v} \setminus G''| \leq \beta$ . Let  $E' = P_{u,v} \setminus G''$ . For each  $e = (x, y) \in E'$ , we have that:

$$\operatorname{dist}_{G'}(x,y) \leq \operatorname{dist}_{G^*}(x,y) \leq O(2^k) \cdot \operatorname{dist}_G(x,y) \leq O(2^k \cdot W_{\max}),$$

where the last inequality follows by the fact that the maximum edge weight in *G* is at most  $W_{\text{max}}$ . For unweighted graphs,  $W_{\text{max}} = 1$ . Since  $|E'| \le \beta$ , we get that

$$\operatorname{dist}_{G''}(u,v) \leq \operatorname{len}(P_{u,v}) + \sum_{(x,y)\in E'} \operatorname{dist}_{G'}(x,y) \leq (1+\epsilon')\operatorname{dist}_{G}(u,v) + O(W_{\max}\cdot\beta\cdot2^k) \ .$$

The stretch argument follows.

This can be compared with the known bounds for weighted additive weighted spanners, e.g., by Elkin [Elk01] which obtained a similar tradeoff. In addition, Elkin, Gitlitz and Neiman [EGN22] provided  $(1 + \epsilon, \beta W)$  spanners with  $\beta = O(k/\epsilon)^{k-1}$  and  $O(kn + n^{1/(3/4)^k-1})$  edges, of in which W is the maximum weight on the given *u*-*v* shortest path.

Turning to near-exact preservers, by using the transformation from missing spanners to preservers of Lemma 2.4, we get the following for weighted and undirected graphs:

**Lemma 4.7** (Near Exact Hopsets  $\rightarrow$  Near-Exact Preservers). Given  $(\beta = (k/\epsilon)^k, \epsilon)$  hopsets from Lemma 4.4 and Theorem 4.3 and a set of p demand pairs  $P \subseteq V \times V$ , one can compute a  $(1 + \epsilon')$ approximate preserver  $\widehat{G}$  for P with  $\epsilon' = O(\epsilon \cdot 2^k \cdot (\log \log n - k))$  and  $|\widehat{G}| = \widetilde{O}(n^{1+1/(2^{k+1}-1)} \cdot (ck/\epsilon)^{2k} + p \cdot \beta)$  edges, for some constant c > 1.

*Proof.* Let *G'* be a  $\beta$ -missing  $(1 + \epsilon')$ -spanner  $G' \subseteq G$  obtained by Lemma 4.5. For every  $(u, v) \in P$ , let  $P_{u,v} \subseteq G$  be a *u*-*v* path satisfying that  $|P_{u,v} \setminus G'| \leq \beta$  and  $len(P_{u,v}) \leq (1 + \epsilon')dist_G(u, v)$ . Those

paths can be computed in polynomial time based on Lemma 2.1. Then, the output preserver  $\widehat{G}$  is given by  $\widehat{G} = G' \cup \bigcup_{(u,v) \in P} (P_{u,v} \setminus G)$ .

We then have that  $P_{u,v} \subseteq \widehat{G}$  for every  $(u,v) \in P$ , and therefore  $\widehat{G}$  is indeed a  $(1 + \epsilon')$  *P*-preserver. The size bound follows as  $|\widehat{G}| \leq |G'| + p \cdot \beta$  and by plugging the bound on *G'* from Lemma 4.5.

Consequently, by setting  $k = \log \log n$ , we complete the proof of Theorem 1.3 which almost matches the bounds obtained for the unweighted case.

**Sourcewise Spanners.** Recall that for a source vertex set  $S \subseteq V$  in a (possibly weighted) graph G(V, E), an *S*-sourcewise *t* spanner is a subgraph  $H \subseteq G$  such that  $dist_H(u, v) \leq t \cdot dist_G(u, v)$  for every  $u, v \in S \times V$  (see Def. 1.3). Throughout, we denote  $\sigma = |S|$ . The following Theorem was proven in [EFN17] (formally it is stated for unweighted graphs but the proof holds for weighted graphs without any changes).

**Theorem 4.8.** [*Theorem 3 of* [EFN17]] For every  $k \ge 1$ , given an *n*-vertex (possibly weighted) graph G = (V, E) and  $S \subseteq V$ , there exists a S-sourcewise (4k - 1)-spanner with  $O(n + \sqrt{n} \cdot \sigma^{1+1/k})$  edges.

Our improved constructions of sourcewise spanners described next are obtained by simply replacing the exact preservers of [CE06] used in the proof of Theorem 4.8 with the near-exact preserver of Theorem 1.3. We prove Theorem 1.6(2) by showing:

**Theorem 4.9.** For any parameter  $k \ge 1$  and any fixed  $\epsilon > 0$ , for an n-vertex (possibly weighted) graph G = (V, E) and  $S \subseteq V$  such that  $|S| = \sigma$ , there exists a S-sourcewise (4k - 1 + o(1))-spanner with  $(n + \sigma^{1+1/k}) \cdot n^{o(1)}$  edges.

We also provide an alternative construction which strictly improves upon Theorem 4.8 of [EFN17] both in terms of the stretch bound and in the number of edges. Note that in comparison to Theorem 4.9, the obtained stretch is smaller, at the cost of increasing the size.

**Theorem 4.10.** For any parameter  $k \ge 2$  and any fixed  $\epsilon > 0$ , for an n-vertex (possibly weighted) graph G = (V, E) and  $S \subseteq V$  such that  $|S| = \sigma$ , there exists a S-sourcewise  $(4k - 5 + \epsilon)$ -spanner with  $(n + \sigma \cdot n^{1/k}) \cdot n^{o(1)}$  edges.

*Proof.* Set  $t = n^{\frac{k-1}{k}}$  and partition the set *S* into  $q = \lceil \frac{\sigma}{t} \rceil$  sets  $S_1, S_2, \ldots, S_q$  such that for all  $1 \le i \le q$  we have  $|S_i| \le t$ . Now for each  $1 \le i \le q$ , apply Theorem 4.9 with parameter k - 1 on graph *G* and the set  $S_i$ , to obtain a  $S_i$ -sourcewise  $(4k - 5 + \epsilon)$ -spanner with the following upper bound on the number of edges

$$(n + t^{1+1/(k-1)}) \cdot n^{o(1)} = n^{1+o(1)}.$$

Hence the total number of edges in the collection of spanners is  $q \cdot n^{1+o(1)} = (n + \sigma \cdot n^{1/k}) \cdot n^{o(1)}$ . The stretch argument is immediate.

Finally, we state a conditional lower bound which implies that *S*-sourcewise (2k-1) spanners require  $\Omega(|S| \cdot n^{1/k})$  edges. This is in particular interesting in light of the considerably improved size bound obtained in Theorem 4.9.

**Theorem 4.11.** [Conditional Lower Bounds] Assuming the Erdős girth conjecture, then for every  $n, \sigma$  and  $k \ge 1$  there is a bipartite graph G = (V, S, E) with |V| = n and  $|S| = \sigma$  such that  $|E| = \Omega(\sigma \cdot n^{1/k})$  and the removal of any edge (u, v) from G results in a graph G' for which dist<sub>G'</sub> $(u, v) \ge 2k + 1$ .

*Proof.* For each  $k \ge 1$ , by the Erdős girth conjecture there is a bipartite graph G' = (V, U, E') such that |V| = |U| = n with  $|E'| = \Omega(n^{1+1/k})$  and girth 2k + 2. Create a subgraph G = (V, S, E) of G' by picking each vertex in U into S independently with probability  $\frac{\sigma}{n}$ . Hence the expected size of S is  $\sigma$  and the expected number of edges in G is  $\Omega(\frac{\sigma}{n} \cdot n^{1+1/k}) = \Omega(\sigma \cdot n^{1/k})$ . The claim follows as the girth of G (and thus also of G') is at least 2k + 2.

Note that the lower bound of Theorem 4.11 implies that Theorem 4.10 is nearly tight for k = 2. **Spanners with Slack.** The notion of spanners with slack introduced by Chan, Dinitz and Gupta [CDG06] provides bounded stretch guarantees for all but a *small* fraction of the vertex pairs.

**Definition 4.1** (Spanners with Slack). Given a (possibly weighted) graph G, a subgraph  $G' \subseteq G$  is an  $\epsilon$ -slack t-spanner if for each vertex  $v \in V(G)$ , the farthest  $(1 - \epsilon)n$  vertices w from v satisfy  $\operatorname{dist}_{G'}(v, w) \leq t \cdot \operatorname{dist}_G(v, w)$ .

On a high level, the construction of  $\epsilon$ -slack *t*-spanner presented in [CDG06] is based on sampling a small sample of  $1/\epsilon$  vertices N, such that each vertex is sufficiently close to some sampled vertex. Then, the algorithm computes a multiplicative spanner H on the complete  $N \times N$  graph. The final spanner is obtained by (i) connecting each vertex to its nearest vertex in N and (ii) computing an exact preserver for the pairs P = E(H). Specifically, in the lack of near-exact preservers [CDG06] used the exact preservers of Coppersmith-Elkin [CE06]. Consequently, their work leaves the interesting gap between the size of  $\epsilon$ -slack emulators vs. spanners. Using the near-exact preservers of Theorem 1.3, we almost match this bound. Since the proof is almost identical to that [CDG06] and the only difference is in pairwise preserver that we use, we defer the complete analysis to Appendix A where we show:

**Lemma 4.12.** *Given an n-vertex (possibly weighted) graph* G = (V, E) *and parameter*  $\epsilon \in (0, 1)$ *, there is an*  $\epsilon$ *-slack*  $(5 + (6 + o(1))\alpha(1/\epsilon))$ *-spanner with*  $O(n^{1+o(1)} + T(1/\epsilon) \cdot n^{o(1)})$  *edges.* 

This should be compared with the bound of  $O(n + T(1/\epsilon) \cdot \sqrt{n})$  edges shown in [CDG06]. By plugging Lemma 1.10 into Lemma 4.12, we get the following which proves Theorem 1.6(3):

**Corollary 4.13.** For any  $\epsilon \in (0,1)$  and integer  $k \ge 1$  and n-vertex (possibly weighted) graph G, there exists an  $\epsilon$ -slack (12k - 1 + o(1))-spanner  $G^* \subseteq G$  with  $O(n^{1+o(1)} + (1/\epsilon)^{1+1/k} \cdot n^{o(1)})$  edges.

### 5 New Hopset Lower Bounds

In this section, we provide a collection of lower bound results on the hopbound of directed and undirected hopsets, and establish Theorem 1.7.

**Exact Weighted Undirected Hopsets.** We start by considering *exact* hopsets and show how to derive improved hopbound lower bounds by translating lower bounds for exact distance preservers. Specifically, we use the following theorem of [CE06].

**Theorem 5.1.** [Section 3 in [CE06]] For any sufficiently large positive integer numbers n, p where  $\Omega(\sqrt{n}) = p = O(n^2)$ , there exists a weighted undirected graph G = (V, E) and  $P \subseteq V \times V$  of cardinality |P| = p, such that any exact preserver for P requires  $\Omega((n \cdot p)^{2/3})$  edges.

We use the relation between hopsets and preservers shown in Sec. 3, in particular we need the following corollary (restatement of Case 1 in the proof of Theorem 1.1).

**Corollary 5.2.** Let  $0 \le b < \min(1, 1/(a-1))$  and a > 1 be fixed parameters. Then, given an algorithm  $\mathcal{A}$  for computing  $(\beta, \epsilon)$  hopsets for *n*-vertex directed (possibly weighted) graphs with  $\widetilde{O}_{W,\epsilon}(n^2/\beta^a)$  edges for  $\beta = n^b$ , the following holds. For any *n*-vertex directed (possibly weighted) graph G = (V, E) and a set P of  $p \le n^{2-a \cdot b}$  demand pairs, one can compute an  $(1 + \epsilon')$ -distance preserver  $G^* \subseteq G$  with

$$\widetilde{O}_{W,\epsilon'}(n \cdot p^{1-1/k})$$
 edges where  $k = \frac{2-a \cdot b}{1-b}$  and  $\epsilon' = O(\epsilon \cdot \log \log n)$ 

We then show the following for exact hopsets (i.e.,  $(\beta, 0)$  hopsets):

**Theorem 5.3.** [*Exact Hopsets, Weighted and Undirected*] For any large enough n, there exists an n-vertex weighted undirected graph for which any  $(\beta, 0)$  hopset H with linear size (i.e., O(n)) must satisfy that  $\beta \ge n^{(1/3-o(1))}$ .

*Proof.* Assume towards a contradiction that for any *n*-vertex weighted undirected graph there is a  $(\beta, 0)$  hopset *H* such that |H| = O(n) and  $\beta = n^{(1/3-\delta)}$  for some fixed  $\delta > 0$ . Hence by Corollary 5.2 with the following parameters:  $b = 1/3 - \delta$ ,  $a = b^{-1}$ , p = n and  $k = (2 - a \cdot b)/(1 - b) = 1/(1 - b)$ , we have that for any *n*-vertex undirected weighted graph G = (V, E) and a set *P* of *n* demand pairs, one can compute an exact preserver  $G^* \subseteq G$  with  $\widetilde{O}(n \cdot p^{1-1/k}) = \widetilde{O}(n \cdot n^b) = \widetilde{O}(n^{4/3-\delta})$  edges. This contradicts Theorem 5.1 for p = n, as  $(n \cdot p)^{2/3} = n^{4/3}$ .

**Exact Unweighted Undirected Hopsets.** We use the following lower bound result for exact preservers in unweighted undirected graphs by [CE06].

**Theorem 5.4.** [*The* d = 2 *case of Sec.* 4 *in* [CE06]] For any sufficiently large positive integer numbers *n*, *p* where  $\Omega(\sqrt{n}) = p = O(n)$ , there exists an unweighted undirected graph G = (V, E) and a set  $P \subseteq V \times V$  of cardinality |P| = p, such that any exact preserver w.r.t. *P* requires  $|E| = \Omega(n^{4/5} \cdot p^{2/5})$ edges.

By combining this lower bound with Corollary 5.2, we get:

**Theorem 5.5.** [Exact Hopsets, Unweighted and Undirected] For any large enough *n*, there exists an *n*-vertex **unweighted and undirected** graph for which any  $(\beta, 0)$  hopset *H* with linear number of edges must satisfy that  $\beta \ge n^{1/5-o(1)}$ .

*Proof.* Assume toward a contradiction that for any *n*-vertex unweighted undirected graph there is a  $(\beta, 0)$  hopset *H* such that |H| = O(n) and  $\beta = n^{(1/5-\delta)}$  for some fixed  $\delta > 0$ . Hence by Corollary 5.2 with the following parameters  $b = 1/5 - \delta$ ,  $a = b^{-1}$ , p = n and  $k = (2 - a \cdot b)/(1 - b) = 1/(1 - b)$ , we have that for any *n*-vertex undirected weighted graph G = (V, E) and a set *P* of *n* demand pairs, one can compute an exact preserver  $G^* \subseteq G$  with  $\widetilde{O}(n \cdot p^{1-1/k}) = \widetilde{O}(n \cdot n^b) = \widetilde{O}(n^{6/5-\delta})$  edges. This contradicts Theorem 5.4 for p = n, as  $n^{4/5} \cdot p^{2/5} = n^{6/5}$ .

**Near-Exact Hopsets.** We next show that lower bounds for near-exact spanners and emulators can be used to derive lower bounds for hopsets. Abboud, Bodwin and Pettie [ABP18] provided a collection of such lower bound results for both near-exact spanners and hopsets. While the lower bound constructions of [ABP18] obtained for hopsets have a very similar flavor to those provided for spanners and emulators, this is done in a white-box manner. The state-of-the-art lower bound results for spanners and emulators are still somewhat stronger for spanners.

We provide a black-box reduction that converts emulators (or spanners) lower bounds into hopset lower bounds. We start by stating the state-of-the-art lower bound results for near-exact hopsets and near-additive spanners by [ABP18]<sup>19</sup>.

**Theorem 5.6.** [*Theorem 4.6 of* [*ABP18*]] Any  $(\beta, \epsilon)$  hopset construction with worst-case size at most  $n^{1+1/(2^k-1)-\delta}$ ,  $\delta > 0$ , has  $\beta = \Omega(1/(\epsilon \cdot exp(k))^k)$ .

For near-additive emulators, [ABP18] provided stronger lower bound results, e.g., for larger values of  $\beta = \Omega(1/(\epsilon k)^k)$  and for *sublinear* stretch functions. As we will see, our reduction from hopsets to emulators allows us to translate in a black-box manner emulators' lower bounds into hopset lower bounds of nearly the same quality. We use the following lower bound for emulators (which should be compared with Theorem 5.6).

Theorem 5.7. [Theorem 2.11, [ABP18]] The following holds.

- 1. Any emulator with  $f(d) \leq d + O(k \cdot d^{1-1/k}) + \widetilde{O}(1)$  requires  $\Omega(n^{1+1/(2^k-1)-o(1)})$  edges.
- 2. Any  $(1 + \epsilon, \beta)$  emulator construction with worst-case size at most  $n^{1+1/(2^k-1)-o(1)}$  has  $\beta = \Omega(1/(\epsilon \cdot (k-1))^{k-1})$ .

We next derive the following which can be compared with Theorem 5.6 of [ABP18].

**Theorem 5.8.** [Improved Near-Exact Hopset Lower Bound] Any  $(\beta, \epsilon)$  hopset construction with worstcase size at most  $n^{1+1/(2^k-1)}/2$  has  $\beta \ge (d/\epsilon \cdot (k-1))^{k-1}$  for some sufficiently small constant d > 0.

*Proof.* Assume towards a contradiction that for every *n*-vertex graph, one can compute a  $(\beta, \epsilon)$  hopset *H* with at most  $n^{1+1/(2^k-1)-o(1)}/2$  edges and  $\beta < (d/\epsilon \cdot (k-1))^{k-1}$ . Then, by Lemma 4.1, one can compute a  $(1 + \epsilon, \beta')$  emulator *H'* with

$$|H'| \leq |H| + 0.5 \cdot n^{1+1/(2^k-1)-o(1)} \leq n^{1+1/(2^k-1)}$$
 and  $\beta' = 2^{k+2} \cdot \beta \leq (8d/\epsilon \cdot (k-1))^{k-1}$ .

By picking *d* to be a small enough constant, this leads to a contradiction to Thm. 5.7 (2) (I.e., Theorem 2.11, [ABP18]). ■

**Theorem 5.9.** Any hopset with  $f(d) \le d + O(kd^{1-1/k}) + \tilde{O}(1)$  and  $\beta = O(kd^{1-1/k}) + \tilde{O}(1)$  requires  $\Omega(n^{1+1/(2^k-1)-o(1)})$  edges.

*Proof.* Set  $\epsilon = k \cdot d^{-1/k}$ . Assume towards a contradiction that for every *n*-vertex graph, one can compute a  $(1 + \epsilon, O(k/\epsilon)^{k-1} + \widetilde{O}(1))$  hopset *H* with  $o(n^{1+1/(2^k-1)-o(1)})$  edges. Then, by Lemma 4.1, one can compute a  $(1 + \epsilon, \beta')$  emulator *H'* with

$$|H'| \leq |H| + 0.5 \cdot n^{1+1/(2^k-1)-o(1)} \leq n^{1+1/(2^k-1)-o(1)}$$
 and  $\beta' = O(k/\epsilon)^{k-1} + \widetilde{O}(1)$  ,

where the bound on  $\beta'$  follows as  $2^k = O(\log n)$  in our context. By picking the constants hidden in the *O*-notation to be a small enough, this leads to a contradiction to Thm. 5.7 (1) (I.e., Theorem 2.11, [ABP18]).

<sup>&</sup>lt;sup>19</sup>Theorem 4.6 of [ABP18] has a minor typo, the fixed statement below has been provided by a personal communication with the authors of [ABP18].

Theorem 1.7 follows by Theorems 5.3,5.5,5.9 and 5.8. Finally, we show that by using the general reduction of Theorem 1.1, we can also recover known lower bounds for reachability preservers.

**Lower Bounds for Reachability Preservers Imply Lower Bounds for Directed Shortcuts.** Finally, we show how to translate lower bounds for reachability preservers into lower bounds on the diameter of linear size shortcuts.

*Proof of Theorem* 1.8. Assume by contradiction that for any *n*-vertex directed graph there is a *d*-shortcut *H* such that |H| = O(n) and  $d = n^{(\alpha + \gamma - 1 - \delta)}$  for some fixed  $\delta > 0$ . Hence by Corollary 5.2 with the following parameters:

$$b = \alpha + \gamma - 1 - \delta$$
,  $a = b^{-1}$ ,  $p = n$  and  $k = \frac{2 - a \cdot b}{1 - b} = \frac{1}{1 - b}$ .

We have that for any *n*-vertex directed graph G = (V, E) and a set *P* of *n* demand pairs, one can compute a reachability preserver  $G^* \subseteq G$  with

$$\widetilde{O}(n \cdot p^{1-1/k}) = \widetilde{O}(n \cdot n^b) = \widetilde{O}(n^{\alpha+\gamma-\delta})$$

edges. This contradicts the assumed lower bound of  $\Omega(n^{\alpha} \cdot p^{\gamma})$  edges for reachability preservers (for p = n) and thus we are done.

The following lower bound for reachability preservers have been shown by Abboud and Bodwin [AB18].

**Theorem 5.10.** [*Theorem 2.6 in* [*AB18*]] *There is an infinite family of n-node directed graphs and pair* sets *P* for which every reachability preserver of *G* with respect to the set *P* requires  $|E| = \Omega(n^{2/3} \cdot p^{1/2})$  edges.

By plugging Theorem 5.10 in Theorem 1.8 (i.e., taking  $\alpha = 2/3$ ,  $\gamma = 1/2$ ), we recover the state-of-the-art lower bound by Huang and Pettie [HP21]:

**Theorem 5.11.** [*Theorem 2.1 in* [*HP21*]] For any large enough n, there exists an n-vertex directed graph for which any d-shortcut H with linear number of edges must satisfy that  $d = n^{(1/6-o(1))}$ .

**Acknowledgment.** We are grateful to Amir Abboud, Aaron Bernstein, Greg Bodwin and Nicole Wein for useful discussions.

## References

- [AB16] Amir Abboud and Greg Bodwin. Error amplification for pairwise spanner lower bounds. In Robert Krauthgamer, editor, *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016, pages 841–854. SIAM, 2016.*
- [AB17] Amir Abboud and Greg Bodwin. The 4/3 additive spanner exponent is tight. J. ACM, 64(4):28:1–28:20, 2017.

- [AB18] Amir Abboud and Greg Bodwin. Reachability preservers: New extremal bounds and approximation algorithms. In Artur Czumaj, editor, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1865–1883. SIAM, 2018.
- [ABP18] Amir Abboud, Greg Bodwin, and Seth Pettie. A hierarchy of lower bounds for sublinear additive spanners. *SIAM J. Comput.*, 47(6):2203–2236, 2018.
- [ABS<sup>+</sup>20a] Abu Reyan Ahmed, Greg Bodwin, Faryad Darabi Sahneh, Keaton Hamm, Mohammad Javad Latifi Jebelli, Stephen G. Kobourov, and Richard Spence. Graph spanners: A tutorial review. *Comput. Sci. Rev.*, 37:100253, 2020.
- [ABS<sup>+</sup>20b] Abu Reyan Ahmed, Greg Bodwin, Faryad Darabi Sahneh, Stephen G. Kobourov, and Richard Spence. Weighted additive spanners. In Isolde Adler and Haiko Müller, editors, Graph-Theoretic Concepts in Computer Science - 46th International Workshop, WG 2020, Leeds, UK, June 24-26, 2020, Revised Selected Papers, volume 12301 of Lecture Notes in Computer Science, pages 401–413. Springer, 2020.
- [ABS<sup>+</sup>21] Abu Reyan Ahmed, Greg Bodwin, Faryad Darabi Sahneh, Keaton Hamm, Stephen G. Kobourov, and Richard Spence. Multi-level weighted additive spanners. In David Coudert and Emanuele Natale, editors, 19th International Symposium on Experimental Algorithms, SEA 2021, June 7-9, 2021, Nice, France, volume 190 of LIPIcs, pages 16:1– 16:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.
- [ACIM99] Donald Aingworth, Chandra Chekuri, Piotr Indyk, and Rajeev Motwani. Fast estimation of diameter and shortest paths (without matrix multiplication). SIAM J. Comput., 28(4):1167–1181, 1999.
- [ADDJ90] Ingo Althöfer, Gautam Das, David P. Dobkin, and Deborah Joseph. Generating sparse spanners for weighted graphs. In John R. Gilbert and Rolf G. Karlsson, editors, SWAT 90, 2nd Scandinavian Workshop on Algorithm Theory, Bergen, Norway, July 11-14, 1990, Proceedings, volume 447 of Lecture Notes in Computer Science, pages 26–37. Springer, 1990.
- [Alo02] Noga Alon. Testing subgraphs in large graphs. *Random Struct. Algorithms*, 21(3-4):359–370, 2002.
- [BCE03] Béla Bollobás, Don Coppersmith, and Michael Elkin. Sparse distance preservers and additive spanners. In Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA, pages 414–423. ACM/SIAM, 2003.
- [BKMP05] Surender Baswana, Telikepalli Kavitha, Kurt Mehlhorn, and Seth Pettie. New constructions of (alpha, beta)-spanners and purely additive spanners. In *Proceedings of* the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2005, Vancouver, British Columbia, Canada, January 23-25, 2005, pages 672–681. SIAM, 2005.
- [Bod17] Greg Bodwin. Linear size distance preservers. In Philip N. Klein, editor, Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19, pages 600–615. SIAM, 2017.

- [Bod21] Greg Bodwin. New results on linear size distance preservers. *SIAM J. Comput.*, 50(2):662–673, 2021.
- [BP20] Uri Ben-Levy and Merav Parter. New ( $\alpha$ ,  $\beta$ ) spanners and hopsets. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020,* pages 1695–1714. SIAM, 2020.
- [BW16] Greg Bodwin and Virginia Vassilevska Williams. Better distance preservers and additive spanners. In Robert Krauthgamer, editor, Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016, pages 855–872. SIAM, 2016.
- [CC20] Shiri Chechik and Sarel Cohen. Distance sensitivity oracles with subcubic preprocessing time and fast query time. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, *Proceedings of the 52nd Annual* ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020, pages 1375–1388. ACM, 2020.
- [CDG06] T.-H. Hubert Chan, Michael Dinitz, and Anupam Gupta. Spanners with slack. In Yossi Azar and Thomas Erlebach, editors, *Algorithms - ESA 2006, 14th Annual European Symposium, Zurich, Switzerland, September 11-13, 2006, Proceedings,* volume 4168 of Lecture Notes in Computer Science, pages 196–207. Springer, 2006.
- [CE06] Don Coppersmith and Michael Elkin. Sparse sourcewise and pairwise distance preservers. *SIAM J. Discret. Math.*, 20(2):463–501, 2006.
- [CFR20] Nairen Cao, Jeremy T. Fineman, and Katina Russell. Efficient construction of directed hopsets and parallel approximate shortest paths. In Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy, editors, Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020, Chicago, IL, USA, June 22-26, 2020, pages 336–349. ACM, 2020.
- [CGK13] Marek Cygan, Fabrizio Grandoni, and Telikepalli Kavitha. On pairwise spanners. In Natacha Portier and Thomas Wilke, editors, 30th International Symposium on Theoretical Aspects of Computer Science, STACS 2013, February 27 - March 2, 2013, Kiel, Germany, volume 20 of LIPIcs, pages 209–220. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2013.
- [Che13] Shiri Chechik. New additive spanners. In Sanjeev Khanna, editor, *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013, pages 498–512. SIAM, 2013.*
- [Coh00] Edith Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest paths. *J. ACM*, 47(1):132–166, 2000.
- [EFN17] Michael Elkin, Arnold Filtser, and Ofer Neiman. Terminal embeddings. Theor. Comput. Sci., 697:1–36, 2017.

- [EGN21] Michael Elkin, Yuval Gitlitz, and Ofer Neiman. Improved weighted additive spanners. In Seth Gilbert, editor, 35th International Symposium on Distributed Computing, DISC 2021, October 4-8, 2021, Freiburg, Germany (Virtual Conference), volume 209 of LIPIcs, pages 21:1–21:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.
- [EGN22] Michael Elkin, Yuval Gitlitz, and Ofer Neiman. Almost shortest paths with nearadditive error in weighted graphs. In Artur Czumaj and Qin Xin, editors, 18th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2022, June 27-29, 2022, Tórshavn, Faroe Islands, volume 227 of LIPIcs, pages 23:1–23:22. Schloss Dagstuhl -Leibniz-Zentrum für Informatik, 2022.
- [Elk01] Michael Elkin. Computing almost shortest paths. In Ajay D. Kshemkalyani and Nir Shavit, editors, Proceedings of the Twentieth Annual ACM Symposium on Principles of Distributed Computing, PODC 2001, Newport, Rhode Island, USA, August 26-29, 2001, pages 53–62. ACM, 2001.
- [EM21] Michael Elkin and Shaked Matar. Deterministic PRAM approximate shortest paths in polylogarithmic time and slightly super-linear work. In Kunal Agrawal and Yossi Azar, editors, SPAA '21: 33rd ACM Symposium on Parallelism in Algorithms and Architectures, Virtual Event, USA, 6-8 July, 2021, pages 198–207. ACM, 2021.
- [EN16] Michael Elkin and Ofer Neiman. Hopsets with constant hopbound, and applications to approximate shortest paths. In Irit Dinur, editor, IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA, pages 128–137. IEEE Computer Society, 2016.
- [EN17] Michael Elkin and Ofer Neiman. Efficient algorithms for constructing very sparse spanners and emulators. In Philip N. Klein, editor, Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19, pages 652–669. SIAM, 2017.
- [EN19] Michael Elkin and Ofer Neiman. Linear-size hopsets with small hopbound, and constant-hopbound hopsets in RNC. In Christian Scheideler and Petra Berenbrink, editors, *The 31st ACM on Symposium on Parallelism in Algorithms and Architectures*, SPAA 2019, Phoenix, AZ, USA, June 22-24, 2019, pages 333–341. ACM, 2019.
- [EN20] Michael Elkin and Ofer Neiman. Near-additive spanners and near-exact hopsets, A unified view. *CoRR*, abs/2001.07477, 2020.
- [EP04] Michael Elkin and David Peleg. (1+epsilon, beta)-spanner constructions for general graphs. *SIAM J. Comput.*, 33(3):608–631, 2004.
- [EP16] Michael Elkin and Seth Pettie. A linear-size logarithmic stretch path-reporting distance oracle for general graphs. *ACM Trans. Algorithms*, 12(4):50:1–50:31, 2016.
- [Erd65] Paul Erdös. On some extremal problems in graph theory. *Israel Journal of Mathematics*, 3(2):113–116, 1965.
- [Fin20] Jeremy T. Fineman. Nearly work-efficient parallel algorithm for digraph reachability. *SIAM J. Comput.*, 49(5), 2020.

- [FN18] Sebastian Forster and Danupon Nanongkai. A faster distributed single-source shortest paths algorithm. In Mikkel Thorup, editor, 59th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2018, Paris, France, October 7-9, 2018, pages 686–697.
   IEEE Computer Society, 2018.
- [Hes03] William Hesse. Directed graphs requiring large numbers of shortcuts. In *Proceedings* of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 12-14, 2003, Baltimore, Maryland, USA, pages 665–669. ACM/SIAM, 2003.
- [HKN14] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. Sublinear-time decremental algorithms for single-source reachability and shortest paths on directed graphs. In David B. Shmoys, editor, Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014, pages 674–683. ACM, 2014.
- [HP19] Shang-En Huang and Seth Pettie. Thorup-zwick emulators are universally optimal hopsets. *Inf. Process. Lett.*, 142:9–13, 2019.
- [HP21] Shang-En Huang and Seth Pettie. Lower bounds on sparse spanners, emulators, and diameter-reducing shortcuts. *SIAM J. Discret. Math.*, 35(3):2129–2144, 2021.
- [Kav15] Telikepalli Kavitha. New pairwise spanners. In Ernst W. Mayr and Nicolas Ollinger, editors, 32nd International Symposium on Theoretical Aspects of Computer Science, STACS 2015, March 4-7, 2015, Garching, Germany, volume 30 of LIPIcs, pages 513–526. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2015.
- [KLX22] Nicole Wein Kevin Lu, Virginia Vassilevska Williams and Zixuan Xu. Better lower bounds for shortcut sets and additive spanners via an improved alternation product. In Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2022, 2022.
- [KP22] Shimon Kogan and Merav Parter. New diameter-reducing shortcuts and directed hopsets: Breaking the  $\sqrt{n}$  barrier. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA, 2022, 2022.*
- [KS97] Philip N. Klein and Sairam Subramanian. A randomized parallel algorithm for single-source shortest paths. *J. Algorithms*, 25(2):205–220, 1997.
- [KV13] Telikepalli Kavitha and Nithin M. Varma. Small stretch pairwise spanners. In Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg, editors, *Automata, Languages, and Programming - 40th International Colloquium, ICALP 2013, Riga, Latvia, July 8-12, 2013, Proceedings, Part I, volume 7965 of Lecture Notes in Computer Science*, pages 601–612. Springer, 2013.
- [KV15] Telikepalli Kavitha and Nithin M. Varma. Small stretch pairwise spanners and approximate d-preservers. *SIAM J. Discret. Math.*, 29(4):2239–2254, 2015.
- [LJS19] Yang P. Liu, Arun Jambulapati, and Aaron Sidford. Parallel reachability in almost linear work and square root depth. In David Zuckerman, editor, 60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019, pages 1664–1686. IEEE Computer Society, 2019.

- [Par14] Merav Parter. Bypassing erdős' girth conjecture: Hybrid stretch and sourcewise spanners. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, Automata, Languages, and Programming 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part II, volume 8573 of Lecture Notes in Computer Science, pages 608–619. Springer, 2014.
- [Pet09] Seth Pettie. Low distortion spanners. *ACM Trans. Algorithms*, 6(1):7:1–7:22, 2009.
- [PS89] David Peleg and Alejandro A. Schäffer. Graph spanners. J. Graph Theory, 13(1):99–116, 1989.
- [RTZ05] Liam Roditty, Mikkel Thorup, and Uri Zwick. Deterministic constructions of approximate distance oracles and spanners. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings, volume 3580 of Lecture Notes in Computer Science, pages 261–272. Springer, 2005.
- [SN21] Idan Shabat and Ofer Neiman. A unified framework for hopsets and spanners. *CoRR*, abs/2108.09673, 2021.
- [SS99] Hanmao Shi and Thomas H. Spencer. Time-work tradeoffs of the single-source shortest paths problem. *J. Algorithms*, 30(1):19–32, 1999.
- [Tho92] Mikkel Thorup. On shortcutting digraphs. In Ernst W. Mayr, editor, Graph-Theoretic Concepts in Computer Science, 18th International Workshop, WG '92, Wiesbaden-Naurod, Germany, June 19-20, 1992, Proceedings, volume 657 of Lecture Notes in Computer Science, pages 205–211. Springer, 1992.
- [TZ06] Mikkel Thorup and Uri Zwick. Spanners and emulators with sublinear distance errors. In Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2006, Miami, Florida, USA, January 22-26, 2006, pages 802–809. ACM Press, 2006.
- [UY91] Jeffrey D. Ullman and Mihalis Yannakakis. High-probability parallel transitiveclosure algorithms. *SIAM J. Comput.*, 20(1):100–125, 1991.
- [Woo06] David P. Woodruff. Lower bounds for additive spanners, emulators, and more. In 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings, pages 389–398. IEEE Computer Society, 2006.
- [Zwi02] Uri Zwick. All pairs shortest paths using bridging sets and rectangular matrix multiplication. *J. ACM*, 49(3):289–317, 2002.

### **A** Missing Proofs

*Proof of Theorem* 4.9. Consider the complete graph G' on the set of vertices S, where each edge  $(s,s') \in E(G')$  is weighted by the *s*-*s'* shortest path distance in G. Let H' by a (2k - 1)-spanner H' of G' obtained by using Lemma 1.10. We then apply the algorithm of Theorem 1.3 to compute a  $(1 + \epsilon = o(1))$  distance preserver  $G'' \subseteq G$  with the pair set P = E(H'), hence p = |E(H')|. Altogether, we get that G'' is a *S*-subsetwise  $(2k - 1)(1 + \epsilon)$ -spanner for some  $\epsilon = o(1)$ . I.e., for every  $s, s' \in S$  it holds that  $\operatorname{dist}_{G''}(s, s'') \leq (2k - 1)(1 + \epsilon)\operatorname{dist}_G(s, s'')$ . Finally, the output sourcewise spanner  $G^*$  is obtained by computing a shortest path tree T rooted at a dummy vertex  $s^*$  in a graph  $\widehat{G} = (V \cup \{s^*\}, E \cup \{(s^*, s) \mid s \in S\}, W')$  where W'(e) = W(e) for every  $e \in E(G)$  and  $W'((s^*, s)) = 1$  for every  $s \in S$ . Define  $G^* = G'' \cup (T \cap E(G))$ . Adding the edges of  $T \cap E(G)$  guarantees that the spanner  $G^*$  contains for each vertex v the shortest path from v to its closest source in S. It is easy to see that  $|G^*| \leq |G''| + n = O(n^{o(1)} \cdot (n + \sigma^{1+1/k}))$  by Lemma 4.7.

We next bound the *s*-*u* distance in  $G^*$  for  $s, u \in S \times V$ . Let  $s_u$  be the closest vertex to *u* in *S*, then

$$\operatorname{dist}_G(s_u, s) \leq \operatorname{dist}_G(s_u, u) + \operatorname{dist}_G(u, s) \leq 2\operatorname{dist}_G(u, s)$$
.

We therefore have:

where the second inequality follows as  $s_u$  is closer to u than s, and by the fact that G'' is a *S*-subsetwise  $(2k-1)(1+\epsilon)$ -spanner.

**Proof of Lemma 4.12.** The construction of succinct spanners with slack is based on the notion of density nets. For completeness, we provide the following definitions. Let  $B(x, r) = \{y \mid \text{dist}_G(x, y) \le r\}$  and define  $R(x, \epsilon)$  to be the minimum distance r such that  $|B(x, r)| \ge \epsilon n$ .

**Definition A.1** (Density Net). For  $\epsilon \in (0, 1)$ , an  $\epsilon$ -net is a set  $N \subseteq V$  such that  $|N| \leq 1/\epsilon$  and for all  $x \in V$  there exists  $y \in N$  such that  $\text{dist}_G(x, y) \leq 2R(x, \epsilon)$ .

The following lemma, implicit in [CDG06], provides the general recipe for computing  $\epsilon$ -slack spanners using density nets.

**Lemma A.1** (Implicit in [CDG06]). *Given algorithms for computing*  $\alpha(n)$ *-spanners with* T(n) *edges and*  $\gamma$ *-multiplicative pairwise spanners for* p *pairs with* D(p) *edges, there is an algorithm for computing an*  $\epsilon$ *-slack*  $(5 + 6\alpha(1/\epsilon) \cdot \gamma)$ *-spanner*  $G^*$  *with*  $O(n + D(T(1/\epsilon)))$  *edges.* 

*Proof.* We provide a proof sketch, see Theorem 5 of [CDG06] for additional details. Let *N* be an  $\epsilon$ -net for *G*, and let *H'* be an  $\alpha(1/\epsilon)$ -spanner for the graph  $\hat{G} = (N, N \times N, W)$  where  $W((z, z')) = \text{dist}_G(z, z')$ . Notes that *H'* is not necessarily a subgraph of *G*.

Let G' be the set of edges connecting each vertex  $v \in V$  to its nearest vertex in N. Note that |G'| = O(n). Finally, the output spanner  $G^*$  is obtained by taking G' and a  $\gamma$ -multiplicative pairwise spanner G'' for the pairs of E(H'). Since  $p = |H'| = T(1/\epsilon)$ , the size bound follows.

For the stretch argument, consider a pair u, v where  $v \notin B(u, \epsilon)$  and let u', v' be their nearest vertices in the set N. By the definition of N, we have that  $\operatorname{dist}_G(u, u') \leq 2R(u, \epsilon) \leq 2\operatorname{dist}_G(u, v)$  and  $\operatorname{dist}_G(v, v') \leq 3\operatorname{dist}_G(u, v)$ . Also,  $\operatorname{dist}_{G^*}(u', v') \leq \alpha(1/\epsilon) \cdot \gamma$ . Consequently, we have that  $\operatorname{dist}_{G^*}(u, v) \leq (5 + 6\alpha(1/\epsilon) \cdot \gamma)\operatorname{dist}_G(u, v)$ .

The construction presented in [CDG06] is based on the using the exact preservers of [CE06], leading  $\epsilon$ -slack (5 + 6 $\alpha$ (1/ $\epsilon$ ))-spanner with  $O(n + T(1/\epsilon)\sqrt{n})$  edges. Lemma 4.12 follows by plugging the near-exact preservers of Theorem 1.3 in Lemma A.1.