

# Parallel Breadth-First Search and Exact Shortest Paths and Stronger Notions for Approximate Distances

Václav Rozhoň\*

ETH Zurich

rozhoňv@inf.ethz.ch

Ⓡ<sup>†</sup>

Bernhard Haeupler ‡

Carnegie Mellon University &amp; ETH Zurich

bernhard.haeupler@inf.ethz.ch

Anders Martinsson

ETH Zurich

anders.martinsson@inf.ethz.ch

Ⓡ

Christoph Grunau \*

ETH Zurich

cgrunau@inf.ethz.ch

Ⓡ

Goran Zuzic

ETH Zurich

goran.zuzic@inf.ethz.ch

## Abstract

We introduce stronger notions for approximate single-source shortest-path distances, show how to efficiently compute them from weaker standard notions, and demonstrate the algorithmic power of these new notions and transformations. One application is the first work-efficient parallel algorithm for computing exact single-source shortest paths graphs — resolving a major open problem in parallel computing.

Given a source vertex in a *directed graph* with polynomially-bounded nonnegative integer lengths the algorithm computes an *exact shortest path tree* in  $m \log^{O(1)} n$  work and  $n^{1/2+o(1)}$  depth. Previously, no parallel algorithm improving the trivial linear depths of Dijkstra’s algorithm without significantly increasing the work was known, even for the case of undirected and unweighted graphs (i.e., for computing a BFS-tree).

Our main result is a black-box transformation that uses  $\log^{O(1)} n$  standard approximate distance computations to produce approximate distances which also satisfy the subtractive triangle inequality (up to a  $(1 + \varepsilon)$  factor) and even induce an exact shortest path tree in a graph with only slightly perturbed edge lengths. These strengthened approximations are algorithmically significantly more powerful and overcome well-known and often encountered barriers for using approximate distances. In directed graphs they can even be boosted to exact distances. This results in a black-box transformation of any (parallel or distributed) algorithm for approximate shortest paths in directed graphs into an algorithm computing exact distances at essentially no cost. Applying this to the recent breakthroughs of Fineman et al. for compute approximate SSSP-distances via approximate hopsets gives new parallel and distributed algorithm for exact shortest paths.

---

\*Supported by the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No. 853109).

<sup>†</sup>The author ordering was randomized using <https://www.aeaweb.org/journals/policies/random-author-order/generator>. It is requested that citations of this work list the authors separated by `\textcircled{r}` instead of commas

<sup>‡</sup>Supported in part by NSF grants CCF-1814603, CCF-1910588, NSF CAREER award CCF-1750808, a Sloan Research Fellowship, funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (ERC grant agreement 949272), and the Swiss National Foundation (project grant 200021-184735).

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	The Breadth-First Search and Shortest-Path Problems . . . . .	2
1.2	Prior Work and Approximate SSSP Algorithms . . . . .	2
1.3	Issues with Standard Notions of Distance Approximations: Monotonicity and Subtractive Triangle Inequality . . . . .	4
1.4	Stronger Notions of Distance Approximations . . . . .	5
1.5	Our Results . . . . .	7
1.5.1	Directed Graphs . . . . .	7
1.5.2	Undirected Graphs . . . . .	8
<b>2</b>	<b>Related Work</b>	<b>9</b>
2.1	Parallel and Distributed Algorithms for Approximate Shortest Paths . . . . .	10
<b>3</b>	<b>Preliminaries and Notation</b>	<b>10</b>
3.1	Boosting for directed SSSP . . . . .	11
<b>4</b>	<b>Different Notions of Approximate Distances</b>	<b>12</b>
4.1	Graph Potentials . . . . .	12
4.2	Smoothly Approximate Distances . . . . .	13
4.3	Tree-Like Distances . . . . .	14
4.4	Strongly approximate distances . . . . .	16
4.5	Do Known Algorithms give Strong Distance Guarantees? . . . . .	18
<b>5</b>	<b>The Smoothing Algorithm</b>	<b>19</b>
5.1	Intuition Behind the Algorithm . . . . .	19
5.2	The Formal Analysis . . . . .	22
<b>6</b>	<b>Tree-Constructing Algorithm</b>	<b>27</b>
6.1	Ball Cutting . . . . .	28
6.2	Constructing the Tree . . . . .	30
<b>A</b>	<b>What Can and Cannot be Done with Strong Distances</b>	<b>38</b>
A.1	A Success Story: Ball Growing . . . . .	38
A.2	A Cautionary Tale: MPX . . . . .	39
<b>B</b>	<b>Undirected APSP Variant of Smoothing</b>	<b>40</b>
<b>C</b>	<b>Applications in Different Computational Models</b>	<b>49</b>
C.1	Directed Graphs . . . . .	49
C.2	Undirected Graphs . . . . .	50

# 1 Introduction

This paper gives the first *work-efficient* parallel algorithms with *sublinear depth* for computing the *exact* distances and shortest paths in weighted graphs — resolving major open problems in parallel computing and breaking a long-standing barrier where no such algorithm was known even for breadth-first search (BFS). The simplest version of our main results are easy to state formally:

**Theorem 1.1.** *There exists a parallel algorithm that takes as input any directed or undirected graph  $G = (V, E)$  with  $m = |E|$  edges,  $n = |V|$  nodes, a length function  $\ell$  that assigns to every edge  $e$  a non-negative polynomially-bounded integer lengths  $\ell(e)$ , and any source vertex  $s \in V$ . The algorithm computes for every vertex the exact distance to  $s$  and an exact shortest path tree rooted at  $s$ . The algorithm is randomized, requires  $m \log^{O(1)} n$  work,  $n^{1/2+o(1)}$  depth, and works with high probability.*

Note that computing a breadth-first-search (BFS) tree corresponds to the much simpler special case of a shortest-path tree in an undirected graph where all edge lengths are one.

We develop our results in a surprisingly indirect way: we develop *stronger notions of approximate distances* that are algorithmically more powerful than the standard notions used throughout the literature. We give efficient black-box reductions computing these stronger approximations using only polylogarithmic black-box calls to algorithms computing the (weaker) *standard* distance approximations.

For directed graphs, our stronger approximations can be efficiently boosted to exact distances [KS97]. Using the parallel shortest-path approximation algorithm of [CFR20a] in our reductions yields Theorem 1.1.

For undirected graphs, much faster work-efficient parallel approximation algorithms with polylogarithmic depth are known [Li20, ASZ20, RGH<sup>+</sup>22]. While we do not know how to use these results to develop fast and work-efficient exact distance algorithms on undirected graphs, we show how to compute the next best thing: an exact shortest path tree in a slightly perturbed graph.

**Theorem 1.2.** *There exists a parallel algorithm that takes as input any undirected graph  $G = (V, E)$  with  $m = |E|$  edges,  $n = |V|$  nodes, a length function  $\ell$  that assigns to every edge  $e$  a non-negative polynomially-bounded integer lengths  $\ell(e)$ , any source vertex  $s \in V$ , and any  $\varepsilon \in (0, 1]$ . The algorithm computes a  $(1 + \varepsilon)$ -perturbed length function  $\ell'$ , which satisfies  $\ell(e) \leq \ell'(e) \leq (1 + \varepsilon)\ell(e)$  for every edge  $e \in E$ , and outputs  $\ell'$  together with exact distances to  $s$  and an shortest path tree rooted at  $s$  under this new length function  $\ell'$ . The algorithm is deterministic, requires near-linear work of  $m \cdot \left(\frac{\log n}{\varepsilon}\right)^{O(1)}$ , and polylogarithmic depth of  $\left(\frac{\log n}{\varepsilon}\right)^{O(1)}$ .*

While our headline result is for directed graphs, we expect the newly-developed techniques — stronger notions of approximate distances together with the near-optimal algorithms to compute them — to be conceptually and algorithmically useful even in undirected graphs. Specifically, many (undirected) graph-theoretic structures and algorithms are crucially built on top of exact distances and our concepts often allow approximate distances (satisfying our additional guarantees) to be used. This enables leveraging state-of-the-art parallel algorithms for approximate distances which are work-efficient and have near-optimal polylogarithmic depth, a quality that seems out of reach if one requires exact distances.

## Organization

We give some motivation and background on parallel algorithms for the breadth-first-search and shortest path problems in Section 1.1, and give a brief summary of related work in Section 1.2. We explain key insights in the subtle problems making standard distance approximations hard to use algorithmically in Section 1.3. These issues directly motivate the definitions of our stronger notions of distance approximation given in Section 1.4. Section 1.5 gives a more detailed account of our results. This is followed by the technical part of the paper starting with preliminaries in Section 3.

### 1.1 The Breadth-First Search and Shortest-Path Problems

Big-data analytics on petabyte-sized network representations and data-intensive computations on graphs have become pervasive and their scale keeps exploding exponentially. Indeed, many applications involve graph abstractions with billions of edges and massively-parallel systems with millions of cores working on these graphs.

Unfortunately even very basic graph algorithmic sub-routines can be notoriously hard to parallelize. The most poignant and common example is simple breadth-first search (BFS). While a linear-time BFS implementation for a single core taking  $O(m)$  time to process any  $m$ -edge graph is trivial, no parallel algorithm reducing the time using more cores without drastically increasing the total work is known — even for undirected graphs.

The complete lack of algorithms with provable guarantees for this basic problem is particularly shocking given the decades-long continued theoretical and practical interest in this problem [HNR68, UY91, Spe91, Tho92, STV95, Coh96, Coh00, TZ05, FSR06, SB13, MBBC07, DBS17, AFGW10, US12, SSP<sup>+</sup>14, Fin18, CFR20b, CFR20a, CFR21, KT19, EN19, Li20, ASZ20, HP21, RGH<sup>+</sup>22, KP22b]. For example, the well-known Graph500 benchmark [MWBA10], which is used to rank supercomputers, consists of merely two problems, firstly computing a BFS tree on graphs with up to  $2^{45}$  nodes, and secondly computing single-source shortest paths (SSSP), i.e., essentially a weighted BFS.

The single-source shortest-path problem (SSSP) asks to compute the minimum-length path from a source vertex to all other vertices in a graph with positive edge lengths. It is one of the most important optimization problems in computer science and parallel SSSP algorithms have been intensely studied for decades but also particularly recently, owing to the increased practical need for graph algorithms suitable for modern massively-parallel systems.

While Dijkstra’s classic 1956 algorithm [Dij59] computes *exact* distances and SSSP trees in directed graphs with lengths in essentially linear work, it does so in an inherently sequential way. Again, the only known algorithms improving over the trivial linear depth of Dijkstra’s algorithm do so at the cost of significantly increasing the total work and the best known time-work trade-offs remain the algorithms of Ullman & Yannakakis [UY91] and Spencer [Spe91] from the 90’s with  $\tilde{O}(m\rho + \frac{\rho^4}{n})$  and  $\tilde{O}(m + n\rho^2)$  work for a depth of  $O(n/\rho)$  for any  $\rho \in [1, n]$ .

### 1.2 Prior Work and Approximate SSSP Algorithms

Given the difficulty of designing parallel BFS and SSSP algorithms, research and algorithm design efforts have turned to heuristics [HNR68, FSR06] and algorithms with provable guarantees for restricted graph classes or other special cases, including planar graphs [HKRS97, FR06], random graphs [FR84, MS98], or graphs arising in road networks [AFGW10]. Another intensely-pursued

research direction has been the design of parallel and distributed algorithms that compute approximate instead of exact distances [Coh00, RGH<sup>+</sup>22, ZGY<sup>+</sup>22, ASZ20, Li20, HL18, BFKL21]. Approximation algorithms for SSSP are most closely related and relevant to this paper and we focus the remaining discussion of related work on this direction.

The intuition why approximate distances are much easier to compute is that they break up the rigid long-range interactions that seem to make BFS or exact SSSP computations so inherently sequential and hard to parallelize. In particular, even the most minor length or distance change on a single node or edge potentially causes not just all vertices its subtree to require distance updates but can completely changing the structure of the BFS or SSSP tree. Even for (unit length) undirected graphs, this makes it very hard to do any meaningful BFS computations in parts of a graph far away from the source without being sure which node will be reached first as even the most minor change can completely change the outcome. In the approximate setting one can often avoid such rigid long-range interactions by computing approximate solutions that have build-in slacks that can absorb minor updates or changes.

Many ideas have been brought forward to utilize this flexibility. Overall, the quest for parallel and distributed SSSP approximation algorithms has been a tremendous inspiration for innovation resulting in a wide variety of fundamental and powerful paradigms and structures that have become a crucial part of the modern algorithm design toolbox. Among many others these include (approximate) hopsets [Coh00, EN19], spanners and emulators [TZ06, HP19], low-diameter decompositions [MPX13, BGK<sup>+</sup>14, EHRG22], diameter-reducing shortcuts [Tho92, HP21, KP22b], and distance oracles [TZ05, Che14]. The search for better parallel approximation algorithms for the shortest-path problem has also lead to a mutually beneficial exchange of ideas and cross-fertilization between continuous and discrete optimization. With research borrowing and further developing ideas from continuous optimization, such as the multiplicative weight methods [Li20, ASZ20, BFKL21, RGH<sup>+</sup>22, ZGY<sup>+</sup>22]. Many of the above tools have been specifically developed for parallelizing shortest-path computations and all state-of-the-art work-efficient parallel algorithms for computing approximate distances crucially utilize many of these tools.

For undirected graphs the state of the art is given by the algorithms of [RGH<sup>+</sup>22, Li20, ASZ20] and for directed graphs the algorithm of Cao, Fineman, and Russel [CFR20a]. These algorithms all compute  $(1+\varepsilon)$ -approximations using near-linear work of  $\tilde{O}(\varepsilon^{-O(1)}m)$ . The algorithm for undirected graphs achieve a near-optimal polylogarithmic depth of  $(\frac{\log n}{\varepsilon})^{O(1)}$  and the algorithm of [CFR20a] for directed graphs has depth  $n^{1/2+o(1)}$ .

These algorithms of [RGH<sup>+</sup>22] and [CFR20a] also have distributed message-passing implementations (i.e., in the standard CONGEST model of distributed computing) — another intensely studied related area of work with tremendous recent activity and progress.

We conclude our brief summary by noting that basically all existing tools and algorithms seem inherently limited to approximation algorithms (with at best a polynomial dependency on  $\varepsilon$  for  $(1+\varepsilon)$ -approximations). Moreover, the vast majority of structures, tools, and algorithms seem inherently restricted to undirected graphs. These limitations explains the complete lack of work-efficient parallel or distributed algorithms for exact shortest paths as well as the large gap between what is known for approximation algorithms in undirected graphs versus directed graph.

### 1.3 Issues with Standard Notions of Distance Approximations: Monotonicity and Subtractive Triangle Inequality

While approximate distances provide flexibilities that makes them easier to compute, these flexibilities cause several subtle issues if we try to use approximate distances as algorithmic tools. This holds even for quite precise  $(1 + \varepsilon)$ -approximations. We discuss these issues next.

The most immediate notion of approximation when discussing the SSSP problem on a graph  $G$  is to ask for a distance estimate  $(\tilde{d}(v))_v$  that  $\alpha$ -approximates the exact distances from the source to each node  $v \in V(G)$ . We formalize the notion below.

**Definition 1.3** (Shortest-Path Distances). Let  $G = (V, E, \ell)$  be a graph with edge lengths  $\ell : E \rightarrow \mathbb{R}_{\geq 0}$ . We denote with  $d_G(u, v)$  the shortest-path distance between  $u$  and  $v$  in the graph  $G$ .

**Definition 1.4** ( $\alpha$ -Approximate Distance Estimate). An arbitrary function  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  is called a distance estimate. Given a graph  $G = (V, E, \ell)$  and a node  $s \in V$ , a distance estimate  $\tilde{d} : V \rightarrow \mathbb{R}_{\geq 0}$  is  $\alpha$ -approximate (with respect to the source  $s$ ) if the following two conditions are satisfied:

$$\forall u \in V : d_G(s, u) \leq \tilde{d}(u) \tag{1.1}$$

$$\forall u \in V : \tilde{d}(u) \leq \alpha \cdot d_G(s, u). \tag{1.2}$$

We sometimes refer to the first condition as noncontractivity. When  $\alpha = 1$ , the inequalities become equalities and we then call  $\tilde{d}$  *exact* distances (from  $s$ ).

Unfortunately, while  $(1 + \varepsilon)$ -approximations seem might seem “almost as good as exact” on first glance, replacing exact distance computations by approximate ones often completely breaks algorithms and correctness proofs in subtle and unexpected ways. This is already true when restricted to unweighted and undirected graphs. We identify two main culprits that break many algorithms which build upon (exact) shortest path computations if one were to replace exact computations with  $(1 + \varepsilon)$ -approximate ones: subtractive triangle inequality and non-monotonicity of distances along shortest paths.

**Issue: subtractive triangle inequality.** In particular, if we label each node  $u$  with its exact distance  $d(u)$  to  $s$ , then for any edge  $e = \{u, v\}$  we have that  $|d(u) - d(v)| \leq 1$  in unweighted graphs and at most  $\ell(e)$  in weighted graphs. For  $(1 + \varepsilon)$ -approximate distances  $\tilde{d}$ , the value  $|\tilde{d}(u) - \tilde{d}(v)|$  can essentially be unboundedly large (e.g., as large as  $\varepsilon n$  in unweighted graphs). The subtractive triangle inequality is crucially used in many algorithms: examples include the MPX algorithm [MPX13] for low-diameter decompositions (see Section A.2 for a discussion), boosting approximate to exact distances in directed graphs [KS97], or Bourgain’s algorithm for embedding a graph into the  $\ell_1$  metric with low distortion [LLR95].

**Issue: non-monotonicity.** The other problem is lack of monotonicity along shortest paths. In particular, approximate distances can have local minima and can be non-monotone along every single path from  $s$  to  $u$ . We illustrate these issues on the following simple example.

**Example: ball growing.** Let  $G$  be an undirected (and even unweighted) graph,  $s \in V(G)$  be a node, and  $R > 0$  be some radius. Let  $r$  be uniformly randomly chosen in  $[R/2, R]$  and let  $B(s, r) = \{u \in V \mid d_G(s, u) \leq r\}$  be the ball around  $s$  with radius  $r$ . It is easy to see that  $B$  is indeed a “nice ball”: e.g., the induced subgraph  $G[B]$  is connected and has diameter  $r$ . Furthermore any

(unit) edge  $\{u, v\}$  in the graph is cut by this ball with probability at most  $\frac{|d_G(s,u) - d_G(s,v)|}{R/2} = O(1/R)$ , giving us a convenient way to bound the number of cut edges.

If we replace  $d$  by a  $(1 + \varepsilon)$ -approximate distance  $\tilde{d}$  then  $B$  can be disconnected and connected components in  $G[B]$  can have arbitrarily large diameter due to the non-monotonicity. There is also no guarantee that an edge is cut with a small probability: an edge  $\{u, v\}$  is cut with probability at most  $\frac{|\tilde{d}(u) - \tilde{d}(v)|}{R/2}$  which can be as large as  $\Theta(\frac{\varepsilon n}{R})$  due to the absence of triangle inequality. Hence, the expected number of cut edges might be much larger compared to the exact case.

These above issues with approximate distances in general and the ball growing example in particular have been noted and pointed out in many prior works on approximate SSSP including among others in [BEL20, FN18, Li20]. The most explicit treatment is the paper by Becker, Emek, and Lenzen [BEL20] which solely and explicitly concerned with the problems the lack of an (approximate) subtractive triangle inequality causes. They show, with lot of effort, that weaker versions of the above “ball cutting” can be achieved algorithmically by designing a randomized “blurry ball growing” technique which can tolerate approximate distances. A slightly simpler and deterministic equivalent of this was presented in [EHRG22]. The notions and algorithms of this paper can be seen as a more general and powerful way to solve problems like the above (see Section A.1 for details).

## 1.4 Stronger Notions of Distance Approximations

An important conceptual contribution of this paper is the analysis of the following strengthened notions of approximate distance aimed squarely at avoiding the just discussed issues. We first introduce two properties, smoothness and tree-likeness, that strengthen the conditions (1.2) and (1.1) in the definition of  $\alpha$ -approximate distance estimates in Definition 1.4.

**Definition 1.5** (Smoothness and Tree-Likenes). Given a graph  $G = (V, E, \ell)$  and (so-called source)  $s \in V$ , we say a distance estimate  $\tilde{d} : V \rightarrow \mathbb{R}_{\geq 0}$  is:

- **$\alpha$ -smooth** (w.r.t.  $s$ ) iff both  $\tilde{d}(s) = 0$  and  $\forall u, v \in V : \tilde{d}(v) - \tilde{d}(u) \leq \alpha \cdot d_G(u, v)$ ,
- **tree-like** (w.r.t.  $s$ ) iff both  $\tilde{d}(s) = 0$  and  $\forall v \neq s : \exists e = (u, v) \in E : \tilde{d}(u) \leq \tilde{d}(v) - \ell(e)$ .

Next, we define the following strengthened notions of  $\alpha$ -approximate distances.

**Definition 1.6** ( $\alpha$ -approximate distances). Given a graph  $G = (V, E, \ell)$  and (so-called source)  $s \in V$ , we say a distance estimate  $\tilde{d} : V \rightarrow \mathbb{R}_{\geq 0}$  is:

- **smoothly  $\alpha$ -approximate** (w.r.t.  $s$ ) if we replace (1.2) in Definition 1.4 by the  $\alpha$ -smooth condition.
- **strongly  $\alpha$ -approximate** (w.r.t.  $s$ ) if we replace both (1.1) by the tree-like condition and (1.2) by the  $\alpha$ -smooth condition in Definition 1.4.

An example showing how smoothly  $\alpha$ -approximate distances strengthen  $\alpha$ -approximate distances is given in Figure 1. Figure 2 shows another diagram explaining the definitions from Definition 1.6.

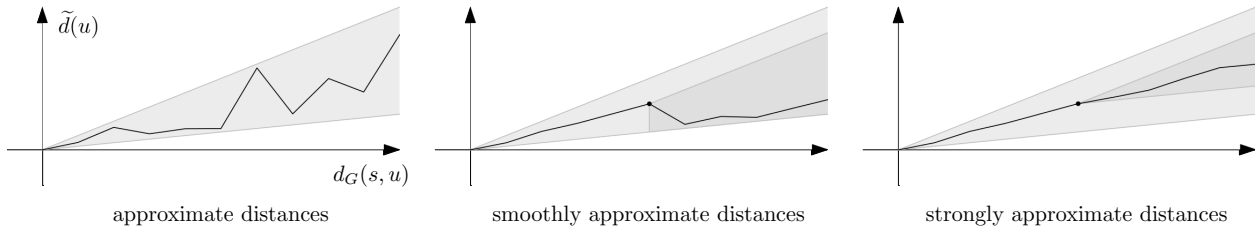
We remark that for our results for directed graphs (e.g., Theorem 1.1), only the  $\alpha$ -smoothness property is required. On the other hand, for our results in undirected graphs it is crucial that a single distance function simultaneously satisfies both properties of a strongly  $\alpha$ -approximate distance simultaneously. Only then does one obtain the powerful characterization of strongly  $\alpha$ -approximate distances used in Theorem 1.2, i.e., strongly  $\alpha$ -approximate approximate distances are exactly equivalent to exact distances induced by a slightly perturbed edge-length function. A detailed discussion

1.01-approximate

smoothly 1.01-approximate



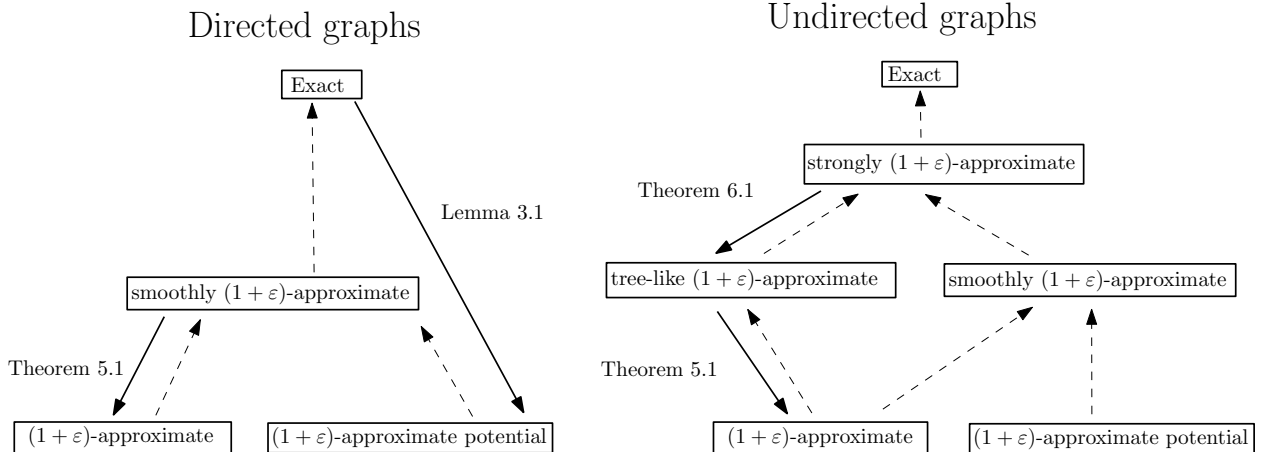
**Figure 1:** Example showing possible  $\alpha$ -approximate and smoothly  $\alpha$ -approximate distance estimates (estimates are blue). Note that in the  $\alpha$ -approximate case (left picture), the distance estimate can increase abruptly over a short edge. This is forbidden if the estimate is smoothly  $\alpha$ -approximate (right picture).



**Figure 2:** The diagram shows an example of approximate distance estimates in case the input graph is an unweighted oriented path and the source is its first node. In all three cases we highlight two rays of slopes 1 and  $\alpha$ , distance estimates need to lie within the grey cone that the rays define.

- 1) In the first picture ( $\alpha$ -approximate estimate), the values are arbitrary within the grey cone.
- 2) In the second picture (smoothly  $\alpha$ -approximate estimate), the distances cannot increase abruptly. Namely, the values of distances after the highlighted node can lie only in the darker area. Distance estimate can, however, drop abruptly.
- 3) In the third picture (strongly  $\alpha$ -approximate estimate), the distances cannot decrease abruptly and they also increase at least at the rate of exact distances (see the cone after a highlighted node). Observe that if we perturb the edge lengths of the input graph multiplicatively by at most  $\alpha$ , the distance estimate becomes exact (see Lemma 4.9).





**Figure 3:** The figure shows the relations between the various stronger approximate distance notions that we consider in this paper. The dashed arrows show the trivial inclusions while the solid arrows show reductions of stronger notions to simpler ones.

**Directed graphs:** [KS93] (Lemma 3.1) shows how to reduce the exact distance computation to the computation of approximate potential. Since smoothly approximate distances strengthen approximate potentials, our reduction from smoothly approximate distances to approximate distances in Theorem 5.1 then reduces the exact distance computation to approximate one.

**Undirected graphs:** We define the notion of strongly approximate distances that combine smoothness with tree-likeness. Theorem 5.1 reduces the computation of strongly approximate distances to the computation of tree-like approximate distances. Next, Theorem 6.1 reduces the latter problem further down to the problem of computing approximate distances.

and formal treatment of these notions and how they compare, e.g., to standard notions of approximation, is provided in Section 4.

## 1.5 Our Results

Our main technical contribution is a procedure for computing smoothly  $(1 + \varepsilon)$ -approximate distance estimates using  $O(\log n)$  calls to an approximate distance oracle. The procedure works both for directed and undirected graphs. However, the implications of this transformation are quite different in the directed and undirected setting. We therefore split the result section into a directed and undirected part. Our reductions from this section are also explained in Figure 3.

### 1.5.1 Directed Graphs

Our main result is an algorithm that turns approximate distances into smoothly approximate ones.

**Theorem 1.7** (Computing smoothly  $(1 + \varepsilon)$ -approximate distance estimates in directed graphs). *Given a directed graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , and accuracy  $\varepsilon \in (0, 1]$ , Algorithm 3 computes smoothly  $(1 + \varepsilon)$ -approximate distance estimates from  $s$  in  $G$  using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate distance oracle  $\mathcal{O}$  on directed graphs.*

By using a reduction of Klein and Subramanian [KS97], one can compute exact distances with  $O(\log n)$  calls to a smoothly 2-approximate distance estimate oracle. To make this paper self-contained and to better match the language we use in this paper, we give a self-contained treatment of this reduction in Section 3.1. By combining this reduction from exact distances to smoothly

approximate distance estimates with our reduction from smoothly approximate distance estimates to approximate distance estimates, we obtain an efficient reduction reducing exact distances to the computation of  $(1 + O(1/\log n))$ -approximate distance estimates, modulo some low-level technical details that we handle and discuss in more detail in Section C. This reduction allows us to turn any parallel algorithm for computing  $(1 + O(1/\log n))$ -approximate distance estimates in directed graphs into a parallel algorithm for computing exact distances with almost no overhead. In a recent breakthrough, Cao, Fineman and Russell gave a parallel algorithm for computing  $(1 + O(1/\log n))$ -approximate distances with  $\tilde{O}(m\rho^2 + n\rho^4)$  work and  $n^{1/2+o(1)}/\rho$  depth, for a given parameter  $\rho \in [1, n^{1/2}]$  [CFR20a]. We therefore obtain as a corollary the following result.

**Corollary 1.8.** *[Directed SSSP, PRAM] There exists a parallel algorithm that takes as input any directed graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\rho \in [1, n^{1/2}]$ . The algorithm computes for every vertex the exact distance to  $s$  and an exact shortest path tree rooted at  $s$ . The algorithm is randomized, requires  $\tilde{O}(m\rho^2 + n\rho^4)$  work and  $n^{1/2+o(1)}/\rho$  depth, and works with high probability.*

In a similar vein, using the state-of-the-art CONGEST algorithm for computing  $(1 + O(1/\log n))$ -approximate distance estimates by Cao, Fineman and Russell [CFR21] and observing that their algorithm works if the source node is a virtual node connected to all other nodes in the graph, we obtain the following result.

**Corollary 1.9.** *[Directed SSSP, CONGEST] There exists a distributed algorithm that takes as input any directed graph  $G$  with nonnegative polynomially bounded integer edge lengths and any source vertex  $s \in V(G)$ . The algorithm computes for every vertex the exact distance to  $s$  and an exact shortest path tree rooted at  $s$ . The algorithm is randomized, runs in  $\tilde{O}(n^{1/2} + D + n^{2/5+o(1)}D^{2/5})$  rounds where  $D$  denotes the undirected hop-diameter of  $G$ , and works with high probability.*

For some range of parameters, this result improves upon the previous fastest exact shortest path CONGEST algorithm with a round complexity of  $\tilde{O}(\sqrt{n}D^{1/4} + D)$  [CM20] by a polynomial factor.

## 1.5.2 Undirected Graphs

We next present our results on undirected graphs. As mentioned above, our procedure for computing smoothly  $(1 + \varepsilon)$ -approximate distance estimates also works for undirected graphs. Moreover, the reduction preserves tree-likeness. That is, if the oracle returns tree-like approximate distances, then the final distance estimate will also be tree-like, on top of being smoothly  $(1 + \varepsilon)$ -approximate, and therefore by definition strongly  $(1 + \varepsilon)$ -approximate.

**Theorem 1.10** (Computing smoothly  $(1 + \varepsilon)$ -approximate distance estimates in undirected graphs while preserving tree-likeness). *Given an undirected graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , and accuracy  $\varepsilon \in (0, 1]$ . Algorithm 3 computes smoothly  $(1 + \varepsilon)$ -approximate distance estimates from  $s$  in  $G$  using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate distance oracle  $\mathcal{O}$  on undirected graphs. Moreover, if  $\mathcal{O}$  returns tree-like  $(1 + \varepsilon)$ -approximate distance estimates from  $s$  in  $G$ , then Algorithm 3 computes strongly  $(1 + \varepsilon)$ -approximate distance estimates from  $s$  in  $G$ .*

We note that we can get a similar result in the more complicated all pairs shortest paths setting, we leave the statement and the proof to Section B.

We also give a procedure for computing tree-like  $(1 + \varepsilon)$ -approximate distance estimates which only relies on approximate distance estimates.

**Theorem 1.11** (Computing tree-like  $(1 + \varepsilon)$ -approximate distance estimates in undirected graphs).

Given an undirected graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , and accuracy  $\varepsilon \in (0, 1]$ . Algorithm 7 computes tree-like  $(1 + \varepsilon)$ -approximate distance estimates from  $s$  in  $G$  using  $O(\log^2 n)$  calls to a  $(1 + O(\varepsilon/\log n))$  approximate distance oracle  $\mathcal{O}$  on undirected graphs.

By combining the two reductions, we obtain a procedure for turning approximate distance estimates into strongly approximate distance estimates, modulo some low-level details that we discuss in Section C.

By using the recent deterministic parallel and distributed approximate shortest path algorithms of Rozhon (R) Grunau (R) Haeupler (R) Zuzic (R) Li [RGH<sup>+</sup>22], we then obtain the following two corollaries.

**Corollary 1.12.** [*Undirected Strong Distance Estimates, PRAM*] *There exists a parallel algorithm that takes as input any undirected graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\varepsilon \in (0, 1]$ . The algorithm computes a strongly  $(1 + \varepsilon)$ -approximate distance estimate from  $s$ . The algorithm is deterministic, requires  $m \cdot \left(\frac{\log n}{\varepsilon}\right)^{O(1)}$  work and depth  $\left(\frac{\log n}{\varepsilon}\right)^{O(1)}$ .*

**Corollary 1.13.** [*Directed Strong Distance Estimates, CONGEST*] *There exists a distributed algorithm that takes as input any undirected graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\varepsilon \in (0, 1]$ . The algorithm computes a strongly  $(1 + \varepsilon)$ -approximate distance estimate from  $s$ . The algorithm is deterministic and runs in  $(\sqrt{n} + D) \left(\frac{\log n}{\varepsilon}\right)^{O(1)}$  rounds.*

## 2 Related Work

The related work for the shortest path problem is vast, hence this section only tries to focus on the most relevant papers.

**Exact and Directed SSSP.** The current state of literature suggests that techniques for computing exact techniques are intricately tied to directed graphs (up to gaps that we close in this paper). For the sequential case, Dijkstra’s algorithm gives an essentially-optimal solution. In the parallel setting, work-span trade-offs were given by Ullman&Yannakakis [UY91] and Spencer [Spe91] from the 90’s with  $\tilde{O}(m\rho + \frac{\rho^4}{n})$  and  $\tilde{O}(m + n\rho^2)$  work for a depths of  $O(n/\rho)$  for any  $\rho \in [1, n]$ . This was improved by Klein and Subramanian [KS97] who gave a parallel  $\tilde{O}(\sqrt{n})$ -depth and  $\tilde{O}(m\sqrt{n})$ -work exact SSSP algorithm by showing how to boost the approximate variant of the problem to the exact one. More recently, there have been improvements in construction for directed hopsets [Fin18, JLS19, CFR20a], an important combinatorial structure used for most parallel and distributed algorithms, culminating in a parallel  $(1 + \varepsilon)$ -approximate SSSP with work-span tradeoff of  $\tilde{O}(m\rho^2 + n\rho^4)$  work and  $n^{1/2+o(1)}/\rho$  depth for  $\rho \in [1, \sqrt{n}]$ . Exciting progress in directed hopset construction started by Kogan and Parter [KP22b, KP22a, BW22] has recently broken the  $\sqrt{n}$ -threshold, with their construction being the bottleneck that prevents them from improving on the state-of-the-art.

In the distributed setting (specifically, for the standard message-passing model CONGEST), the first sublinear algorithm for exact directed SSSP was given by Elkin with a  $\tilde{O}(n^{5/6} + n^{2/3}D^{1/3})$ -round CONGEST algorithm, where  $D$  is the hop-diameter of the graph. Forster and Nanongkai [FN18] gave two algorithms taking  $\tilde{O}(\sqrt{nD})$  and  $\tilde{O}(\sqrt{nD}^{1/4} + n^{3/5} + D)$  rounds, while Ghaffari and Li [GL18] gave a  $\tilde{O}(n^{3/4+o(1)}D^{1/6} + D)$ -round and  $\tilde{O}(n^{6/7} + D)$ -round algorithms, both improving on Elkin’s

bound. Chechik and Muhtar [CM20] gave a  $\tilde{O}(\sqrt{n}D^{1/4} + D)$ -round algorithm for exact directed SSSP. Finally, for directed  $(1 + \varepsilon)$ -approximate SSSP, Cao, Russell, and Fineman [CFR21] gave a  $\tilde{O}(\sqrt{n} + n^{2/5+o(1)}D^{2/5} + D)$ -round algorithm.

**Approximate Undirected SSSP.** The progress for approximate distance computation on undirected graphs has significantly outpaced the progress on directed or exact ones. The breakthrough work of Cohen [Coh00] gave an  $m^{1+o(1)}$ -work and  $\text{poly}(\log n)$ -depth algorithm. This was the first truly sublinear work-efficient parallel algorithm (somewhat mirroring our result), and was achieved by introducing the (undirected) *hopset*, a combinatorial structure that spurred significant research interest [EN19, HP19, BLP20, EN20]. The state-of-the-art for  $(1 + \varepsilon)$ -SSSP in PRAM is given by the randomized algorithms of Andoni etc [ASZ20] and Li [Li20] for undirected graphs with  $\tilde{O}(\frac{1}{\varepsilon^2}m)$  work and  $(\frac{\log n}{\varepsilon})^{O(1)}$  depth. A deterministic algorithm with the same guarantees was given by Rozhon [Roz20], Grunau [Gru20], Haeupler [Hae20], Zuzic [Zuz20], Li [Li20] [RGH<sup>+</sup>22]. In the distributed setting (specifically, for the standard message-passing model CONGEST), Becker, Forster, Karrenbauer, and Lenzen [BFKL21] gave an existentially-optimal  $\tilde{O}(D + \sqrt{n})$  algorithm for  $(1 + \varepsilon)$ -SSSP, where  $D$  is the hop-diameter of the graph. This was improved to a universally-optimal algorithm (one that is as fast as possible on any network topology) by Zuzic [Zuz20], Goranci [Gor20], Ye [Ye20], Haeupler [Hae20], Sun [Sun20] [ZGY<sup>+</sup>22], and made deterministic in the aforementioned [RGH<sup>+</sup>22].

**Work-efficient parallel algorithms in special graph classes.** For Erdos-Renyi random graphs, Crauser, Mehlhorn, Meyer, and Sanders [CMMS98] gave a work-efficient  $\tilde{O}(n^{1/3})$ -depth exact SSSP. For planar and genus-bounded directed graphs, Klein and Subramanian [KS93] gave a parallel work-efficient  $\tilde{O}(\sqrt{n})$ -depth exact SSSP algorithm. For bounded-treewidth graphs, Chaudhuri and Zaroliagis [CZ98] gave a work-efficient  $O(\log^2 n)$ -depth algorithm for exact SSSP.

## 2.1 Parallel and Distributed Algorithms for Approximate Shortest Paths

The investigation into efficient computation of approximate distances in the parallel and distributed setting has flourished and been a tremendous source of fundamental ideas. These ideas include

## 3 Preliminaries and Notation

Let  $G$  be a directed or an undirected graph with  $n := |V(G)|$  nodes and  $m := |E(G)|$  edges. We use  $N_G^{\text{in}}(v), N_G^{\text{out}}(v)$  to denote the in- and out-neighborhood of a node  $v$  in  $G$ . For undirected graphs,  $N_G^{\text{in}}(v) = N_G^{\text{out}}(v)$ . All our graphs are weighted. More precisely, they come with a length function  $\ell_G : E(G) \mapsto \mathbb{R}_{\geq 0}$  which assigns each edge  $e \in E(G)$  a nonnegative length  $\ell_G(e)$ ; if  $e = (u, v)$ , we simplify the notation and write  $\ell_G(u, v)$ . However, most of our results need to assume that the input graph has only **positive integer** lengths that are polynomially bounded, albeit this property might be lost throughout the algorithm. We will make it clear in our formal statements when do we assume this is the case. We also define the **maximum distance**  $\Delta(G) := \max_{u, v \in V(G)} d_G(u, v)$ , or just  $\Delta$  when  $G$  is clear from context.

Unless otherwise stated, all of our results hold for both directed and undirected graphs. In the latter case, we think of undirected graphs as directed graphs where each undirected edge is replaced by two opposite directed edges with the same weight.

### 3.1 Boosting for directed SSSP

In this section, we show how to *boost* smoothly  $\alpha$ -approximate distance estimates to exact distances with  $O(\alpha \log n)$  calls to the approximate oracle (and an insignificant amount of additional processing). In other words, finding an exact solution for SSSP on directed graphs is essentially of the same hardness as finding an (appropriate) approximate solution. Moreover, this reduction is fairly general in that it works in (at least) sequential, parallel, and even distributed models.

The core of the method is the following graph transformation: given a directed graph  $G$  and a smoothly  $\alpha$ -approximate distance estimate  $\tilde{d}$  (with respect to a source  $s$ ), we define a new directed graph  $G'$  by changing edge length of each edge  $e = (u, v)$  to  $\ell'(e) := \ell(e) - \tilde{d}(v)/\alpha + \tilde{d}(u)/\alpha$ . It is easy to show that the new graph  $G'$  retains nonnegative edge weights due to  $\alpha$ -smoothness of  $\tilde{d}$ . Furthermore, the shortest path of  $G'$  tree is also the same as the one for  $G$ . Various versions of this method have appeared throughout the literature [HNR68, Joh77, BNWN22].

Algorithm 1 uses the transformation to boost any approximate oracle to an exact solution: we show that repeatedly finding a smoothly  $\alpha$ -approximate distance estimate  $\tilde{d}$  via the oracle and transforming the graph using  $\tilde{d}$  yields a (nonnegative) graph where the shortest paths from the source  $s$  to all other nodes are 0, making the (exact) shortest path tree trivial to find. Since all of the graphs share a common shortest path tree (property of the transformation), this is sufficient to solve exact SSSP.

---

**Algorithm 1** Constructing exact distances from  $\alpha$ -approximate ones for directed graphs

---

**Input:** A directed graph  $G$  with weights  $\mathbb{Z}_{\geq 0}$  and maximum distance upper bound  $\Delta$ , source  $s \in V(G)$ , and an approximate oracle  $\mathcal{O}$ .

**Oracle:**  $\mathcal{O}(G, s)$  returns a smoothly  $\alpha$ -approximate estimate  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  from  $s$  in  $G$ .

**Output:** Exact distances  $\tilde{d}' : V(G) \rightarrow \mathbb{Z}_{\geq 0}$  from  $s$  in  $G$ .

- 1:  $\ell_1 \leftarrow \ell_G$  be the edge lengths of  $G = (V, E, \ell_G)$
  - 2: **for**  $i \in 1, 2, \dots, I$ , where  $I \leftarrow O(\alpha \log(n\Delta))$  **do**
  - 3:      $\tilde{d}_i \leftarrow \mathcal{O}((V, E, \ell_i), s)$
  - 4:      $\forall e = (u, v) \in E, \ell_{i+1}(e) \leftarrow \ell_i(e) - \tilde{d}_i(v)/\alpha + \tilde{d}_i(u)/\alpha$
  - 5: **return**  $\tilde{d}'$  defined by  $\tilde{d}'(v) := \sum_{i=1}^I \tilde{d}_i(v)/\alpha$ , rounded to the nearest integer.
- 

**Lemma 3.1** (Implicit in [KS97]). *Given an  $n$ -node directed graph  $G$  with nonnegative integer weights and maximum distance  $\Delta$ , a source  $s \in V(G)$ , and access to an oracle  $\mathcal{O}$  which returns smoothly  $\alpha$ -approximate distance estimates from  $s$ , Algorithm 1 computes the exact distances from  $s$  to each node in  $G$ . The oracle  $\mathcal{O}$  is called  $O(\alpha \log(n\Delta))$  times.*

*Proof.* We note that all, throughout the algorithm, edge lengths are nonnegative since  $\ell'(e) = \ell(e) - \tilde{d}(v)/\alpha + \tilde{d}(u)/\alpha \geq 0$  can be rewritten as  $\tilde{d}(v) - \tilde{d}(u) \leq \alpha \cdot \ell(e)$ , which is implied by  $\alpha$ -smoothness. Moreover, note that for any path from  $s$  the identity  $\ell_{i+1}(u, v) = \ell_i(u, v) - \tilde{d}_i(v)/\alpha + \tilde{d}_i(u)/\alpha$  telescopes. Formally, writing  $G(i) := (V, E, \ell_i)$  and telescoping across the shortest path, we get:

$$d_{G(i+1)}(s, u) = d_{G(i)}(s, u) - \tilde{d}_i(u)/\alpha + \tilde{d}_i(s)/\alpha = d_{G(i)}(s, u) - \tilde{d}_i(u)/\alpha. \quad (3.1)$$

Using the fact that  $\tilde{d}$  is  $\alpha$ -approximate we have  $\tilde{d}_i(u) \geq d_{G(i)}(s, u)$  (for this proof, we only use noncontractivity of  $\tilde{d}$ , with  $\alpha$ -smoothness this is equivalent to being  $\alpha$ -approximate). We argue that

$\sum_{u \in V} d_{G^i}(s, u)$  drops by a factor of  $(1 - 1/\alpha)$  in each iteration:

$$\sum_{u \in V} d_{G^{(i+1)}}(s, u) = \sum_{u \in V} [d_{G^{(i)}}(s, u) - \tilde{d}_i(u)/\alpha] \leq (1 - \frac{1}{\alpha}) \sum_{u \in V} d_{G^{(i)}}(s, u).$$

Therefore, after  $I = O(\alpha \cdot \log(n\Delta))$  steps, we have  $\sum_{u \in V} d_{G^{(I)}}(s, u) < 1/3$ , implying  $d_{G^{(I)}}(s, u) < 1/3$  since all distances are nonnegative. Then, by summing up (3.1) over  $i$ , we get  $d_G(s, u) = d_{G^{(I)}}(s, u) + \tilde{d}'(u)$ . This implies  $|d_G(s, u) - \tilde{d}'(u)| \leq d_{G^{(I)}}(s, u) < 1/3$ , hence rounding  $\tilde{d}'(v)$  to the nearest integer gives the correct (integer) distance  $d_G(u)$ .  $\square$

We remark that the oracle is always called on the same graph  $G$ , with the exception that the lengths of the edges are changed in-between calls. Notably, these changes cause the altered graph to generally become directed (i.e., has different lengths on anti-parallel edges), even if the original graph was undirected. Finally, we remark that in many parallel or distributed models (and many others), if the model supports an efficient implementation of  $\mathcal{O}$ , then Algorithm 1 can also be implemented efficiently. Therefore, exact and (smoothly) approximate distances on (integer weighted) directed graphs are essentially equivalent.

## 4 Different Notions of Approximate Distances

This section is devoted to the exposition of all kinds of different notions of approximate distances that are either standard or are considered in this paper. We also prove fundamental results for the smooth and tree-like property and their combination. All of the notions discussed in this section work for directed graphs. However, we note that the notions of tree-like, strong, and to some extent also smooth distances are interesting mostly for the undirected case, since in the directed graphs we can use the boosting technique from Section 3.1 to get even exact shortest paths.

The roadmap for this section is as follows. In Section 4.1 we first briefly review the standard notion of graph potentials. In Section 4.2 we discuss smoothly approximate distances and explain in which sense smooth distances strengthen both approximate distances and potentials. Next, in Section 4.3 we discuss the notion *tree-like* distances, i.e., those coming from an underlying approximate shortest path tree. Finally, in Section 4.4 we argue that having these two properties together leads to strong distances that have a simple alternative definition.

### 4.1 Graph Potentials

A notion closely related to that of approximate distances is that of graph potentials. These are distance estimates (i.e., functions  $V(G) \rightarrow \mathbb{R}_{\geq 0}$ ) that are dual to approximate distances as can be formalized by the following linear program (this program will not play a role in the subsequent discussion).

<b>Variables:</b>	<b>(Primal)</b> $f : E \rightarrow \mathbb{R}_{\geq 0}$	<b>(Dual)</b> $\phi : V \rightarrow \mathbb{R}$
<b>Optimize:</b>	$\max. \sum_{e \in E} \ell(e) f(e)$	$\min. \phi(t) - \phi(s)$
<b>Such that:</b>	$\sum_{v \in N^{\text{out}}(u)} f(u, v) - \sum_{v \in N^{\text{in}}(u)} f((v, u)) =$ $= \mathbb{1}[u = s] - \mathbb{1}[u = t] \quad \forall u \in V$	$\phi(v) - \phi(u) \leq \ell(u, v)$ $\forall (u, v) \in E$

Formally, we define  $\beta$ -approximate potential as follows.

**Definition 4.1** ( $\beta$ -approximate potential). A function  $\phi : V(G) \mapsto \mathbb{R}_{\geq 0}$  is a potential with respect to  $s$  if  $\phi(s) = 0$  and

$$\forall (u, v) \in E(G) : \phi(v) \leq \phi(u) + \ell_G(u, v). \quad (4.1)$$

A potential  $\phi$  is  $\beta$ -approximate (with respect to a source  $s$ ) for some  $\beta \geq 1$  if

$$\beta \cdot \sum_{u \in V(G)} \phi(u) \geq \sum_{u \in V(G)} d_G(s, u). \quad (4.2)$$

Again, note that for  $\beta = 1$  we recover exact shortest-path distances. In general, potentials always underapproximate the actual distances in the sense that  $\phi(v) - \phi(u) \leq d_G(u, v)$ . Hence, they are  $\beta$ -approximate iff they do not underapproximate by more than a  $\beta$  factor, on average.

We note that the boosting result Lemma 3.1 from Section 3.1 in fact does not require to start with smooth distance functions. It can boost any  $\beta$ -approximate potential to the exact shortest-path distance function in  $O(\beta \log n)$  iterations.

## 4.2 Smoothly Approximate Distances

The most important concept analysed in this paper is that of *smoothly*  $\alpha$ -approximate distances, as defined next. The discussion of their applications is deferred to Section A. For convenience we now repeat the definition of the  $\alpha$ -smooth condition that was given in Definition 1.5.

**Definition 4.2** (Smooth distances). Given a graph  $G = (V, E, \ell)$  and a source  $s \in V$ , we say a distance estimate  $\tilde{d} : V \rightarrow \mathbb{R}_{\geq 0}$  is  $\alpha$ -smooth if  $\tilde{d}(s) = 0$  and

$$\forall u, v \in V : \tilde{d}(v) - \tilde{d}(u) \leq \alpha \cdot d_G(u, v),$$

Moreover, we say that a distance estimate is smoothly  $\alpha$ -approximate if it is both  $\alpha$ -smooth and noncontractive, i.e., it moreover satisfies

$$\forall u \in V : d_G(s, u) \leq \tilde{d}(u). \quad (4.3)$$

Observe that smoothly  $\alpha$ -approximate distances strengthen  $\alpha$ -approximate distances. Indeed, if we set  $u = s$  in the above  $\alpha$ -smoothness definition, we exactly recover the bound  $\tilde{d}(v) \leq \tilde{d}(s) + \alpha \cdot d_G(s, v) = \alpha \cdot d_G(s, v)$  for  $\alpha$ -approximate distance estimate.

**Subtractive triangle inequality.** The  $\alpha$ -smoothness requirement  $\tilde{d}(v) \leq \tilde{d}(u) + \alpha \cdot d_G(u, v)$  can be seen as requiring an *approximate triangle inequality* to hold for distances to  $s$ . In particular, for exact distances (i.e.,  $\alpha = 1$ ) the triangle inequality stipulates correctly that  $d_G(s, v) \leq d_G(s, u) + d_G(u, v)$ . However, this inequality completely fails to hold — even approximately — if exact distances are replaced by approximate ones (see Figure 2). Our strengthening demands that this inequality remains to hold  $\alpha$ -approximately. This interpretation explains most easily why smoothly approximate distances can replace exact distances in some contexts: Any use of the above triangle inequality by a proof of correctness still holds approximately if smoothly approximate distances are used to replace exact distances (e.g., see Section A for an example).

**Smoothness vs. potentials.** Next, we discuss the relationship of smoothly approximate functions with potentials. Observe having the  $\alpha$ -smooth property is, up to rescaling, the same as being a potential.

**Fact 4.3.** A distance estimate  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  is  $\alpha$ -smooth if and only if  $\tilde{d}/\alpha$  is a potential.

*Proof.* Recall that potentials satisfy  $\phi(v) - \phi(u) \leq d_G(u, v)$  for all  $u, v \in V(G)$ . Then the result follows because  $\tilde{d}(v) - \tilde{d}(u) \leq \alpha \cdot d_G(u, v)$  is equivalent to  $\frac{\tilde{d}(v)}{\alpha} \leq \frac{\tilde{d}(u)}{\alpha} + d_G(u, v)$ .  $\square$

A helpful view of smoothly  $\alpha$ -approximate distances is that they get the “best of the both worlds”: they are exactly the distances that strengthen both  $\alpha$ -approximate distances and  $\alpha$ -approximate potentials.

**Lemma 4.4.**  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  is a smoothly  $\alpha$ -approximate distance estimate (w.r.t.  $s$ ) if and only if

1.  $\tilde{d}$  is an  $\alpha$ -approximate distance estimate (w.r.t.  $s$ ), and
2.  $\tilde{d}/\alpha$  is an  $\alpha$ -approximate potential function (w.r.t.  $s$ ).

*Proof.* The “only if” direction follows from our observations that smoothly  $\alpha$ -approximate distances are  $\alpha$ -approximate and Fact 4.3. To prove the “if” direction, we need to prove that  $\tilde{d}$  is noncontractive ((1.1)) and  $\alpha$ -smooth (Definition 4.2). The first property follows from the fact that  $\tilde{d}$  is  $\alpha$ -approximate. The second property follows from Fact 4.3.  $\square$

To summarize, the smoothly approximate distances generalize approximate distances since they additionally satisfy “subtractive triangle inequality”. But, rescaled by  $\alpha$ -factor, they also generalize approximate potentials, since instead of satisfying the aggregate condition  $\sum_{u \in V(G)} \tilde{d}(u) \geq \sum_{u \in V(G)} d_G(s, u)/\alpha$ , they in fact satisfy a stronger, individual condition: for each  $u$  they have  $\tilde{d}(u)/\alpha \geq d_G(s, u)/\alpha$ .

### 4.3 Tree-Like Distances

In this subsection we will discuss the  $\alpha$ -approximate shortest path trees — the object that many approximate shortest path algorithms recover together with  $\alpha$ -approximate distance functions. Formally,  $\alpha$ -approximate shortest path trees are defined as follows.

**Definition 4.5** ( $\alpha$ -approximate shortest path tree). A spanning tree  $T$  of a graph  $G$  (where  $T \subseteq G$ ) and source node  $s$  is an  $\alpha$ -approximate shortest path tree for some  $\alpha \geq 1$  if

$$\forall u : d_T(s, u) \leq \alpha \cdot d_G(s, u).$$

We remark that any subtree-induced distances can only overapproximate the true distances, hence  $T$  being an ( $\alpha$ -approximate) shortest path tree automatically implies that the function  $d_T(s, \cdot)$  is noncontractive.

We find it useful to define *tree-like* distance functions — roughly speaking these are the distance functions such that we can simply recover their shortest path tree. The usefulness of this concept stems from the fact that it is sometimes easier to work with a simple algebraic condition than to work with a special tree-structure on top of the distance function. Also, we will see in Section 4.4 that being tree-like strengthens being noncontractive from (1.1) in Definition 1.4 similarly to how being  $\alpha$ -smooth strengthens the property (1.2). For convenience, we now repeat the definition of the tree-like property from Definition 1.5.



**Definition 4.6** (Tree-Like Distances). Given a graph  $G = (V, E, \ell)$  and a source  $s \in V$ , we say a distance estimate  $\tilde{d} : V \rightarrow \mathbb{R}_{\geq 0}$  is: *tree-like* (w.r.t.  $s$ ) if  $\tilde{d}(s) = 0$  and

$$\forall v \neq s : \exists e = (u, v) \in E : \tilde{d}(u) \leq \tilde{d}(v) - \ell(e).$$

Observe that tree-like condition strengthens noncontractivity of (1.1) in Definition 1.4 since from any node  $v$  we can use the tree-like condition repeatedly to construct a sequence of nodes starting from  $v = v_1, v_2, \dots, v_k = s$  with  $\tilde{d}(v_{i+1}) \leq \tilde{d}(v_i) - \ell(v_i, v_{i+1})$ . Together with the requirement that  $\tilde{d}(s) = 0$ , this implies  $\tilde{d}(v) \geq d_G(s, v)$ .

Note that above argument requires that all edge lengths are strictly positive. However, one can often simply transform a graph with nonnegative edge-lengths so that its edges then have only positive lengths: we can add a very small length to every edge or, in case of undirected graphs, we first find connected components of 0-length edges and contract them to get exclusively positive lengths.

Let us now show in which sense tree-like distance functions correspond to approximate shortest path trees.

**Lemma 4.7.** *Let  $G$  be a graph with edges of positive weights.*

1. *Let  $T$  be an  $\alpha$ -approximate spanning tree of  $G$  rooted at  $s$ . Then  $d_T(s, \cdot)$  is a tree-like,  $\alpha$ -approximate distance estimate.*
2. *Let  $\tilde{d}$  be a tree-like distance estimate and define a tree  $T$  as follows. For each  $v \in V(G), v \neq s$ , consider an arbitrary  $u \in N_G^{\text{in}}(v)$  satisfying  $\tilde{d}(u) \leq \tilde{d}(v) - \ell_G(u, v)$ ; add  $(u, v)$  to  $E(T)$ .*

*We have  $d_T(s, v) \leq \tilde{d}(v)$  and, in particular, if  $\tilde{d}$  is  $\alpha$ -approximate then  $T$  is an  $\alpha$ -approximate shortest path tree.*

Note that we are again requiring positive and not nonnegative edge weights here.

*Proof.* We start with the first claim. By definition, for every  $v \neq s$  the predecessor  $u$  of  $v$  in  $T$  satisfies  $d_T(s, u) = d_T(s, v) - \ell_G(u, v)$  and hence  $d_T(s, \cdot)$  is tree-like.

We continue with the second claim. First, note that  $T$  is a well defined tree since we assume there are no 0-length edges. The property  $d_T(s, v) \leq \tilde{d}(v)$  then holds by induction on the unique path from  $s$  to  $v$  in  $T$ . Whence, whenever  $\tilde{d}(v) \leq \alpha \cdot d_G(s, v)$  for all  $v \in V(G)$ , we conclude that  $d_T(s, v) \leq \alpha d_G(s, v)$  and  $T$  is an  $\alpha$ -approximate shortest path tree.  $\square$

Finally, let us observe that the tree-like property does not perfectly match the intuition that the function corresponds to a shortest path tree. It may seem that a better definition of tree-likeness would strengthen  $\leq$  to  $=$  in Definition 4.6, i.e., perhaps we should require that

$$\forall v : \exists u \in N^{\text{in}}(v) : \tilde{d}(u) = \tilde{d}(v) - \ell_G(u, v). \quad (4.4)$$

Above property would allow us to write  $d_T(s, v) = \tilde{d}(v)$  in the second part of Lemma 4.7, i.e., such a function would be directly induced by  $T$ . However, this alternative definition does not seem to behave as nicely as our definition. For example, the property defined by (4.4) is not closed under element-wise min-operations, unlike tree-likeness (Lemma 4.10). Also, our main technical contribution, the smoothing algorithm from Section 5, preserves the tree-like property but does not preserve (4.4).

## 4.4 Strongly approximate distances

When a single distance estimate  $\tilde{d}$  has both the  $\alpha$ -smooth property and is tree-like, we give it a special label of being a *strongly*  $\alpha$ -approximate distance. This is to emphasize the synergy between the two properties, which we will discuss in this section.

**Definition 4.8.** We say that a distance estimate is strongly  $\alpha$ -approximate if it has both the  $\alpha$ -smooth property and is tree-like.

Strong distances are especially suited for applications — a case study is given in Section A.

We argue that strong distances are “as close as it gets” to exact distances. In fact, we can make this claim precise: a strong distance estimate in  $G$  is, in fact, the exact distance in  $G$  with edge lengths perturbed by at most  $\alpha$ .

**Lemma 4.9.** *Let  $G = (V, E, \ell)$  be a graph with positive lengths,  $s \in V(G)$ , and  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  be a distance estimate. The following two claims are equivalent.*

- *The distance estimate  $\tilde{d}$  is strongly  $\alpha$ -approximate.*
- *There exist edge multipliers  $1 \leq \lambda(e) \leq \alpha$  for every  $e \in E(G)$  such that the following holds. Let  $G' = (V, E, \ell')$  with  $\ell'(e) := \lambda(e)\ell(e)$  for every  $e \in E(G)$  be the graph  $G$  with edges multiplicatively perturbed by  $(\lambda(e))_{e \in E(G)}$ . Then,  $\tilde{d}$  is an exact distance function on  $G'$  in the sense that  $\tilde{d}(u) = d_{G'}(s, u)$  for all  $u \in V(G)$ .*

*Proof.* We prove the result for directed graphs, it is simple to adapt it to undirected graphs. We first prove that the first claim implies the second. We are going to define the multipliers in such a way that the tree given by the tree-like property of  $\tilde{d}$  becomes the exact shortest path tree in  $G'$ .

More concretely, we define the multipliers  $\lambda$  as follows. First, we use that  $\tilde{d}$  is tree-like and let every node  $v$  choose its neighbor  $u$  such that  $\tilde{d}(u) \leq \tilde{d}(v) - \ell_G(e)$  where  $e = (u, v)$ . The smoothness of  $\tilde{d}$  implies that  $\tilde{d}(u) \geq \tilde{d}(v) - \alpha\ell_G(e)$ . Putting the two properties together, we infer that

$$\ell_G(e) \leq \tilde{d}(v) - \tilde{d}(u) \leq \alpha\ell_G(e).$$

Thus we can choose  $\lambda(e)$  such that  $\ell_{G'}(e) = \lambda(e)\ell_G(e) = \tilde{d}(v) - \tilde{d}(u)$ .

For all other edges  $e = (u, v)$  we choose  $\lambda(e) = \alpha$ ; note that this implies

$$\tilde{d}(v) - \tilde{d}(u) \leq \alpha\ell_G(e) = \ell_{G'}(e).$$

We claim  $\tilde{d}$  are exact distances on  $G'$  which we show by proving that  $d_{G'}(s, v) \leq \tilde{d}(v) \leq d_{G'}(s, v)$ . To observe that  $\tilde{d}(v) \leq d_{G'}(s, v)$  for all  $v \in V(G)$ , note that by construction we have that  $\tilde{d}(v) - \tilde{d}(u) \leq \ell_{G'}(e)$ . That is,  $\tilde{d}$  is a potential function on  $G'$  and in particular  $\tilde{d}(v) \leq d_{G'}(s, v)$  for all  $v \in V(G)$ .

On the other hand, for every  $v$  there is  $u$  with  $\tilde{d}(u) < \tilde{d}(v)$  (as all edges have nonzero lengths) where by construction we have  $\tilde{d}(v) = \tilde{d}(u) + \ell_{G'}(u, v)$ . Hence, we can inductively find a path from  $s$  to  $v$  such that  $\tilde{d}(v) = d_{P \cap G'}(s, v)$ . Consequently, we infer that  $\tilde{d}(v) \geq d_{G'}(s, v)$  (i.e., is noncontractive on  $G'$ ) and this completes the proof of the first part of the statement.

It remains to prove that the second claim implies the first one in the lemma statement. Since  $\tilde{d}$  is exact on  $G'$ , for an edge  $(u, v) \in E(G)$  we have  $\tilde{d}(v) - \tilde{d}(u) \leq \ell'(u, v) \leq \alpha \cdot \ell(u, v)$ . Hence,  $\tilde{d}$  has the  $\alpha$ -smooth property. Furthermore, let  $T'$  be the (exact) shortest path tree from the source  $s$  in  $G'$ . For each node  $v$  let  $(u, v) \in E(G)$  be the parent edge of  $v$  in  $T'$ . By definition, we have

$\tilde{d}(u) = \tilde{d}(v) - \ell'(u, v) \geq \tilde{d}(v) - \ell(u, v)$  since  $\ell(u, v) \leq \ell'(u, v)$ , thus  $\tilde{d}$  is tree-like w.r.t.  $s$ . Therefore,  $\tilde{d}$  is strongly  $\alpha$ -approximate, as needed.  $\square$

**Closure properties.** We observe that the discussed distance notions have useful closure properties, which will be useful later.

**Lemma 4.10.** *Let  $\tilde{d}_1, \tilde{d}_2$  be two distance estimates. Let us define  $\tilde{d}_{\min} = \min(\tilde{d}_1, \tilde{d}_2)$ ,  $\tilde{d}_{\max} = \max(\tilde{d}_1, \tilde{d}_2)$ ,  $\tilde{d}_{\text{avg}} = \lambda \tilde{d}_1 + (1 - \lambda) \tilde{d}_2$  for some  $0 \leq \lambda \leq 1$  (element-wise). We have:*

1. *If both  $\tilde{d}_1, \tilde{d}_2$  are  $\alpha$ -approximate, so are  $\tilde{d}_{\min}, \tilde{d}_{\max}, \tilde{d}_{\text{avg}}$ .*
2. *If both  $\tilde{d}_1, \tilde{d}_2$  have the  $\alpha$ -smooth property so does  $\tilde{d}_{\min}, \tilde{d}_{\max}, \tilde{d}_{\text{avg}}$ .*
3. *If both  $\tilde{d}_1, \tilde{d}_2$  are tree-like, so is  $\tilde{d}_{\min}$ .*

*Proof.* The item (1) follows from the chain of inequalities  $d_G(s, v) \leq \tilde{d}_{\min}(v) \leq \tilde{d}_{\text{avg}}(v) \leq \tilde{d}_{\max}(v) \leq \alpha d_G(s, v)$ .

We continue with item (2). We start with  $\tilde{d}_{\min}$ , let  $u, v$  be any two nodes of  $G$  and without loss of generality assume that  $\tilde{d}_{\min}(u) = \tilde{d}_1(u)$ . We then have  $\tilde{d}_{\min}(v) - \tilde{d}_{\min}(u) \leq \tilde{d}_1(v) - \tilde{d}_1(u) \leq \alpha d_G(u, v)$  as needed. For  $\tilde{d}_{\max}$ , let us similarly without loss of generality assume  $\tilde{d}_{\max}(v) = \tilde{d}_1(v)$ . We then have  $\tilde{d}_{\max}(v) - \tilde{d}_{\max}(u) \leq \tilde{d}_1(v) - \tilde{d}_1(u) \leq \alpha d_G(u, v)$  as needed. Finally, for  $\tilde{d}_{\text{avg}}$  we can write

$$\tilde{d}_{\text{avg}}(v) - \tilde{d}_{\text{avg}}(u) = \lambda \left( \tilde{d}_1(v) - \tilde{d}_1(u) \right) + (1 - \lambda) \left( \tilde{d}_2(v) - \tilde{d}_2(u) \right) \leq (\lambda + (1 - \lambda)) \cdot \alpha d_G(u, v)$$

and we are done.

We finish with item (3). Fix a node  $v$  and without loss of generality assume  $\tilde{d}_{\min}(v) = \tilde{d}_1(v)$ . Consider the node  $u$  such that  $\tilde{d}_1(u) \leq \tilde{d}_1(v) - d_G(u, v)$ . We have  $\tilde{d}_{\min}(u) \leq \tilde{d}_1(u) \leq \tilde{d}_1(v) - d_G(u, v) = \tilde{d}_{\min}(v) - d_G(u, v)$  as needed.  $\square$

We note that  $\tilde{d}_{\text{avg}}, \tilde{d}_{\max}$  are not necessarily tree-like even if  $\tilde{d}_1, \tilde{d}_2$  are. This can happen e.g. whenever  $v$  has two predecessors  $u_1, u_2$  such that  $\tilde{d}_1(u_1) \leq \tilde{d}_1(v) - d_G(u, v)$  but  $\tilde{d}_1(u_1) > \tilde{d}_1(v) - d_G(u, v)$  and analogously for  $\tilde{d}_2$ .

**Local checkability.** What do  $\alpha$ -smoothness and tree-likeness have in common? We observe that  $\alpha$ -smoothness is a locally checkable proof that a distance function satisfies the property (1.2), whereas tree-likeness is a locally checkable proof that a distance function satisfies the noncontractiveness property (1.1). Therefore, strongly  $\alpha$ -approximate distances are  $\alpha$ -approximate distances where both required properties can be checked locally.

We briefly discuss the topic of local checkability: Given a distance estimate  $\tilde{d}$  that is presumably  $\alpha$ -approximate, it is often convenient to be able to prove its approximation guarantee without trusting the algorithm; or, e.g., see whether it still holds after a graph changes to avoid re-running the computation. One way to make this certification precise is to look for local algorithms that verify the property (see [Feu19] for an introduction to the topic of local certification). More precisely, each node is given the values in its (small-hop) neighborhood, and outputs YES/NO; the property holds if and only if all nodes output YES.

It is not possible to directly check whether  $\tilde{d}$  is  $\alpha$ -approximate without computing the exact distances first. On the other hand, one can easily check whether a distance estimate  $\tilde{d}$  is strongly  $\alpha$ -approximate via a local algorithm. Concretely, for the  $\alpha$ -smoothness property, each node  $u$

checks for all outgoing edges  $(u, v)$  whether  $\tilde{d}(v) \leq \tilde{d}(u) + \alpha d_G(u, v)$  (outputs NO if any fail). This works since requiring that the  $\alpha$ -smooth condition holds for all edges  $(u, v)$  already implies it holds for any pair  $u, v$ . Tree-likeness is similar; together, we conclude that one can locally verify whether  $\tilde{d}$  is strongly  $\alpha$ -approximate.

The fact that a property can be easily checked is often helpful. For example, it allows us to turn any randomized Monte Carlo algorithm for strongly approximate distance computation into a Las Vegas one.

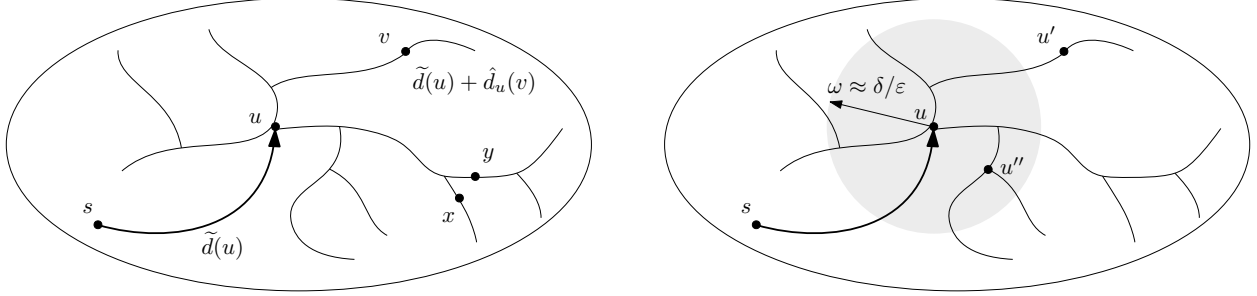
#### 4.5 Do Known Algorithms give Strong Distance Guarantees?

Many important algorithms use exact shortest path as an algorithmic building block (e.g., low-diameter decompositions,  $\ell_1$ -embeddings, hopsets, emulators, various versions of min-cost maximum flows, etc.). However, obtaining exact distances is hard in many settings (e.g., parallel or distributed), and replacing exact distances with an arbitrary  $(1 + \varepsilon)$ -approximation is insufficient. For this reason, there is a long list of papers which obtain approximate distances with additional guarantees. However, as far as the authors as aware, only a few approaches obtain strong distances, and they only do it in very specific settings that cannot be readily generalized. In this section, we review several algorithms from the parallel and distributed literature and clarify their properties.

**Hopsets.** A hopset with hopbound  $\beta \geq 1$  of a graph  $G = (V, E, \ell)$  is a (sparse) graph  $H = (V, E', \ell')$  such that  $\beta$ -hop paths in  $G \cup H$   $(1 + \varepsilon)$ -approximate shortest path of  $G$ . More precisely,  $d_G(u, v) \leq d_{G \cup H}^{(\beta)}(u, v) \leq (1 + \varepsilon)d_G(u, v)$  for all  $u, v \in V$ , where  $d^{(\beta)}(u, v)$  is the shortest path using paths of at most  $\beta$  hops [Coh00, EN19, HP19, CFR20a]. Simple  $O(\beta)$ -time algorithms like Bellman-Ford can compute  $\tilde{d}(v) := d_{G \cup H}^{(\beta)}(s, v)$ , which  $(1 + \varepsilon)$ -approximate SSSP distances from  $s$ . Moreover, every edge in  $H$  corresponds to a path in  $G$ , hence it is not hard to prove that such  $\tilde{d}$  is tree-like. On the other hand,  $\tilde{d}$  is generally not smooth. Hence, hopsets do not directly give strong distances.

**Algorithms computing approximate potentials.** Computing approximate potentials (i.e., smooth and approximate distance) can be used to boost the approximation guarantees for distances in both directed [KS97] and undirected settings [She17, Zuz21]. Therefore, many papers develop tools to compute approximate potentials. For example, the state-of-the-art  $(1 + \varepsilon)$ -approximate undirected shortest path papers in parallel and distributed setting [Li20, BFKL21, RGH<sup>+</sup>22, ZGY<sup>+</sup>22] all can compute a primal-dual pair  $(f, \phi)$  where  $f$  is a  $(1 + \varepsilon)$ -approximate shortest path tree, and  $\phi$  is a  $(1 + \varepsilon)$ -approximate potential. However, these approaches do not yield a single distance estimate  $\tilde{d}$  that is strong (i.e., has primal-dual guarantees)—the primal tree is not smooth, and the potentials are not tree-like.

**Emulator approach.** Several papers do obtain strong approximate distances using emulators, which are graph  $H$  such that  $d_H(u, v) \leq d_G(u, v) \leq \alpha \cdot d_H(u, v)$  for all  $u, v$  (i.e., an emulator is an  $\alpha$ -spanner without the subgraph requirement). These papers construct an emulator with additional structural properties in a way that it is possible to compute *exact* distances on them. These are strong distances per Lemma 4.9. For example, Andoni, Stein, and Zhong [ASZ20] compute a  $\text{poly}(\log n)$ -emulator where  $O(\log \log n)$ -hop paths exactly match the shortest paths, hence can be computed in their parallel setting. In another approach, Forster and Nanongkai [FN18] (and several other papers that follow this line of work) randomly sample about  $\sqrt{n}$  nodes, compute a 2-emulator on it, and then compute exact distances on this emulator (which can be extended to all nodes using small number of hops). Unfortunately, it is unclear how to make these approaches work outside of their respective settings.



**Figure 4:** Left: In the first try, we simply try to “fix” a node  $u$  (by that we mean making all pairs  $u, v$  satisfy the smoothness condition) by computing approximate distances  $\hat{d}_u$  from  $u$ . Adding  $\tilde{d}(u)$  to all computed values then creates a distance estimate  $\tilde{d}_u$  measuring distances from  $s$ . This distance estimate fixes  $u$ ; even if we set  $\tilde{d}' = \min(\tilde{d}, \tilde{d}_u)$ ,  $u$  remains fixed in  $\tilde{d}'$ . The problem is that even though a pair  $x, y$  may have satisfied the smoothness condition in  $\tilde{d}$ , we may unfix it in  $\tilde{d}'$ .

Right: To deal with the problem, we use the same construction but slow down the edges of  $G$  by a factor of  $(1 + \varepsilon)\alpha$ . We also do not aim to make  $u$  satisfy the  $\alpha$ -smoothness condition right away but instead only go from  $(\alpha, \delta)$ -smoothness to  $((1 + \varepsilon)\alpha, \delta/2)$ -smoothness guarantee. Because of the slowdown and assumption that  $\tilde{d}$  is  $(\alpha, \delta)$ -smooth, any node  $u'$  that is further than  $\omega = 2\delta/(\alpha\varepsilon) = \Theta(\delta/\varepsilon)$  from  $u$  has  $\tilde{d}(u') \geq \tilde{d}_u(u')$  which means that  $\tilde{d}'(u) = \tilde{d}(u)$  and we do not unfix it. On the other hand, whenever a node  $u''$  is closer than  $\omega = \Theta(\delta/\varepsilon)$  to  $u$ , we use the fact that the approximation oracle can make the value  $\hat{d}(u'')$  distorted from the truth only additively by  $\frac{\varepsilon}{100} \cdot \omega \ll \delta/2$ . Hence, up to a small error dominated by  $\delta/2$  we can use the same argument that shows that  $u$  gets fixed for  $u''$ .

## 5 The Smoothing Algorithm

This section is devoted to the proof of our main technical result, Theorem 5.1 that we restate here for convenience.

**Theorem 5.1.** *Obtaining smoothness* Given a graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , accuracy  $\varepsilon \in (0, 1]$ , and an approximate oracle  $\mathcal{O}$ , Algorithm 3 computes  $(1 + \varepsilon)$ -approximate distances that are  $(1 + \varepsilon)$ -smooth  $\tilde{d}$  in a directed graph  $G$  using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate distance oracle  $\mathcal{O}$  on directed graphs. Moreover, if  $\mathcal{O}$  returns tree-like distances, the algorithm returns strongly  $(1 + \varepsilon)$ -approximate distances. Moreover, if  $G$  is undirected, it suffices if  $\mathcal{O}$  works in undirected graphs.

First, in Section 5.1 we build the intuition behind the algorithm. Next, in Section 5.2 we prove Theorem 5.1 formally.

### 5.1 Intuition Behind the Algorithm

Let us now explain the intuition behind our algorithm that uses calls to  $(1 + \varepsilon)$ -approximate distance oracle to compute smooth approximate distances in directed graphs. We will build the algorithm in three steps.

**Step 1: The first try.** Recall that we wish to compute the  $(1 + \varepsilon)$ -approximate smooth distances from some source node  $s$ . We can start by running the approximate distance oracle from  $s$  to get some approximate distance estimate  $\tilde{d}$ . The problem of course is that there can be two nodes  $u, v$  ( $v$  further from  $s$  than  $u$ ) such that  $\tilde{d}(v) - \tilde{d}(u)$  is much larger than  $(1 + \varepsilon)d_G(u, v)$ . This is because

our oracle is allowed to make an additive error of  $\varepsilon \cdot d_G(s, u)$  for the distance labels  $\tilde{d}(u), \tilde{d}(v)$ ; this can be much larger than the distance  $d_G(u, v)$ .

Let us first discuss how we can use one more call to the approximate distance oracle to satisfy the smoothness condition for some fixed pair of nodes  $u, v$ : We run the approximate distance oracle for the second time, but we start at  $u$ . We get an approximate distance estimate  $\hat{d}_u$  from  $u$  that defines a new noncontractive distance estimate  $\tilde{d}_u$  from  $s$  by defining  $\tilde{d}_u(v) = \tilde{d}(u) + \hat{d}_u(v)$  (see Figure 4).

This new distance estimate  $\tilde{d}_u$  is not necessarily  $(1 + \varepsilon)$ -approximate, as it considers only routes through  $u$ . However, the pair of nodes  $u$  and  $v$  (for any  $v$ ) satisfies the smoothness condition: we have  $\tilde{d}_u(v) - \tilde{d}_u(u) = (\tilde{d}(u) + \hat{d}_u(v)) - \tilde{d}(u) = \hat{d}_u(v) \leq (1 + \varepsilon)d_G(u, v)$ .

Finally, if we define  $\tilde{d}'(\cdot) = \min(\tilde{d}(\cdot), \tilde{d}_u(\cdot))$ , the new estimate takes the best of the both worlds:  $\tilde{d}'$  remains a noncontractive  $(1 + \varepsilon)$ -approximate distance estimate that also additionally satisfies the smoothness condition for  $u$  and any other node  $v$  (to verify this, compare with the proof of Lemma 4.10).

So what is the bad news? It may seem that we can already construct a simple, albeit too slow, smoothing algorithm that simply repeats above construction for all nodes  $u \in V(G)$ , until all pairs of nodes satisfy the smoothness condition. However, such an algorithm does not work! While the above procedure for some  $u$  “fixes” all the pairs  $(u, v)$ , it may also “break” some other pair  $(x, y)$  elsewhere in the graph that satisfied the smooth condition in  $\tilde{d}$ .

**Step 2: A (formally) correct but slow algorithm.** We will now show how we can fight the problem that fixing one pair  $u, v$  by computing distances from  $u$  can break some other pair of nodes in the graph. Our algorithm will still require  $O(n \log n)$  calls to the distance oracle which we improve upon only in the final step.

Our main idea is, roughly, to fix the pairs  $u, v$  in  $O(\log n)$  phases, starting from the pairs with the largest  $d_G(u, v)$ . To be more concrete, let us generalize the smoothness condition as follows:

**Definition 5.2** ( $(\alpha, \delta)$ -smoothness). We say that a distance estimate  $\tilde{d}$  is  $(\alpha, \delta)$ -smooth if for every  $u, v \in V(G)$  we have

$$\tilde{d}(v) \leq \tilde{d}(u) + \alpha \cdot d_G(u, v) + \delta$$

We will show how to refine an  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$  to a  $((1 + \varepsilon)\alpha, \delta/2)$ -smooth estimate  $\tilde{d}_*$ . Then, to get  $(1 + \varepsilon_0)$ -smoothness, we simply choose  $\varepsilon = \Theta(\varepsilon_0 / \log n)$ , start with the trivial  $(1, \text{poly}(n))$ -smooth estimate (the weights are polynomially-bounded integers) and use the above reduction  $O(\log n)$  times until we end up with  $(1 + \frac{\varepsilon_0}{2}, \frac{1}{\text{poly}(n)})$ -smooth distances that are also  $(1 + \varepsilon_0)$ -smooth. On a more intuitive level, we first fix pairs at larger distance scales before moving on to smaller distance scales. Indeed,  $(\alpha, \delta)$ -smooth distances  $\tilde{d}$  are  $(\alpha + \varepsilon)$ -smooth for pairs  $(u, v)$  with  $d_G \geq \delta/\varepsilon$ :

$$\tilde{d}(v) \leq \tilde{d}(u) + \alpha \cdot d_G(u, v) + \delta \leq \tilde{d}(u) + (\alpha + \varepsilon)d_G(u, v).$$

The rest of this explanation focuses on a single refinement step. Our algorithm is very similar to the one discussed in the previous subsection: We simply iterate over each node  $u$  of  $G$  and run our  $(1 + \varepsilon/10)$ -approximate oracle from  $u$ , compute the distance estimates  $\tilde{d}_u(\cdot)$  as before, and output the element-wise minimum. The important difference is that we run our oracle on the “slowed down” graph  $G'$  where each edge length is multiplied by  $(1 + \varepsilon/2)\alpha$ . The method is formally written in Algorithm 2.

Intuitively, the benefit of the slowdown is that node  $u$  will not affect nodes far away from it, thereby not breaking them.

**Lemma 5.3.** *If  $d_G(u, u') > 2\delta/(\alpha\varepsilon)$ , then  $\tilde{d}_u(u') > \tilde{d}(u')$ .*

*Proof.* By assumption of  $(\alpha, \delta)$ -smoothness we have  $\tilde{d}(u') \leq \tilde{d}(u) + \alpha d_G(u, u') + \delta$ . Noncontractivity of distance estimates on the slowed-down graph implies  $\tilde{d}_u(u') \geq \tilde{d}(u) + (1 + \varepsilon/2)\alpha d_G(u, u')$ . Comparing the two expressions and rewriting our assumption as  $(\varepsilon/2) \cdot \alpha d_G(u, u') > \delta$  gives  $\tilde{d}_u(u') > \tilde{d}(u')$ , as required.  $\square$

Furthermore,  $\tilde{d}_u(\cdot)$  “fixes” the smoothness for all nodes close to  $u$  (a feature of our relaxed  $(\alpha, \delta)$  definition of smoothness).

**Lemma 5.4.** *If  $d_G(u, u') \leq 2\delta/(\alpha\varepsilon)$ , then all pairs  $(u', v)$  are  $(\alpha(1 + \varepsilon), \delta/2)$ -smooth in  $\tilde{d}_u(\cdot)$ .*

*Proof.* Indeed, for any node  $v$  we can write

$$\tilde{d}_u(v) - \tilde{d}_u(u') \leq \alpha(1 + \varepsilon/2)(1 + \varepsilon/10)d_G(u, v) - \alpha(1 + \varepsilon/2)d_G(u, u')$$

which, using triangle inequality  $d_G(u, v) - d_G(u, u') \leq d_G(u', v)$ , simplifies to

$$\tilde{d}_u(v) - \tilde{d}_u(u') \leq \alpha(1 + \varepsilon/2)(1 + \varepsilon/10)d_G(u', v) + \alpha(1 + \varepsilon/2)\frac{\varepsilon}{10} \cdot d_G(u, u')$$

The last error term can now be upper bounded by  $\alpha(1 + \varepsilon/2)\frac{\varepsilon}{10} \cdot 2\delta/(\alpha\varepsilon) \leq \delta/2$ , so we get

$$\tilde{d}_u(v) - \tilde{d}_u(u') \leq \alpha(1 + \varepsilon)d_G(u', v) + \delta/2.$$

In other words, the node  $u'$  now satisfies the  $((1 + \varepsilon)\alpha, \delta)$ -smooth condition!  $\square$

We now formally prove the correctness of Algorithm 2.

**Lemma 5.5.** *Every pair  $(u', v)$  is  $(\alpha(1 + \varepsilon), \delta/2)$ -smooth in  $\tilde{d}_*$ .*

*Proof.* Let  $u := \arg \min_u \tilde{d}_u(u')$  be the “parent” of  $u'$ . Hence,  $\tilde{d}_u(u') = \tilde{d}_*(u') \leq \tilde{d}_{u'}(u') = \tilde{d}(u')$ . By Lemma 5.3, we have  $d_G(u, u') \leq 2\delta/(\alpha\varepsilon)$ . Hence, by Lemma 5.4,  $(u', v)$  is  $((1 + \varepsilon)\alpha, \delta)$ -smooth in  $\tilde{d}_u$ . But,  $\tilde{d}_*(v) - \tilde{d}_*(u') = \tilde{d}_*(v) - \tilde{d}_u(u') \leq \tilde{d}_u(v) - \tilde{d}_u(u')$ , hence  $(u', v)$  is also smooth in  $\tilde{d}_*$ . This completes the proof.  $\square$

**Step 3: Speeding up the algorithm.** It remains to show how we can speed up Algorithm 2 so that it uses only 2 queries to the distance oracle per refinement step, instead of  $n$ . To this end, we will use two observations about how we can “bundle” the oracle calls for two different vertices into just one call.

First, consider some two nodes  $u_1, u_2$  such that  $|\tilde{d}(u_1) - \tilde{d}(u_2)| \leq \delta/\varepsilon$ . Instead of running the distance oracle once from  $u_1$  and once from  $u_2$ , we can use the following trick. Assuming  $\tilde{d}(u_1) < \tilde{d}(u_2)$ , construct a new special node  $s$ , connect it to  $u_1$  with an edge of length 0 and with  $u_2$  with an edge of length  $\tilde{d}(u_2) - \tilde{d}(u_1)$ . Run just one oracle call from  $s$  to get a distance estimate  $\tilde{d}_s(\cdot)$ ; define  $\tilde{d}_{u_1, u_2}(v) = \tilde{d}(u_1) + \tilde{d}_s(v)$ . The distance function  $\tilde{d}_{u_1, u_2}(\cdot)$  is basically the same as the function  $\min(\tilde{d}_{u_1}(\cdot), \tilde{d}_{u_2}(\cdot))$  that we want to compute. The difference is that the edge from  $s$  to  $u_2$  has length

---

**Algorithm 2** Slow Partial Smoothing algorithm

---

**Input:** A directed graph  $G$ , source  $s$ , parameter  $\varepsilon$ , approximate oracle  $\mathcal{O}$ ,  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$

**Oracle:**  $\mathcal{O}(G, s, \varepsilon)$  returns  $(1 + \varepsilon)$ -approximate distances  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  from  $s$  in  $G$ .

**Output:**  $((1 + \varepsilon)\alpha, \delta/2)$ -smooth distance estimate  $\tilde{d}'$

- 1: Let  $G'$  be the graph  $G$  with distances multiplied by  $(1 + \varepsilon/2)\alpha$ .
  - 2:  $\tilde{d} \leftarrow \mathcal{O}(G, s, \varepsilon)$
  - 3: **for** each  $u \in V(G)$  **do**
  - 4:      $\hat{d}_u \leftarrow \mathcal{O}(G', u, \varepsilon/10)$
  - 5:      $\tilde{d}_u(\cdot) \leftarrow \tilde{d}(u) + \hat{d}_u(\cdot)$
  - 6: **return**  $\tilde{d}_*(\cdot) = \min_{u \in V(G)} \tilde{d}_u(\cdot)$
- 

$\tilde{d}(u_2) - \tilde{d}(u_1) < \delta/\varepsilon$  and this worsens the argument from previous analysis of the slow smoothing algorithm by an additive  $\frac{\varepsilon}{10} \cdot \frac{\delta}{\varepsilon} = \frac{\delta}{10} \ll \frac{\delta}{2}$ . That is, this additional loss is insignificant.

Second, consider two nodes  $u_1, u_2$  with  $|\tilde{d}(u_1) - \tilde{d}(u_2)| \gg \delta/\varepsilon$ . Again, add a special node  $s$ , connect to  $u_1$  and  $u_2$  with an edge of length 0 and run our distance oracle from  $s$ . We already know from Lemma 5.3 that the distance information we get by calling our distance oracle from  $u_1$  is valuable only for nodes that are at distance at most  $O(\delta/\varepsilon)$  from  $u_1$ , and analogously for  $u_2$ . This again suggests that the information we gain by calling our oracle from  $s$  is sufficient to recover the value of  $\min(\tilde{d}_{u_1}(\cdot), \tilde{d}_{u_2}(\cdot))$ .

We put these two observations together as follows as visualized in Figure 5. We define two *level graphs*  $H_1, H_2$ ; to get the level graph  $H_1$ , we start with  $G$  and slow down all edges by a factor of  $(1 + \varepsilon)\alpha$ . Next, we partition the nodes of  $G$  into “levels” of length  $\omega = \Theta(\delta/\varepsilon)$  based on their value of  $\tilde{d}$ ; we delete edges going between different levels. Finally, we add a new node  $\sigma$  and connect it to every node  $u$  with an edge of length  $\{\tilde{d}(u)\}_1$  which is defined as the unique number between 0 and  $\omega$  such that  $\tilde{d}(u) - \{\tilde{d}(u)\}_1$  is divisible by  $\omega$ .

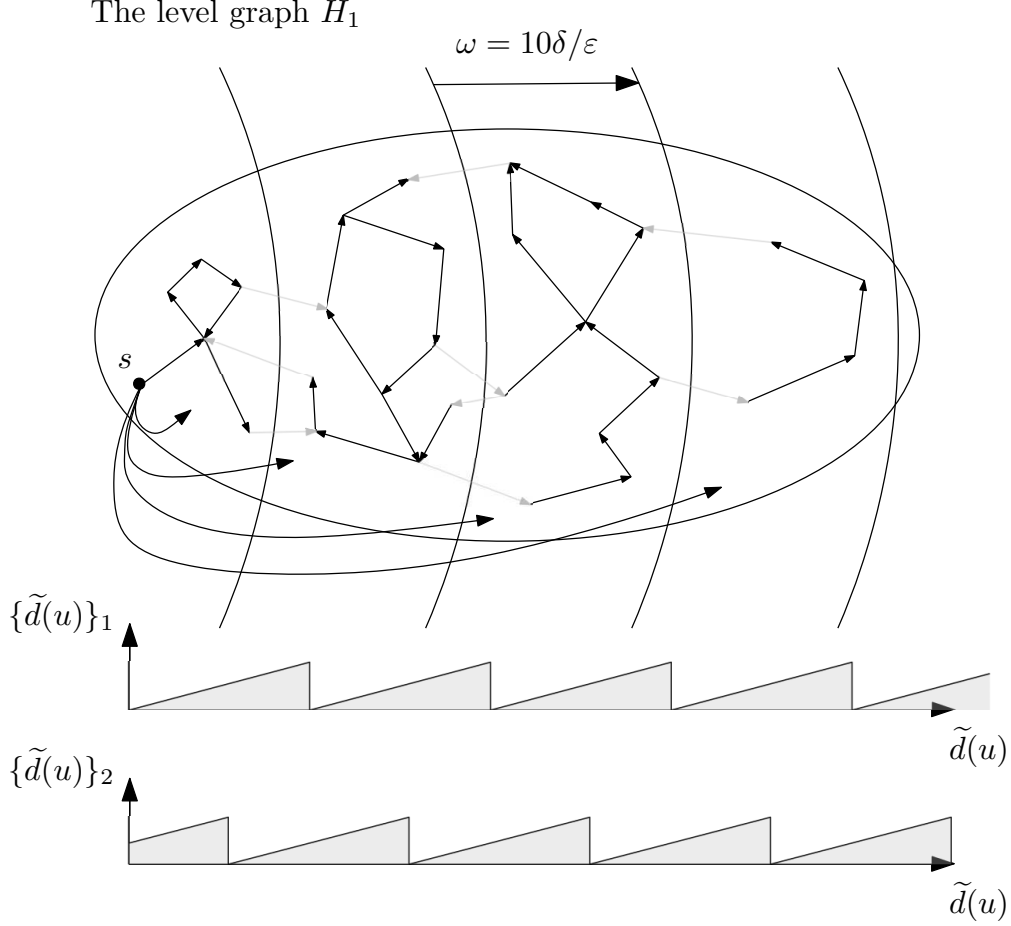
Running the distance oracle on  $H_1$  enables us to (approximately) compute  $\min(d_{u_1}(\cdot), d_{u_2}(\cdot), \dots)$  where the nodes  $u_1, u_2, \dots$  are all the nodes of  $G$  except those that are too close to the boundary of two different levels. We next define a graph  $H_2$  analogously to  $H_1$  but with its levels shifted by  $\omega/2$ . Computing the distances in  $H_1$  and  $H_2$  and taking the minimum of appropriate quantities then allows us to compute the value of  $\min_{u \in V(G)} d_u(\cdot)$ , thereby solving the problem we need to solve in Algorithm 2. However, we need just two calls to the distance oracle instead of  $n$  calls. Hence, the final smoothing algorithm only needs  $O(\log n)$  oracle calls.

## 5.2 The Formal Analysis

In this section we formally describe and analyze the smoothing algorithm described in Algorithm 3. Specifically, we prove its properties, restated below for convenience.

**Theorem 5.1.** *Obtaining smoothness* Given a graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , accuracy  $\varepsilon \in (0, 1]$ , and an approximate oracle  $\mathcal{O}$ , Algorithm 3 computes  $(1 + \varepsilon)$ -approximate distances that are  $(1 + \varepsilon)$ -smooth  $\tilde{d}$  in a directed graph  $G$  using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate distance oracle  $\mathcal{O}$  on directed graphs. Moreover, if  $\mathcal{O}$  returns tree-like distances, the algorithm returns strongly  $(1 + \varepsilon)$ -approximate distances. Moreover, if  $G$  is undirected, it suffices if  $\mathcal{O}$  works in undirected graphs.





**Figure 5:** The figure shows one level graph  $H_1$ . The level graph is created from  $G$  by cutting it based on  $\tilde{d}$  into levels of length  $\omega = 10\delta/\varepsilon$ . We remove edges that go across boundaries (grey edges). We also add edges from  $s$  to all other nodes  $v$  of length  $\{\tilde{d}(v)\}_1$ . This is the “fractional part” of  $\tilde{d}(v)$  as shown below the graph. The graph  $H_2$  is defined analogously but the fractional parts are shifted by  $\omega/2$ .

We start by describing the level graphs from Figure 5.

**Definition 5.6** (Rounding distances). Let  $G$  be a graph and  $\tilde{d}$  a distance estimate on it. Let  $i \in \{1, 2\}$  and  $\omega$  a parameter. We define  $\{\tilde{d}(u)\}_i$  to be the smallest value such that  $\tilde{d}(u) - \{\tilde{d}(u)\}_i - (i-1) \cdot \frac{\omega}{2}$  is divisible by  $\omega$ . We also define  $\tilde{d}(u) = \lfloor \tilde{d}(u) \rfloor_i + \{\tilde{d}(u)\}_i$ .

**Definition 5.7** (Level graph  $H_i$ ). Given an input directed graph  $G$ , a width parameter  $\omega$  and  $i \in \{1, 2\}$ , we define the *level graph*  $H_i$  as follows.

We define  $V(H_i) = V(G) \cup \{\sigma\}$ . The set of edges  $E(H_i)$  contains an edge  $(u, v) \in E(G)$  if and only if  $\lfloor \tilde{d}(u) \rfloor_i = \lfloor \tilde{d}(v) \rfloor_i$ . The weight of the edge  $(u, v)$  in  $H_i$  is defined as

$$\ell_{H_i}(u, v) = (1 + \varepsilon)\alpha \ell_G(u, v).$$

Moreover, for every  $u \in V(G)$  we add an edge  $(\sigma, u)$  to  $E(H_i)$  of weight  $\{\tilde{d}(u)\}_i$ .

To prove Theorem 5.1, we need mainly to analyse the smoothing subroutine of Algorithm 4. To analyse it, we first define the concept of a smoothing path.

---

**Algorithm 3** Smoothing algorithm  $\text{SMOOTHING}(G, s, \varepsilon)$ 

---

**Input:** A directed graph  $G$ , source  $s$ , parameter  $\varepsilon$ , approximate oracle  $\mathcal{O}(G, s, \varepsilon')$  for  $\varepsilon' = \mathcal{O}(\varepsilon/\log n)$

**Oracle:**  $\mathcal{O}(G, s, \varepsilon)$  returns  $(1 + \varepsilon)$ -approximate distances  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  from  $s$  in  $G$ .

**Output:** Smoothly  $(1 + \varepsilon_0)$ -approximate distance estimate  $\tilde{d}$ .

- 1: Let  $\Delta = n^{\mathcal{O}(1)}$  be the maximum distance in  $G$ .
  - 2:  $t \leftarrow 1 + \log_2(n\Delta/\varepsilon_0)$
  - 3:  $\varepsilon \leftarrow \varepsilon_0/(10t)$
  - 4: Define distance  $\tilde{d}_0(u) \leftarrow \mathcal{O}(G, s, 1)$  // this estimate is  $(1, n\Delta)$ -smooth.
  - 5: **for**  $i \leftarrow 1, \dots, t$  **do**
  - 6:      $\tilde{d}_i(u) \leftarrow \text{PARTIALSMOOTHING}(\tilde{d}_{i-1}(u), \varepsilon)$  // this estimate is  $((1 + \varepsilon)^i, n\Delta/2^i)$ -smooth
  - 7: **return**  $\tilde{d}' \leftarrow \tilde{d}_t$
- 

---

**Algorithm 4** Partial Smoothing algorithm  $\text{PARTIALSMOOTHING}$ 

---

**Input:** A directed graph  $G$ , source  $s$ , parameter  $\varepsilon$ , approximate oracle  $\mathcal{O}$ ,  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$ .

**Oracle:**  $\mathcal{O}(G, s, \varepsilon)$  returns  $(1 + \varepsilon)$ -approximate distances  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  from  $s$  in  $G$ .

**Output:**  $((1 + \varepsilon)\alpha, \delta/2)$ -smooth distance estimate  $\tilde{d}'$

- 1: Define  $H_1, H_2$  as in Definition 5.7 for  $\omega = 10\delta/\varepsilon$
  - 2: **for**  $i \leftarrow 1, 2$  **do**
  - 3:      $\hat{d}_i \leftarrow \mathcal{O}(H_i, s, \varepsilon/100)$
  - 4:      $\tilde{d}_i \leftarrow \lfloor \tilde{d}(u) \rfloor_i + \hat{d}_i(u)$
  - 5: **return**  $\tilde{d}'$  defined by  $\tilde{d}'(\cdot) = \min(\tilde{d}(\cdot), \tilde{d}_1(\cdot), \tilde{d}_2(\cdot))$
- 

**Definition 5.8** (Smoothing path). Given a parameter  $\varepsilon$  and an  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$  as in Algorithm 4, we will call a path  $P$  from  $u$  to  $v$  in  $G$  smoothing if

$$\tilde{d}(v) \geq \tilde{d}(u) + \alpha(1 + \varepsilon)\ell_G(P).$$

We next prove that any smoothing path cannot be very long.

**Claim 5.9.** *Any smoothing path  $P$  satisfies  $\ell_G(P) \leq \delta/(\alpha\varepsilon)$ .*

*Proof.* We have by definition that a smoothing path  $P$  from  $u$  to  $v$  satisfies

$$\alpha(1 + \varepsilon)\ell_G(P) \leq \tilde{d}(v) - \tilde{d}(u),$$

where by assumption on  $\tilde{d}(\cdot)$  we have

$$\tilde{d}(v) - \tilde{d}(u) \leq \alpha \cdot d_G(u, v) + \delta \leq \alpha \cdot \ell_G(P) + \delta.$$

Combining these gives the desired inequality. □

The fact that smoothing paths cannot be long then implies that any smoothing path is contained in either  $H_1$  or  $H_2$ .

**Claim 5.10.** *Any smoothing path  $P \subseteq G$  from  $u$  to  $v$  satisfies either  $P \subseteq H_1$  or  $P \subseteq H_2$ .*

*Proof.* Let  $w_{\max} = \operatorname{argmax}_{w \in P} \tilde{d}(w)$  and  $w_{\min} = \operatorname{argmin}_{w \in P} \tilde{d}(w)$ . We upper bound the difference  $\tilde{d}(w_{\max}) - \tilde{d}(w_{\min})$ . We have

$$\begin{aligned} \tilde{d}(w_{\max}) - \tilde{d}(w_{\min}) &\leq \tilde{d}(w_{\max}) - \tilde{d}(w_{\min}) + \tilde{d}(v) - \tilde{d}(u) \\ &= \left( \tilde{d}(w_{\max}) - \tilde{d}(u) \right) + \left( \tilde{d}(v) - \tilde{d}(w_{\min}) \right) \\ &\leq \alpha \ell_G(P) + \delta + \alpha \ell_G(P) + \delta \\ &\leq 2(\varepsilon^{-1} + 1)\delta \leq 3\delta/\varepsilon \end{aligned}$$

where we first used the  $(\alpha, \delta)$ -smoothness assumption on  $(u, w_{\max})$  and  $(w_{\min}, v)$  and then we applied Claim 5.9. Note that whenever  $P \not\subseteq H_1$ , we have necessarily  $\{\tilde{d}(w_{\min})\}_1 \leq \omega$  while  $\{\tilde{d}(w_{\max})\}_1 \geq \omega$ . However, our bound then implies that  $\frac{\omega}{2} \leq \omega - 3\delta/\varepsilon \leq \{\tilde{d}(w_{\min})\}_1 \leq \{\tilde{d}(w_{\max})\}_1 \leq \omega + 3\delta/\varepsilon \leq \frac{3\omega}{2}$  which implies that  $P \subseteq H_2$ .  $\square$

**Proposition 5.11.** *The output distance estimate  $\tilde{d}'$  in Algorithm 4 satisfies*

$$\min_{u \in V(G)} \left( \tilde{d}(u) + \alpha(1 + \varepsilon)d_G(u, v) \right) \leq \tilde{d}'(v) \leq \min_{u \in V(G)} \left( \tilde{d}(u) + \alpha(1 + \varepsilon)d_G(u, v) \right) + \delta/2, \quad (5.1)$$

*Proof.* We first observe that the first inequality holds. If  $\tilde{d}'(v) = \tilde{d}(v)$ , simply choose  $u = v$  in the left hand side minimization. Else, we have  $\tilde{d}'(v) = \tilde{d}_i(v)$  for some  $i \in \{1, 2\}$ . Note that the definition of  $H_i$  for any  $i \in \{1, 2\}$  implies that for any  $v \in V(G)$  we in fact have

$$d_{H_i}(\sigma, v) = \min_{u \in V(G)} \left( \{\tilde{d}(u)\}_i + d_{H_i[V(G)]}(u, v) \right) \quad (5.2)$$

As the computed distance  $\hat{d}_i$  only overestimates  $d_{H_i}(\sigma, v)$ , we thus have

$$\tilde{d}_i(v) = \lfloor \tilde{d}(v) \rfloor_i + \hat{d}_i(v) \geq \lfloor \tilde{d}(v) \rfloor_i + \min_{u \in V(G)} \left( \{\tilde{d}(u)\}_i + d_{H_i[V(G)]}(u, v) \right) \quad (5.3)$$

$$= \min_{u \in V(G)} \left( \tilde{d}(u) + d_{H_i[V(G)]}(u, v) \right) \quad (5.4)$$

where the last equality follows from the fact that there is a path from  $u$  to  $v$  in  $H_i$  only if  $\lfloor \tilde{d}(u) \rfloor_i = \lfloor \tilde{d}(v) \rfloor_i$ . The first inequality of (5.1) then follows from the fact that the distance in  $H_i$  are scaled by  $\alpha(1 + \varepsilon)$ .

So it remains to show the second inequality. Let  $u$  be the minimizer of the right-hand side of (5.1), and let  $P$  be any shortest path from  $u$  to  $v$ . Since we minimize over all nodes of  $V(G)$  including  $v$ , we conclude that

$$\tilde{d}(u) + \alpha(1 + \varepsilon)d_G(u, v) \leq \tilde{d}(v).$$

This means that  $P$  is a smoothing path.

By Claim 5.10 this path is completely contained in  $H_i$  for some  $i \in \{1, 2\}$ . It follows that for that  $i$  we have

$$\begin{aligned} \tilde{d}'(v) &\leq \tilde{d}'(v)_i \leq \lfloor \tilde{d}(u) \rfloor_i + (1 + \varepsilon/100) \left( \{\tilde{d}(u)\}_i + \alpha(1 + \varepsilon)d_G(u, v) \right) \\ &\leq \tilde{d}(u) + \varepsilon \{\tilde{d}(u)\}_i / 100 + \alpha(1 + \varepsilon)(1 + \varepsilon/100)d_G(u, v) \end{aligned}$$

where we used  $\lfloor \tilde{d}(u) \rfloor_i = \lfloor \tilde{d}(v) \rfloor_i$ . Notice that we can bound  $\varepsilon \{\tilde{d}(u)\}_i / 100 \leq \varepsilon \omega / 100 \leq \delta / 10$  and since  $P$  is smoothing, Claim 5.9 yields  $\alpha(1 + \varepsilon) \frac{\varepsilon}{100} d_G(u, v) \leq \delta / 50$ . Thus, above inequality simplifies to

$$\tilde{d}'(v) \leq \tilde{d}(u) + \alpha(1 + \varepsilon) d_G(u, v) + \delta / 2$$

and we are done.  $\square$

**Proposition 5.12.** *If Algorithm 4 is called with an  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$ , then the output  $\tilde{d}'$  is  $(\alpha(1 + \varepsilon), \delta/2)$ -smooth.*

*Proof.* Let  $v, w$  be any vertices in  $G$ . Letting  $u := \operatorname{argmin}_{u' \in V(G)} \tilde{d}(u') + \alpha(1 + \varepsilon) d_G(u', v)$  it follows by Proposition 5.11 that

$$\tilde{d}(u) + \alpha(1 + \varepsilon) d_G(u, v) \leq \tilde{d}'(v),$$

and

$$\tilde{d}'(w) \leq \tilde{d}(u) + \alpha(1 + \varepsilon) d_G(u, w) + \delta / 2.$$

Combining these, we get that

$$\begin{aligned} \tilde{d}'(w) - \tilde{d}'(v) &\leq \alpha(1 + \varepsilon) d_G(u, w) + \delta / 2 - \alpha(1 + \varepsilon) d_G(u, v) \\ &\leq \alpha(1 + \varepsilon) (d_G(u, v) + d_G(v, w)) + \delta / 2 - \alpha(1 + \varepsilon) d_G(u, v) \\ &= \alpha(1 + \varepsilon) d_G(v, w) + \delta / 2. \end{aligned}$$

as desired.  $\square$

After Algorithm 4 is analysed, it is simple to finish the proof of the main theorem Theorem 5.1.

*Proof.* Proposition 5.12 implies by induction that  $\tilde{d}_i(u)$  is  $((1 + \varepsilon)^i, n\Delta/2^i)$ -smooth distance estimate. Using our choice of  $\varepsilon = \varepsilon_0 / (10t)$  we get that  $(1 + \varepsilon)^t \leq 1 + \varepsilon_0 / 2$ . On the other hand,  $n\Delta/2^t < \varepsilon_0 / 2$ . As we assume that all distances in  $G$  are at least 1, we conclude that  $\tilde{d}_t(v) - \tilde{d}_t(u) \leq (1 + \varepsilon_0 / 2) d_G(u, v) + \varepsilon_0 / 2 \leq (1 + \varepsilon_0) d_G(u, v)$ , as needed.

We note that we can assume that  $\varepsilon_0 = \Omega(1/(n\Delta))$  since otherwise one call to approximate distances already gives exact distances as the answer. Together with the assumptions that edge weights are polynomially bounded, we conclude that  $O(\log(n\Delta/\varepsilon_0)) = O(\log n)$ .

It remains to argue about the preservation of the tree-like property and undirected graphs. We will first observe that whenever  $\tilde{d}$  is tree-like in Algorithm 4, so is  $\tilde{d}'$ . To see this, consider any node  $v \in V(G)$ .

If  $\tilde{d}'(v) = \tilde{d}(v)$ , we use the tree-likeness of  $\tilde{d}$  to find  $u$  with  $\tilde{d}(u) \leq \tilde{d}(v) - \ell_G(u, v)$  and the fact  $\tilde{d}'(u) \leq \tilde{d}(u)$  shows that  $u$  also certifies the tree-likeness of  $\tilde{d}'$ .

On the other hand, assume that  $\tilde{d}'(v) = \tilde{d}_i(v) < \tilde{d}(v)$  for some  $i \in \{1, 2\}$ . Use the tree-likeness of  $\hat{d}_i$  to find some  $u$  with  $\hat{d}_i(u) \leq \hat{d}_i(v) - \ell_{H_i}(u, v)$ . Note that if  $u = \sigma$ , the tree-likeness of  $\hat{d}_i$  implies

$$0 = \hat{d}_i(\sigma) \leq \hat{d}_i(v) - \ell_{H_i}(\sigma, v) \leq \tilde{d}_i(v) - \tilde{d}(v) \tag{5.5}$$

where for the second inequality we used  $\tilde{d}_i(v) = \lfloor \tilde{d}(v) \rfloor_i + \hat{d}_i(v)$  and  $\tilde{d}(v) \leq \lfloor \tilde{d}(v) \rfloor_i + \ell_{H_i}(\sigma, v)$ . We thus get  $\tilde{d}_i(v) \geq \tilde{d}(v)$  which contradicts our assumption of  $\tilde{d}_i(v) < \tilde{d}(v)$ . So  $u$  has to be some node

of  $V(G)$  with  $\lfloor \tilde{d}(u) \rfloor_i = \lfloor \tilde{d}(v) \rfloor_i$ . But this implies  $\tilde{d}_i(u) \leq \tilde{d}_i(v) - \ell_G(u, v)$  and as  $\tilde{d}'(u) \leq \tilde{d}_i(u)$ , the tree-likeness for  $v$  follows.

Finally, one can check that all above results work also for undirected graphs. There, in the definition of  $H_i$  we simply add an undirected edge instead of a directed one. The undirectedness of the added edges may decrease distances between nodes of  $H_i$ . However, notice that in the proof above, whenever we argue about  $H_i$ , we only argue about  $d_{H_i[V(G)]}(u, v)$  or about  $d_{H_i}(\sigma, u)$  for some  $u, v \in V(G)$ . These distances are not distorted, hence the same proof applies to the undirected case.  $\square$

## 6 Tree-Constructing Algorithm

This section is dedicated to prove Theorem 6.1.

**Theorem 6.1.** *Obtaining tree-likeness* Given an undirected graph  $G$  with (real) weights in  $[1, \text{poly}(n)]$ , a source  $s \in V(G)$ , accuracy  $\varepsilon \in (0, 1]$ ,  $D = \text{poly}(n)$  an upper bound on the diameter of  $G$ , and an approximate oracle  $\mathcal{O}$ , Algorithm 7 computes a tree  $T$  such that the distance estimate  $\tilde{d}$  from  $s$  induced by  $T$  is  $(1 + \varepsilon)$ -approximate (and tree-like) in  $G$ , using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate distance oracle and  $O(\log^2 n)$  calls to a 1.1-approximate distance oracle.

We refer the reader to the discussion after Lemma 4.7 to see that in this setup constructing tree-like distances is really the same as constructing an approximate shortest path tree  $T$ . We in fact do the latter.

The algorithm is based on a simple idea: if  $G$  has radius  $D$  from  $s$ , we “split” the graph in the middle into the set  $S$  of nodes with distance roughly  $D/2$  from  $s$  and those that are further. We get a graph with radius roughly  $D/2$  where we solve the problem recursively and after this is done we “stitch” the trees together to get the final output tree (see Figure 6).

There is an important issue with this approach. Constructing the tree of  $S$  and of  $V(G) \setminus S$  can work only if there is no shortest path that would repeatedly go outside and inside of  $S$ . If this case happens, stitching the trees together can seriously distort distances.

Therefore, the set  $S$  cannot be simply chosen by first calling an approximate distance oracle and then adding all nodes with a distance estimate of at most  $D/2$  to  $S$ . Instead, we need to use as  $S$  the solution of the following ball cutting problem.

**Definition 6.2** (Simple ball cutting problem). Given an undirected graph  $G$  with nonnegative edge weights, a non-empty source set  $S \subseteq V(G)$ , and parameters  $D > 0$  and  $\varepsilon > 0$ . We ask for a set  $S^{sup} \subseteq V(G)$  with  $S^{sup} \supseteq S$  such that

1. for every  $v \in V(G)$ , whenever  $d_G(S, v) \leq D$ , we have  $v \in S^{sup}$  and
2. for every  $v \in S^{sup}$  we have  $d_{G[S^{sup}]}(S, v) \leq (1 + \varepsilon) \min(d_G(S, v), D)$ .

Note that this problem would be easy to solve with a tree-like approximate distance oracle: then the simple approach of running the approximate distance oracle once and taking all nodes with approximate distance at most  $(1 + \varepsilon)D$  works.

All algorithms in the rest of the section work also for directed graphs but we state them only for undirected graphs, as for directed graphs we can get exact distances and exact distances are clearly tree-like.

## 6.1 Ball Cutting

In this subsection we show how to solve the ball cutting problem from Definition 6.2.

**Lemma 6.3.** *Given an undirected graph  $G$  with minimum edge weight at least 1, a source set  $S \subseteq V(G)$ , accuracy  $\varepsilon \in (0, 1]$ ,  $D > 0$ , and an approximate oracle  $\mathcal{O}$ , Algorithm 5 solves the simple ball cutting problem from Definition 6.2 (with inputs  $G, S, D$  and  $\varepsilon$ ), using a single call to a  $(1 + \varepsilon/4)$ -approximate distance oracle and  $O(\max(\log D, 1))$  calls to a 1.1-approximate distance oracle.*

We use the following algorithm.

---

**Algorithm 5** Ball cutting algorithm BALLCUTTING( $G, S, \varepsilon, D$ )

---

**Input:** An undirected graph  $G$  with minimum edge weight at least 1, a source set  $S \subseteq V(G)$ , parameters  $\varepsilon \in (0, 1]$  and  $D > 0$

**Oracle:** For  $S' \subseteq V(G)$ ,  $\mathcal{O}(G, S', \varepsilon')$  returns  $(1 + \varepsilon')$ -approximate distance estimates in the graph we obtain from  $G$  by adding a supersource node and connect it to each node in  $S$  with an edge of length 0.

**Output:** A set  $S^{sup}$  satisfying properties from Definition 6.2.

- 1:  $t \leftarrow O(\max(\log D, 1))$
  - 2: Define  $\tilde{d}_0 \leftarrow \mathcal{O}(G, S, \varepsilon/4)$
  - 3:  $S_0 = \{v \in V(G) : \tilde{d}_0(v) \leq (1 + \varepsilon/4)D\}$
  - 4: **for**  $i \leftarrow 1, \dots, t$  **do**
  - 5:      $\tilde{d}_i(u) \leftarrow \mathcal{O}(G, S_{i-1}, 0.1)$
  - 6:      $S_i = \{v \in V(G) : \tilde{d}_i(v) \leq 0.5^{i+1} \cdot \varepsilon D\}$
  - 7: **return**  $S^{sup} \leftarrow S_t$
- 

We first observe that clearly whenever  $d_G(S, v) \leq D$ , we have  $\tilde{d}_0(v) \leq (1 + \varepsilon/4)D$  and hence  $v \in S_0$ . So the first property in Definition 6.2 is satisfied; we need to focus on the other one.

We also observe that we have  $S_{t-1} = S_t$  since to construct  $S_t$  we add nodes such that their approximate distance to  $S_{t-1}$  is at most  $(0.5)^{(t-1)+1} \varepsilon D \ll 1$ .

We note that for any  $v \in S_i$  it can even be  $d_{G[S_i]}(S, v) = +\infty$ . However, we prove that this problem is always fixed in the following step by enlarging  $S_i$  to  $S_{i+1}$ . This is made precise in the following two claims.

**Claim 6.4.** *For any  $v \in S_0$  we have*

$$d_{G[S_1]}(S, v) = d_G(S, v) \tag{6.1}$$

*Proof.* Let  $v \in S_0$ . By definition of  $S_0$  we have  $\tilde{d}_0(v) \leq (1 + \varepsilon/4)D$ , hence  $d_G(S, v) \leq \tilde{d}_0(v) \leq (1 + \varepsilon/4)D$ . Let  $P$  be a shortest path from  $S$  to  $v$  of length  $d_G(S, v)$ . Note that  $P$  is not necessarily a subset of  $S_0$ . However, since every node  $u$  with  $d_G(S, u) \leq D$  is certainly contained in  $S_0$  because  $\tilde{d}_0(u) \leq (1 + \varepsilon/4)d_G(S, u) \leq (1 + \varepsilon/4)D$ , we have for every node  $w \in P$  that  $d_G(S, w) \leq (\varepsilon/4)D$ . This implies that  $\tilde{d}_1(w) \leq 1.1 \cdot (\varepsilon/4)D \leq 0.5^{0+1} \varepsilon D$ , hence  $w \in S_1$ . We conclude that  $P \subseteq S_1$  and we are done.  $\square$

**Claim 6.5.** *Suppose that  $v \in S_i$  for any  $0 \leq i \leq t-1$ . Then,*

$$d_{G[S_{i+1}]}(S, v) \leq D + (1 - 0.5^{i+1})\varepsilon D. \quad (6.2)$$

*Proof.* We prove the claim by induction. For  $i = 0$  it follows from Claim 6.4 and the fact that any  $v \in S_0$  has  $d_G(S, v) \leq \tilde{d}_0(v) \leq (1 + \varepsilon/4)D \leq D + (1 - 0.5^{0+1})\varepsilon D$ .

Next, consider some  $i \in \{1, 2, \dots, t-1\}$ . Let  $P$  be any shortest path from some  $u \in S_{i-1}$  to  $v$  in  $G$ . By definition of  $S_i$  and the fact that  $\tilde{d}_i$  is noncontractive, we have  $d_G(u, v) = d_G(S_{i-1}, v) \leq 0.5^{i+1}\varepsilon D$ . We now proceed as in the proof of Claim 6.4 and prove that  $P \subseteq S_{i+1}$ .

To see this, note that any  $w \in P$  has  $d_G(S_{i-1}, w) \leq d_G(S_{i-1}, v) \leq \tilde{d}_i(v) \leq 0.5^{i+1}\varepsilon D$ . All nodes  $w \in P$  with  $d_G(S_{i-1}, w) \leq \frac{0.5^{i+1}\varepsilon D}{1.1}$  have however  $\tilde{d}_i(w) \leq 0.5^{i+1}\varepsilon D$  and hence are taken to  $S_i$ . Since  $P$  is a shortest path from  $S_i$  to  $v$ , this implies that any  $w \in P$  satisfies  $d_G(S_i, w) \leq 0.5^{i+1}\varepsilon D - \frac{0.5^{i+1}\varepsilon D}{1.1} \leq 0.5^{i+3}\varepsilon D$ . This means that  $\tilde{d}_{i+1}(w) \leq 1.1d_G(S_i, w) \leq 0.5^{i+2}\varepsilon D$  and hence  $w \in S_{i+1}$ .

Since  $P \subseteq S_{i+1}$ , we can now infer that  $d_{G[S_{i+1}]}(u, v) = d_G(u, v) \leq 0.5^{i+1}\varepsilon D$ . We finish by using the induction hypothesis on  $u$ ; we have

$$d_{G[S_{i+1}]}(S, v) \leq d_{G[S_i]}(S, u) + d_{G[S_{i+1}]}(u, v) \leq D + (1 - 0.5^{(i-1)+1} + 0.5^{i+1})\varepsilon D = D + (1 - 0.5^{i+1})\varepsilon D \quad (6.3)$$

and we are done.  $\square$

**Corollary 6.6.** *The second property from Definition 6.2 is satisfied; that is, for every  $v \in S$  we have  $d_{G[S^{sup}]}(S, v) \leq (1 + \varepsilon) \min(d_G(S, v), D)$ .*

*Proof.* For any  $v \in S^{sup}$  let  $i$  be the first index such that  $v \in S_i$ . Note that  $i \leq t-1$  since we already observed that  $S_{t-1} = S_t$ . If  $i = 0$ , we use Claim 6.4 to conclude that  $d_{G[S^{sup}]}(S, v) = d_G(S, v)$ . Since  $\tilde{d}_0$  is noncontractive we also have  $d_{G[S^{sup}]}(S, v) \leq (1 + \varepsilon/4)d_G(S, v)$ . Otherwise, if  $i \geq 1$ , we have  $d_G(S, v) > D$  and therefore we can directly use Claim 6.5.  $\square$

Our procedure for computing tree-like distance estimates actually requires a ball cutting algorithm that works on graphs with slightly more general edge weights. In particular, all of the edge weights are at least one except that edges incident to the source might have a nonnegative weight less than one. Algorithm 6 given below uses the ball cutting algorithm given above as a subroutine to handle this slightly more general case.

**Lemma 6.7.** *Given an undirected graph  $G$  with minimum edge weight at least 1 except that edges incident to source node  $s \in V(G)$  can also have edge weights in the interval  $[0, 1]$ , accuracy  $\varepsilon \in (0, 1]$ ,  $D > 0$ , and an approximate oracle  $\mathcal{O}$ , Algorithm 6 solves the simple ball cutting problem from Definition 6.2 (with inputs  $G, \{s\}, D$  and  $\varepsilon$ ), using a single call to a  $(1 + \varepsilon/8)$ -approximate distance oracle and  $O(\max(\log D, 1))$  calls to a 1.1-approximate distance oracle.*

*Proof.* First, observe that Lemma 6.3 is true even if we relax the condition that the minimum edge weight is 1. The only place where we use the minimum weight condition is to prove that  $S_{t-1} = S_t$ . This condition is even satisfied if the minimum edge weight is at least  $\varepsilon$  because  $(0.5)^{(t-1)+1}\varepsilon D \ll \varepsilon$ . We have to check the two conditions of Definition 6.2. If  $D < 1$ , then it follows from the fact that only edges incident to  $s$  can have a length of less than one that  $S = B_G(s, D)$ . Therefore, both conditions of Definition 6.2 are trivially satisfied. Thus, it remains

to consider the case that  $D \geq 1$ . To check the first condition, consider an arbitrary node  $v \in V(G)$ . If  $d_G(s, v) \leq D$ , then also  $d_G(S, v) \leq D$  and thus it follows from the fact that  $S^{sup}$  is a valid output for the ball cutting problem with input  $G, S, \varepsilon/2$  and  $D$  that  $v \in S^{sup}$ . To check the second condition, consider an arbitrary  $v \in S^{sup}$ . As  $S^{sup}$  is a valid output for the ball cutting problem with input  $G, S, \varepsilon/2$  and  $D$ , we get that  $d_{G[S^{sup}]}(S, v) \leq (1 + \varepsilon/2) \min(d_G(S, v), D)$ . We have to show that this implies  $d_{G[S^{sup}]}(s, v) \leq (1 + \varepsilon) \min(d_G(s, v), D)$ . First, if  $v \in S$ , then  $d_{G[S^{sup}]}(s, v) = \ell_G(s, v) \leq \min(d_G(s, v), D)$ . Next, consider the case that  $v \notin S$ . If  $d_G(s, v) \geq 1$ , then

$$d_{G[S^{sup}]}(s, v) \leq \varepsilon/2 + d_{G[S^{sup}]}(S, v) \leq \varepsilon/2 + (1 + \varepsilon/2) \min(d_G(S, v), D) \leq (1 + \varepsilon) \min(d_G(s, v), D).$$

On the other hand, if  $v \notin S$  and  $d_G(s, v) < 1$ , then

$$d_G(s, v) = d_G(S, v) \leq (1 + \varepsilon/2) \min(d_G(S, v), D) \leq (1 + \varepsilon) \min(d_G(s, v), D),$$

which finishes the proof. □

---

**Algorithm 6** Slightly more general ball cutting algorithm BALLCUTTING'(G, s, ε, D)

---

**Input:** An undirected graph  $G$  with minimum edge weight at least 1 except that edges incident to source node  $s \in V(G)$  can also have edge weights in the interval  $[0, 1]$ , parameters  $\varepsilon \in (0, 1]$  and  $D > 0$

**Output:** A set  $S^{sup}$  satisfying properties from Definition 6.2 with inputs  $G, \{s\}, D$  and  $\varepsilon$ .

- 1: **if**  $D < 1$  **then return**  $S^{sup} \leftarrow \{v \in V(G) : \ell_G(s, v) \leq D\}$
  - 2:  $S \leftarrow \{v \in V(G) : \ell_G(s, v) \leq \varepsilon/2\}$
  - 3: **return**  $S^{sup} \leftarrow \text{BallCutting}(G, S, \varepsilon/2, D)$
- 

## 6.2 Constructing the Tree

We are now ready to prove Theorem 6.1, the main result of this section.

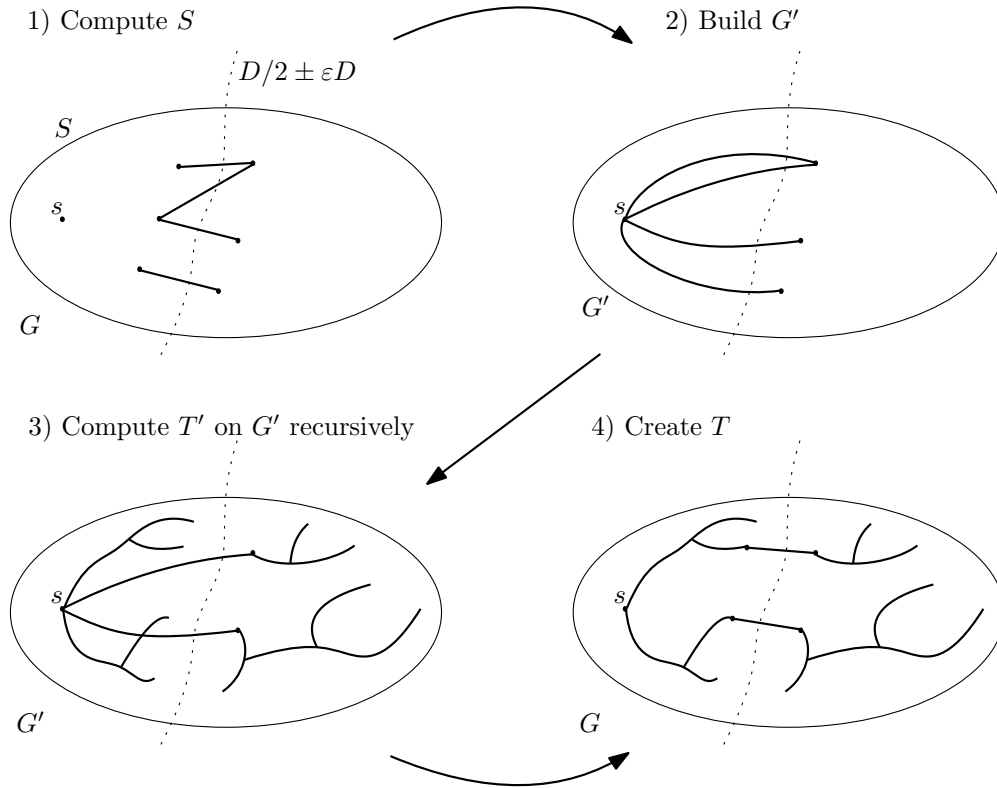
We construct the approximate shortest path tree using the following recursive algorithm in Algorithm 7, see also Figure 6.

We note that the definition of Line 4 in Algorithm 7 never creates negative length edges: whenever  $\ell_G(u, v) + \tilde{d}_0(u) - D/2 < 0$ , we then have  $d_G(s, v) \leq d_G(s, u) + \ell_G(u, v) \leq \tilde{d}_0(u) + \ell_G(u, v) < D/2$  where we first used triangle inequality, then the fact that  $\tilde{d}_0$  is noncontractive and then our assumption. Hence, Definition 6.2 implies that  $v \in S$ .

During the algorithm we may create edges of length 0 that connect a node to  $s$ . In fact, we note that whenever  $D < 1$ , it has to be that  $D = 0$  due to all lengths being integers. This means that all nodes are connected to  $s$  with a zero-length edge. This justifies Line 1 in Algorithm 7: the returned tree  $T$  in that case is the (trivial) shortest path tree that connects  $s$  with all other nodes by a zero-length edge.

Let  $\tilde{d}'$  be the distance estimate induced by  $T'$  and  $\tilde{d}$  the distance estimate induced by  $T$ . Theorem 6.1 follows by repeated application of the following claim for  $O(\log(nD)) = O(\log n)$  steps.





**Figure 6:** The figure describes Algorithm 7. In the first step, we compute the set  $S$  that cuts the graph into two halves, both with diameter about  $D/2$ . The set  $S$  has the additional property that any node  $v \in V(G)$  can be reached from  $s$  by a path that leaves  $S$  only at most once such that the path is not much longer than  $d_G(s, v)$ .

In the second step, we construct a new graph  $G'$  where each edge crossing a boundary from some  $u \in S$  to some  $v \in V(G) \setminus S$  is replaced by a direct edge from  $s$  to  $v$ . The length of the new edge is chosen so that the distance of  $v$  to  $s$  decreases roughly by  $D/2$ .

In the third step, we construct the approximate shortest path tree  $T'$  recursively in  $G'$ .

Finally, we replace all edges of  $T'$  that are not in  $G$  by their counterparts. If all computations were exact, this recursive algorithm would construct a shortest path tree. We lose a  $(1 + \varepsilon)$ -factor in one recursion level due to inexact calculations. In total, we lose  $(1 + \varepsilon)^{O(\log D)} = (1 + O(\varepsilon \cdot \log D))$  as we choose  $\varepsilon < 1/\log D$ .

---

**Algorithm 7** Tree-constructing algorithm TREECONSTRUCTING( $G, D$ )

---

**Input:** An undirected graph  $G$  with minimum edge weight at least 1 and diameter upper bound  $D$ , source  $s$ , parameter  $\varepsilon < O(1/\log D)$ , approximate oracle  $\mathcal{O}(G, s, \varepsilon)$ . (During recursive calls, edges incident to the source  $s$  can have arbitrary nonnegative edge lengths)

**Oracle:**  $\mathcal{O}(G, s, \varepsilon)$  returns  $(1 + \varepsilon)$ -approximate distances  $\tilde{d} : V(G) \rightarrow \mathbb{R}_{\geq 0}$  from  $s$  in  $G$ .

**Output:** A tree  $T$  such that the distance estimate  $\tilde{d}$  from  $s$  induced by  $T$  is  $(1 + \varepsilon \cdot O(\log D))$ -approximate (and tree-like).

- 1: If  $D < 1$  return the star  $T$  connecting  $s$  with every other node of  $G$
  - 2:  $\tilde{d}_0 \leftarrow \mathcal{O}(G, s, \varepsilon)$
  - 3:  $S \leftarrow \text{BALLCUTTING}'(G, s, \varepsilon, D/2)$
  - 4: Define  $G'$  by replacing each edge  $e = \{u, v\}$  with  $u \in S, v \notin S$  by an edge  $e_{u,v}$  from  $s$  to  $v$  with weight  $\ell_G(u, v) + \tilde{d}_0(u) - D/2$
  - 5:  $T' \leftarrow \text{TREECONSTRUCTING}(G', (1 + \varepsilon)D/2)$
  - 6: Construct  $T$  by replacing each edge  $e_{u,v}$  in  $T'$  by the edge  $\{u, v\}$
  - 7: **return**  $T$
- 

**Claim 6.8.** Assume that the distance estimate  $\tilde{d}$  is  $(1 + \varepsilon_0)$ -approximate for  $\varepsilon_0 < 1$ . Then  $\tilde{d}$  is  $(1 + \varepsilon_0 + 5\varepsilon)$ -approximate.

*Proof.* Consider any  $w \in V(G)$ . First assume that  $w \in S$ . In that case by the definition of  $\tilde{d}$  we have  $\tilde{d}(w) = \tilde{d}'(w)$  and the guarantees of  $\tilde{d}'(w)$  say that  $\tilde{d}'(w) \leq (1 + \varepsilon_0)d_{G[S]}(s, w)$ . Using Definition 6.2, we have  $d_{G[S]}(s, w) \leq (1 + \varepsilon)d_G(s, w)$ . We conclude that  $\tilde{d}(w) \leq (1 + \varepsilon_0)(1 + \varepsilon)d_G(s, w)$  as needed.

Next, we assume  $w \notin S$ . Consider a shortest path  $P_0$  from  $s$  to  $w$  and let the edge  $\{u_0, v_0\}$  be the last edge of  $P_0$  such that  $u_0 \in S$  and  $v_0 \in V(G) \setminus S$ . Similarly, consider the path  $P$  that connects  $s$  and  $w$  in  $T$ . Let  $\{u, v\}$  be the unique edge of it such that  $u \in S, v \notin S$  (see Figure 7). Our task is to upper bound the length of  $P$  in terms of the length of  $P_0$ .

Formally, using the definition of  $T$  in terms of  $T'$  in Line 6 we write

$$\tilde{d}(w) = \tilde{d}'(u) + \ell_G(u, v) + \tilde{d}'(w) - \tilde{d}'(v) \quad (6.4)$$

We will bound separately  $\tilde{d}'(w)$  and the rest of the right hand side. See Figure 7, the bound of  $\tilde{d}'(w)$  corresponds to the comparison of  $P$  and  $P_0$  to the right of the boundary between  $S$  and  $V(G) \setminus S$ ; the bound on  $\tilde{d}'(u) + \ell_G(u, v) - \tilde{d}'(v)$  corresponds to the parts of the paths to the left of the boundary.

**First upper bound:** To bound  $\tilde{d}'(w)$ , we first note that  $\tilde{d}'(w) \leq (1 + \varepsilon_0)d_{G'}(s, w)$  so we need to bound the value of  $d_{G'}(s, w)$ . We have

$$d_{G'}(s, w) \leq \ell_{G'}(s, v_0) + d_G(v_0, w) = \ell_G(u_0, v_0) + \tilde{d}_0(u_0) - D/2 + d_G(v_0, w)$$

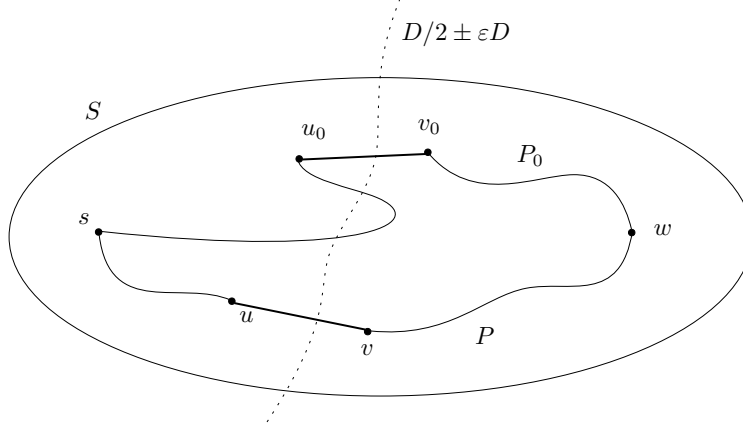
where we used that  $\{u_0, v_0\}$  is the last edge of  $P_0$  crossing  $S$ , hence  $d_{G[V(G) \setminus S]}(v_0, w) = d_G(v_0, w)$ . Moreover, we can bound  $\tilde{d}_0(u_0) \leq (1 + \varepsilon)d_G(s, u_0)$ . So, we can write

$$d_{G'}(s, w) \leq (1 + \varepsilon) (\ell_G(u_0, v_0) + d_G(s, u_0) + d_G(v_0, w)) - D/2$$

and hence we can combine above bounds to get

$$\tilde{d}'(w) \leq (1 + \varepsilon_0) (1 + \varepsilon) (\ell_G(u_0, v_0) + d_G(s, u_0) + d_G(v_0, w)) - (1 + \varepsilon_0)D/2 \quad (6.5)$$

$$= (1 + \varepsilon_0) (1 + \varepsilon) d_G(s, w) - (1 + \varepsilon_0)D/2 \quad (6.6)$$



**Figure 7:** A figure for the proof of Claim 6.8. We consider the optimal path  $P_0$  from  $s$  to  $w$  via  $u_0, v_0$  and the path  $P$  taken by the constructed tree  $T$  that goes through the edge  $u, v$ .

**Second upper bound:** Next, we need to upper bound the rest of the right hand side of (6.4), i.e., the expression  $\tilde{d}'(u) + \ell_G(u, v) - \tilde{d}'(v)$ . To do so, we first write

$$\tilde{d}'(u) \leq (1 + \varepsilon_0)d_{G[S]}(s, u) \leq (1 + \varepsilon_0)(1 + \varepsilon)d_G(s, u)$$

where we used the inductive assumption on  $\tilde{d}'$  together with  $d_{G[S]}(s, u) = d_{G'}(s, u)$  and then the property of  $S$  from Definition 6.2. Then we write

$$\tilde{d}'(v) = \ell_{G'}(\{s, v\}) = \ell_G(u, v) + \tilde{d}_0(u) - D/2 \geq \ell_G(u, v) + d_G(s, u) - D/2$$

where we used that  $\tilde{d}_0$  is noncontractive. Plugging in, we get

$$\tilde{d}'(u) + \ell_G(u, v) - \tilde{d}'(v) \leq (1 + \varepsilon_0)(1 + \varepsilon)d_G(s, u) + \ell_G(u, v) - \ell_G(u, v) - d_G(s, u) + D/2 \quad (6.7)$$

$$\leq (\varepsilon_0 + 2\varepsilon)d_G(s, u) + D/2. \quad (6.8)$$

**Putting it together** Putting (6.6) and (6.8) together, we conclude that

$$\tilde{d}(w) \leq (1 + \varepsilon_0)(1 + \varepsilon)d_G(s, w) - \varepsilon_0 D/2 + (\varepsilon_0 + 2\varepsilon)d_G(s, u) \quad (6.9)$$

Since we have  $d_G(s, u) \leq (1 + \varepsilon)D/2$  by the second property in Definition 6.2, we can bound the error terms of the right hand side by

$$-\varepsilon_0 D/2 + (\varepsilon_0 + 2\varepsilon)(1 + \varepsilon)D/2 \leq 3\varepsilon \cdot D/2.$$

Finally, since  $d_G(s, w) \geq D/2$  by definition of  $S$  (Definition 6.2), we can simplify (6.9) to

$$\tilde{d}(w) \leq (1 + \varepsilon_0 + 3\varepsilon)d_G(s, w). \quad (6.10)$$

□

## References

- [AFGW10] Ittai Abraham, Amos Fiat, Andrew V Goldberg, and Renato F Werneck. Highway dimension, shortest paths, and provably efficient algorithms. In *Proceedings of the 21st Proc. Symposium on Discrete Algorithms (SODA)*, pages 782–793. SIAM, 2010.

- [ASZ20] Alexandr Andoni, Clifford Stein, and Peilin Zhong. Parallel approximate undirected shortest paths via low hop emulators. 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 322–335. ACM, 2020.
- [BEL20] Ruben Becker, Yuval Emek, and Christoph Lenzen. Low diameter graph decompositions by approximate distance computation. In *ITCS*, 2020.
- [BFKL21] Ruben Becker, Sebastian Forster, Andreas Karrenbauer, and Christoph Lenzen. Near-optimal approximate shortest paths and transshipment in distributed and streaming models. *SIAM Journal on Computing*, 50(3):815–856, 2021.
- [BGK<sup>+</sup>14] G. E. Blelloch, A. Gupta, I. Koutis, G. L. Miller, R. Peng, and K. Tangwongsan. Nearly-linear work parallel SDD solvers, low-diameter decomposition, and low-stretch subgraphs. *Theory Comput. Syst.*, 55(3):521–554, 2014.
- [BLP20] Uri Ben-Levy and Merav Parter. New  $(\alpha, \beta)$  spanners and hopsets. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1695–1714. SIAM, 2020.
- [BNWN22] Aaron Bernstein, Danupon Nanongkai, and Christian Wulff-Nilsen. Negative-weight single-source shortest paths in near-linear time, 2022.
- [BW22] Aaron Bernstein and Nicole Wein. Closing the gap between directed hopsets and shortcut sets. *arXiv preprint arXiv:2207.04507*, 2022.
- [CFR20a] Nairen Cao, Jeremy T. Fineman, and Katina Russell. Efficient construction of directed hopsets and parallel approximate shortest paths. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, STOC 2020*, page 336–349, New York, NY, USA, 2020. Association for Computing Machinery.
- [CFR20b] Nairen Cao, Jeremy T Fineman, and Katina Russell. Improved work span tradeoff for single source reachability and approximate shortest paths. In *Proceedings of the 32nd Proc. Symposium on Parallel Algorithms and Architecture (SPAA)*, pages 511–513, 2020.
- [CFR21] Nairen Cao, Jeremy T. Fineman, and Katina Russell. Brief announcement: An improved distributed approximate single source shortest paths algorithm. In *Proceedings of the 2021 ACM Symposium on Principles of Distributed Computing, PODC’21*, page 493–496, New York, NY, USA, 2021. Association for Computing Machinery.
- [Che14] Shiri Chechik. Approximate distance oracles with constant query time. In *Proceedings of the 46th Proc. Symposium on Theory of Computation (STOC)*, pages 654–663, 2014.
- [CM20] Shiri Chechik and Doron Mukhtar. Single-source shortest paths in the congest model with improved bound. In *Proceedings of the 39th Symposium on Principles of Distributed Computing*, pages 464–473, 2020.
- [CMMS98] Andreas Crauser, Kurt Mehlhorn, Ulrich Meyer, and Peter Sanders. A parallelization of dijkstra’s shortest path algorithm. In *International Symposium on Mathematical Foundations of Computer Science*, pages 722–731. Springer, 1998.

- [Coh96] Edith Cohen. Efficient parallel shortest-paths in digraphs with a separator decomposition. *Journal of Algorithms*, 21(2):331–357, 1996.
- [Coh00] Edith Cohen. Polylog-time and near-linear work approximation scheme for undirected shortest paths. *J. ACM*, 47(1):132–166, 2000.
- [CZ98] Shiva Chaudhuri and Christos D Zaroliagis. Shortest paths in digraphs of small treewidth. part ii: Optimal parallel algorithms. *Theoretical Computer Science*, 203(2):205–223, 1998.
- [DBS17] Laxman Dhulipala, Guy Blelloch, and Julian Shun. Julienne: A framework for parallel graph algorithms using work-efficient bucketing. In *Proceedings of the 29th ACM Symposium on Parallelism in Algorithms and Architectures*, pages 293–304, 2017.
- [Dij59] Edsger Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.
- [EHRG22] Michael Elkin, Bernhard Haeupler, Václav Rozhoň, and Christoph Grunau. Deterministic low-diameter decompositions for weighted graphs and distributed and parallel applications, 2022.
- [EN19] Michael Elkin and Ofer Neiman. Hopsets with constant hopbound, and applications to approximate shortest paths. *SIAM Journal on Computing*, 48(4):1436–1480, 2019.
- [EN20] Michael Elkin and Ofer Neiman. Near-additive spanners and near-exact hopsets, a unified view. *arXiv preprint arXiv:2001.07477*, 2020.
- [Feu19] Laurent Feuilloley. Introduction to local certification. *CoRR*, abs/1910.12747, 2019.
- [Fin18] Jeremy T Fineman. Nearly work-efficient parallel algorithm for digraph reachability. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 457–470, 2018.
- [FN18] Sebastian Forster and Danupon Nanongkai. A faster distributed single-source shortest paths algorithm. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 686–697. IEEE, 2018.
- [FR84] Alan M Frieze and Larry Rudolph. *A parallel algorithm for all pairs shortest paths in a random graph*. Management Sciences Research Group, Graduate School of Industrial . . . , 1984.
- [FR06] Jittat Fakcharoenphol and Satish Rao. Planar graphs, negative weight edges, shortest paths, and near linear time. *Journal of Computer and System Sciences*, 72(5):868–889, 2006.
- [FSR06] Liping Fu, Dihua Sun, and Laurence R Rilett. Heuristic shortest path algorithms for transportation applications: State of the art. *Computers & Operations Research*, 33(11):3324–3343, 2006.
- [GL18] Mohsen Ghaffari and Jason Li. Improved distributed algorithms for exact shortest paths. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 431–444, 2018.

- [HKRS97] Monika R Henzinger, Philip Klein, Satish Rao, and Sairam Subramanian. Faster shortest-path algorithms for planar graphs. *Journal of Computer and System Sciences*, 55(1):3–23, 1997.
- [HL18] Bernhard Haeupler and Jason Li. Faster Distributed Shortest Path Approximations via Shortcuts. In *Proceedings of the 32nd Proc. Symposium on DIStributed Computing (DISC)*, pages 33:1–33:14, 2018.
- [HNR68] Peter E Hart, Nils J Nilsson, and Bertram Raphael. A formal basis for the heuristic determination of minimum cost paths. *IEEE transactions on Systems Science and Cybernetics*, 4(2):100–107, 1968.
- [HP19] Shang-En Huang and Seth Pettie. Thorup–zwick emulators are universally optimal hopsets. *Information Processing Letters*, 142:9–13, 2019.
- [HP21] Shang-En Huang and Seth Pettie. Lower bounds on sparse spanners, emulators, and diameter-reducing shortcuts. *SIAM Journal on Discrete Mathematics*, 35(3):2129–2144, 2021.
- [JLS19] Arun Jambulapati, Yang P Liu, and Aaron Sidford. Parallel reachability in almost linear work and square root depth. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1664–1686. IEEE, 2019.
- [Joh77] Donald B Johnson. Efficient algorithms for shortest paths in sparse networks. *J. ACM*, 24(1):1–13, 1977.
- [KP22a] Shimon Kogan and Merav Parter. Beating Matrix Multiplication for  $n^{\{1/3\}}$ -Directed Shortcuts. In *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 82:1–82:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik.
- [KP22b] Shimon Kogan and Merav Parter. New diameter-reducing shortcuts and directed hopsets: Breaking the barrier. In *Proceedings of the 33rd Proc. Symposium on Discrete Algorithms (SODA)*, pages 1326–1341. SIAM, 2022.
- [KS93] Philip N Klein and Sairam Subramanian. *A linear-processor polylog-time algorithm for shortest paths in planar graphs*. IEEE, 1993.
- [KS97] Philip N Klein and Sairam Subramanian. A randomized parallel algorithm for single-source shortest paths. *Journal of Algorithms*, 25(2):205–220, 1997.
- [KT19] Michael Kainer and Jesper Larsson Träff. More parallelism in dijkstra’s single-source shortest path algorithm. *arXiv preprint arXiv:1903.12085*, 2019.
- [Li20] Jason Li. Faster parallel algorithm for approximate shortest path. 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 308–321. ACM, 2020.
- [LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

- [MBBC07] Kamesh Madduri, David A Bader, Jonathan W Berry, and Joseph R Crobak. An experimental study of a parallel shortest path algorithm for solving large-scale graph instances. In *Proceedings of the 9th Proc. ALENEX*, pages 23–35. SIAM, 2007.
- [MPX13] Gary L Miller, Richard Peng, and Shen Chen Xu. Parallel graph decompositions using random shifts. In *Proceedings of the twenty-fifth annual ACM symposium on Parallelism in algorithms and architectures*, pages 196–203, 2013.
- [MS98] Ulrich Meyer and Peter Sanders.  $\delta$ -stepping: A parallel single source shortest path algorithm. In *European symposium on algorithms*, pages 393–404. Springer, 1998.
- [MWBA10] Richard C Murphy, Kyle B Wheeler, Brian W Barrett, and James A Ang. Introducing the graph 500. *Cray Users Group (CUG)*, 19:45–74, 2010.
- [RGH<sup>+</sup>22] Václav Rozhoň, Christoph Grunau, Bernhard Haeupler, Goran Zuzic, and Jason Li. Undirected (1+ $\epsilon$ )-shortest paths via minor-aggregates: Near-optimal deterministic parallel and distributed algorithms. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022, page 478–487, New York, NY, USA, 2022. Association for Computing Machinery.
- [SB13] Julian Shun and Guy E Blelloch. Ligra: a lightweight graph processing framework for shared memory. In *Proceedings of the 18th ACM SIGPLAN symposium on Principles and practice of parallel programming (PPoPP)*, pages 135–146, 2013.
- [She17] Jonah Sherman. Generalized preconditioning and undirected minimum-cost flow. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 772–780. SIAM, 2017.
- [Spe91] Thomas H Spencer. More time-work tradeoffs for parallel graph algorithms. In *Proceedings of the third annual ACM symposium on Parallel algorithms and architectures*, pages 81–93, 1991.
- [SSP<sup>+</sup>14] Nadathur Satish, Narayanan Sundaram, Md Mostofa Ali Patwary, Jiwon Seo, Jongsoo Park, M Amber Hassaan, Shubho Sengupta, Zhaoming Yin, and Pradeep Dubey. Navigating the maze of graph analytics frameworks using massive graph datasets. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pages 979–990, 2014.
- [STV95] Sairam Subramanian, Roberto Tamassia, and Jeffrey Scott Vitter. An efficient parallel algorithm for shortest paths in planar layered digraphs. *Algorithmica*, 14(4):322–339, 1995.
- [Tho92] Mikkel Thorup. On shortcutting digraphs. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 205–211. Springer, 1992.
- [TZ05] Mikkel Thorup and Uri Zwick. Approximate distance oracles. *Journal of the ACM (JACM)*, 52(1):1–24, 2005.
- [TZ06] Mikkel Thorup and Uri Zwick. Spanners and emulators with sublinear distance errors. In *Proceedings of the 17th Proc. Symposium on Discrete Algorithms (SODA)*, pages 802–809, 2006.

- [US12] Koji Ueno and Toyotaro Suzumura. Highly scalable graph search for the graph500 benchmark. In *Proceedings of the 21st international symposium on High-Performance Parallel and Distributed Computing*, pages 149–160, 2012.
- [UY91] Jeffrey D Ullman and Mihalis Yannakakis. High-probability parallel transitive-closure algorithms. *SIAM Journal on Computing*, 20(1):100–125, 1991.
- [ZGY<sup>+</sup>22] Goran Zuzic, Goramoz Goranci, Mingquan Ye, Bernhard Haeupler, and Xiaorui Sun. Universally-optimal distributed shortest paths and transshipment via graph-based l1-oblivious routing. SIAM, 2022.
- [Zuz21] Goran Zuzic. A simple boosting framework for transshipment. *arXiv preprint arXiv:2110.11723*, 2021.

## A What Can and Cannot be Done with Strong Distances

In this section we first show in Section A.1 how the notions of strong distances can be used in practice to greatly simplify proofs of some known results. We then, on the other hand, show in Section A.2 a problem solvable with exact distance oracle that we do not know how to solve even with the power of strong distances on our side.

### A.1 A Success Story: Ball Growing

Here we give a simple demonstration of the strength of strong distances. The problem that we solve is also called “blurring” or “blurry ball growing”. In that problem, we are given a node (or in general a set of nodes) and we want to find a cluster around it of diameter  $D$  such that every edge  $e$  is cut by this ball only with probability  $\ell_G(e)/D$ . The cluster should be implementable in a parallel/distributed manner with a few calls to a distance oracle.

We note that proofs of such results in [BEL20] and [EHRG22] are nontrivial and somewhat technical. This is the case even for [EHRG22] where the authors work with the approximate distance oracle from [RGH<sup>+</sup>22] that, as explained in Section 4.5, gives not only approximate tree-like distances, but also approximate potential as a bonus (so we have both an approximate distance and potentials, but they are two different functions; in contrast a strong distance estimate has both properties at once).

This is in stark contrast with the fact that the result is trivial if we have access to an *exact* distance oracle. With our transformation from off-the-shelf approximations to strong distances (Theorem 1.10) at hand, the proof is now not much more difficult than in the exact distance case as we show next.

**Theorem A.1.** *[A variant of Theorem 5 in [BEL20], Theorem 1.5 in [EHRG22]] Given an undirected weighted graph  $G$ , its node  $s$  and a parameter  $D > 0$ , there is a randomized algorithm computing a set  $S$  containing  $s$  such that*

$$\max_{v \in S} d_{G[S]}(s, v) \leq D \tag{A.1}$$

*and moreover every edge  $e \in E(G)$  has one endpoint in  $S$  and the other in  $V(G) \setminus S$  with probability at most  $O(\ell(e)/D)$ . The algorithm uses poly  $\log n$  calls to a  $(1 + 1/\text{poly } \log n)$ -approximate distance oracle.*

Let us now show a quick proof of this theorem.



*Proof.* We first observe that the result would be simple to prove should we have access to an *exact* distance oracle  $\mathcal{O}$ . In that case, we simply choose a uniformly random number  $\tilde{D} \in [0, D)$  and define  $S$  as the set of nodes of distance at most  $\tilde{D}$  from  $s$ . This clearly makes (A.1) satisfied since whenever a node  $v$  is included in  $S$ , so are all the nodes on the shortest path from  $s$  to  $v$ . Moreover, every edge  $e = \{u, v\}$  is cut with probability  $|d_G(s, u) - d_G(s, v)|/D \leq \ell_G(e)/D$  as needed.

Going back to the approximate distance oracle, we use Theorem 1.10 with  $\varepsilon = 1$  to construct exact distances in some graph  $G'$  with weights  $\ell_G \leq \ell_{G'} \leq 2\ell_G$  and then follow the above algorithm for exact distances but on  $G'$  instead of  $G$ . On one hand, we have  $\max_{v \in S} d_{G'[S]}(S, v) \leq \max_{v \in S} d_{G'[S]}(S, v) \leq D$ , on the other hand the probability of an edge being cut is now at most  $|d_{G'}(s, u) - d_{G'}(s, v)|/D \leq \ell_{G'}(e)/D \leq 2\ell_G(e)/D$  and we are done.  $\square$

## A.2 A Cautionary Tale: MPX

We continue by describing a problem which is simple to solve with exact distance oracle but is not directly solved with approximate distances, even in view of our transformations.

The problem is simulating the randomized clustering algorithm of [MPX13] (called MPX next). In their algorithm, we have an undirected graph  $G$  and a parameter  $D > 0$ . During the algorithm, each node  $v$  first samples a random “head start”  $h(v)$  which is a number from the exponential distribution with mean  $D$ . Next, we add a virtual node  $\sigma$  and connect it with each node  $v \in V(G)$  with an edge of length  $O(D \log n) - h(v)$ , the first term being there only to make the weight positive. Then we run the exact distance oracle from  $\sigma$ . The returned shortest path tree defines clusters of  $G$  and this is the output of the MPX algorithm.

It can be proven that, on one hand, with high probability all head starts satisfy  $h(v) = O(D \log n)$ , hence all edges have positive weight and the diameter of all clusters is at most  $O(D \log n)$ . On the other hand, each edge  $e$  is cut by the clustering (that is, its endpoints are in different clusters) with probability  $O(\ell(e)/D)$  (compare with Theorem A.1).

It is not clear at all how to simulate this algorithm with the access to even a strongly  $(1 + \varepsilon)$ -approximate-distance oracle, that is, the distance oracle that returns exact distances on a slightly perturbed graph.

To see this, let us leave the algorithm exactly the same except of using the strong-distance oracle instead of the exact distance oracle. The problem is that the distances in the graph  $G \cup \{\sigma\}$  are (adversarially) perturbed by only *after* the head starts  $h(\cdot)$  are sampled. This is not a problem for the cluster diameter, which grows only by a  $(1 + \varepsilon)$  multiplicative factor. However, the guarantee that every edge  $e$  is cut only with probability  $O(\ell(e)/D)$  does not hold anymore!

To see this, consider a simple path-graph  $G$  consisting of four nodes  $u_0, u, v_0, v$  with  $\ell_G(u_0, u) = \ell_G(v_0, v) = D$  and  $\ell_G(u, v) = 1$ . With constant probability, the head starts of both  $u_0$  and  $v_0$  are more than  $D$  larger than both head starts  $u, v$ . Moreover, with probability  $\Omega(\varepsilon)$  we additionally have  $|h(u_0) - h(v_0)| \leq \varepsilon D$ . In this case the algorithm constructs two clusters rooted in  $u_0$  and  $v_0$ . Unfortunately, the adversarial perturbation of edge lengths can be such that in the perturbed graph we always cut the edge  $\{u, v\}$ . This means that we cut  $\{u, v\}$  with probability  $\Omega(\varepsilon)$ , while we require it to be cut only with probability  $O(1/D)$  which would be achieved by exact distance oracle.

## B Undirected APSP Variant of Smoothing

In this section we show how to adapt the algorithm of Section 5 so that it works in undirected graphs and in the more general context of APSP. We formally prove the following Theorem B.1.

**Theorem B.1.** *Algorithm 3 and Algorithm 8 compute  $(1 + \varepsilon)$ -approximate smooth APSP distances  $\tilde{d}(u, v)$  in an undirected graph  $G$  using  $O(\log n)$  calls to a  $(1 + O(\varepsilon/\log n))$ -approximate APSP distance oracle  $\mathcal{O}$  on undirected graphs. Moreover,*

1. *If  $\mathcal{O}$  returns tree-like distances,  $\tilde{d}(s, \cdot)$  is also tree-like for every  $s \in V(G)$ .*
2. *If  $\mathcal{O}$  computes only single source approximate distances, the algorithm solves the single source variant of the problem.*

Notice that the second additional part of our theorem (undirected SSSP) is also a special case of our main theorem Theorem 5.1. That is, there are two (slightly) different smoothing algorithms for single source undirected smoothing.

The algorithm we use is again Algorithm 3 but we replace its inner loop from Algorithm 4 by its slightly modified version Algorithm 8. Let us first describe the necessary changes to the algorithm, then we analyse it by closely following the analysis of Section 5.2.

The main challenge of the APSP case is the construction of the level graph  $H_i$ . In the new construction we add a new node  $\sigma_u$  to  $G$  for each node  $u \in V(G)$ , instead of just one node  $\sigma$  as in the SSSP case. The problem is with cutting edges going between different levels — every node  $\sigma_u$  would like the graph to be partitioned into different levels and hence would like to cut a different set of edges! Fortunately, in the undirected case we can make the smoothing construction work without cutting edges: We simply do not update nodes that are in the “upper half” of each level, since the shortest path to those nodes in the level graph can start at the above level (see Figure 8 and Line 6 in Algorithm 8).

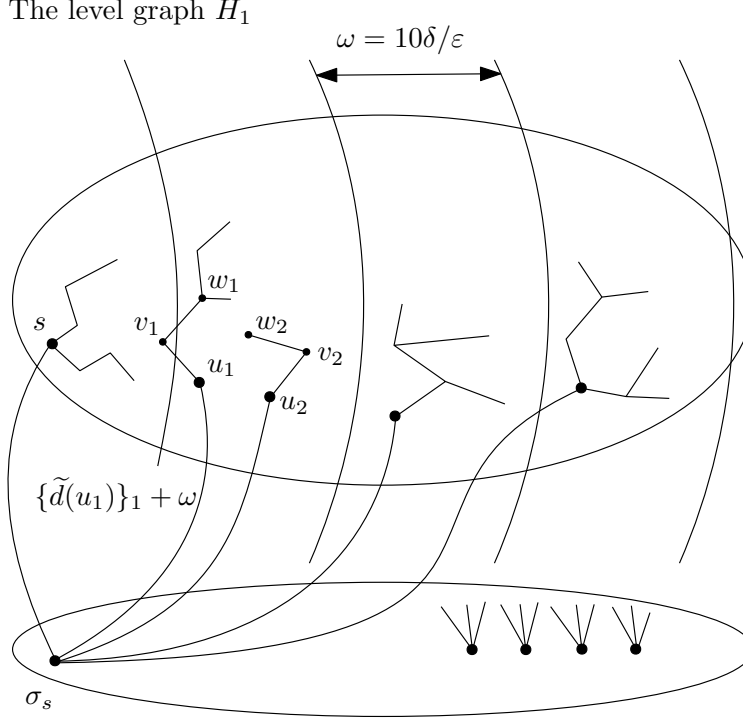
This approach would not work in the directed case where we can have a very short edge  $(u, v)$  from a node  $u$  with very large  $\lfloor \tilde{d}(u) \rfloor$  but  $\{\tilde{d}(u)\} = 0$  to a node  $v$  with smaller  $\tilde{d}(v) \ll \tilde{d}(u)$ . If we do not cut  $(u, v)$  and run the distance oracle in the level graph  $H_i$ , we simply learn that  $\tilde{d}_i(s, v)$  is close to 0, which is a meaningless information. Fortunately, this bad case simply cannot occur in undirected graphs: there, a short edge between  $u$  and  $v$  implies that  $u$  lies in the same level as  $v$  (unless  $v$  lies at the boundary of its level, in which case we need to argue more carefully).

There are a few other minor changes in Algorithm 8 with respect to Algorithm 4. Due to technical reasons, we need three level graphs instead of two and we need to “slow down” also the newly added edges connecting  $V(G)$  with  $V(H_i) \setminus V(G)$ , see Line 1 in Algorithm 8.

In Theorem B.1 we consider the case when the distance oracle returns only SSSP, not APSP distances. This is a special case of the following analysis with APSP oracles, so we next analyse only the case when  $\mathcal{O}$  is an APSP oracle.

**Definitions and the algorithm.** We define rounding analogously to Definition 5.6.

**Definition B.2** (Rounding distances). Let  $G$  be a graph and  $\tilde{d}$  a distance estimate on it. Let  $i \in \{1, 2, 3\}$  and  $\omega$  a parameter. We define  $\{\tilde{d}(s, v)\}_i$  to be the smallest value such that  $\tilde{d}(s, v) - \{\tilde{d}(s, v)\}_i - (i - 1) \cdot \frac{\omega}{3}$  is divisible by  $\omega$ . We also define  $\tilde{d}(s, v) = \lfloor \tilde{d}(s, v) \rfloor_i + \{\tilde{d}(s, v)\}_i$ .



**Figure 8:** The figure shows the construction of the level graph  $H_1$  in the undirected APSP case. We add a node  $\sigma_s$  for each  $s \in V(G)$ . No edges of the original graph  $G$  are cut. The edges from  $\sigma_s$  to other nodes in  $V(G)$  are defined as in the directed case but with additional term  $+\omega$  so that  $\sigma_s$  is not affected by other nodes  $\sigma_{s'}$ .

We want to update in Algorithm 8 only the distance estimate of nodes that lie in the “first half” of each level, i.e.  $0 \leq \{\tilde{d}(v)\}_1 \leq \omega/2$ . This is because nodes close to the upper end of a level are affected by short distances to the low end of the next level. To make the new distances tree-like, whenever we update nodes  $w_1, w_2$  with  $0 \leq \{\tilde{d}(w_1)\}_1, \{\tilde{d}(w_2)\}_1 \leq \omega/2$ , we also need to update their predecessors in the tree constructed from the tree-like property of  $\tilde{d}$ . Namely, we update  $v_1$  although it lies in a smaller level but we behave as if its level is the same as the level of  $u_1$ . We also update the value of  $v_2$  although  $\{\tilde{d}(v_2)\}_1 > \omega/2$ .

Next, let us define the level graph analogously to Definition 5.7 with the main difference being that  $H_i$  contains all edges of  $G$ , i.e., we are not cutting edges (see also Figure 8).

**Definition B.3** (Level graph  $H_i$  for undirected APSP). Given an input undirected graph  $G$ , a width parameter  $\omega$  and  $i \in \{1, 2, 3\}$ , we define the *APSP level graph*  $H_i$  as follows.

We define  $V(H_i) = V(G) \cup \bigcup_{u \in V(G)} \sigma_u$ . The set of edges  $E(H_i)$  is a superset of  $E(G)$ , with  $\ell_{H_i}(u, v) = \alpha(1 + \varepsilon)\ell_G(u, v)$ . Additionally, for each  $\sigma_u$  and  $v$  we add an edge  $\{\sigma_u, v\}$  to  $E(H_i)$  for any  $u \in V(G)$  of weight

$$(1 + \varepsilon/10^3)\{\tilde{d}(u, v)\}_i + \omega.$$

Comparing with Definition 5.7, we have additional term  $+\omega$  in the definition of edges  $\{\sigma_u, v\}$ . This is because we do not want the shortest paths from  $\sigma_u$  to nodes of  $V(G)$  be affected by other edges  $\{\sigma_{u'}, v\}$  in  $H_i$ . This would not be necessary if we were allowed to make the edges of type  $\{\sigma_u, v\}$  directed. The term  $+\omega$  implies additional distortion of distances, but since we choose  $(1 + \varepsilon/10^4)$ -approximate oracle and  $\varepsilon\omega/10^4 \ll \delta/2$ , this additional error term does not create new difficulties.

We also “slow down” the edge by  $(1 + \varepsilon/10^3)$ . This is only needed to argue about the final distances being tree-like at the end of the proof.

---

**Algorithm 8** Partial Smoothing algorithm for undirected APSP

---

**Input:** An undirected graph  $G$ , parameter  $\varepsilon$ , approximate APSP oracle  $\mathcal{O}$ ,  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}(\cdot, \cdot)$ .

**Oracle:**  $\mathcal{O}(G, \varepsilon_0)$  returns  $(1 + \varepsilon_0)$ -approximate APSP distances  $\hat{d} : V(G) \times V(G) \rightarrow \mathbb{R}_{\geq 0}$  in  $G$ .

**Output:**  $((1 + 2\varepsilon)\alpha, \delta/2)$ -smooth distance estimate  $\tilde{d}'(\cdot, \cdot)$

```

1: Define  $H_1, H_2, H_3$  as in Definition B.3 for  $\omega = 100\delta/\varepsilon$  and  $\ell(\sigma_s, u) = (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w$ 
2: for  $i \leftarrow 1, 2, 3$  do
3:    $\hat{d}_i(\cdot, \cdot) \leftarrow \mathcal{O}(H_i, \varepsilon/10^4)$ 
4:   for all  $s \in V(G)$  do
5:     for all  $v \in V(G)$  do
6:       if  $0 \leq \{\tilde{d}(s, w)\}_i < \omega/2$  then
7:          $\tilde{d}_i(s, v) \leftarrow \lfloor \tilde{d}(s, u) \rfloor_i + \hat{d}_i(\sigma_s, v) - \omega$ 
8:       else
9:          $\tilde{d}_i(s, v) \leftarrow +\infty$ 
10: return  $\tilde{d}'$  defined by  $\tilde{d}'(s, u) = \min(\tilde{d}(s, u), \tilde{d}_1(s, u), \tilde{d}_2(s, u), \tilde{d}_3(s, u))$ 

```

---

We will now analyse Algorithm 8. The analysis is very similar to the directed SSSP case from Section 5.2 but we need to argue a bit more carefully about interactions of different levels of the level graph  $H_i$ .

Analogously to Section 5.2, for parameters  $\varepsilon, \alpha$  and a node  $s \in G$  we define a path  $P \subseteq G$  *smoothing* whenever

$$\tilde{d}(s, v) \geq \tilde{d}(s, u) + \alpha(1 + \varepsilon/2)\ell_G(P) \tag{B.1}$$

and use its following properties.

**Claim B.4.** *Analogue of Claim 5.9 Any smoothing path  $P$  satisfies  $\ell_G(P) \leq 2\delta/(\alpha\varepsilon)$ .*

*Proof.* The proof is the same as in Claim 5.9: We have by definition that a smoothing path  $P$  from  $u$  to  $v$  satisfies

$$\alpha(1 + \varepsilon/2)\ell_G(P) \leq \tilde{d}(s, v) - \tilde{d}(s, u).$$

By assumption that  $\tilde{d}(s, \cdot)$  is  $(\alpha, \delta)$ -smooth, we have

$$\tilde{d}(s, v) - \tilde{d}(s, u) \leq \alpha \cdot d_G(u, v) + \delta \leq \alpha \cdot \ell_G(P) + \delta.$$

Combining these gives the desired inequality. □

**Claim B.5.** *Analogue of Claim 5.10 Let  $P$  be a smoothing path. Let  $w_{\max} = \operatorname{argmax}_{w \in P} \tilde{d}(w)$  and  $w_{\min} = \operatorname{argmin}_{w \in P} \tilde{d}(w)$ . Then  $\tilde{d}(w_{\max}) - \tilde{d}(w_{\min}) \leq 6\delta/\varepsilon$ .*

*Proof.* We have

$$\begin{aligned}
\tilde{d}(w_{\max}) - \tilde{d}(w_{\min}) &\leq \tilde{d}(w_{\max}) - \tilde{d}(w_{\min}) + \tilde{d}(v) - \tilde{d}(u) \\
&= \left( \tilde{d}(w_{\max}) - \tilde{d}(u) \right) + \left( \tilde{d}(v) - \tilde{d}(w_{\min}) \right) \\
&\leq \alpha \ell_G(P) + \delta + \alpha \ell_G(P) + \delta \\
&\leq 2(2\varepsilon^{-1} + 1)\delta \leq 6\delta/\varepsilon.
\end{aligned}$$

where we first used the  $(\alpha, \delta)$ -smoothness assumption on  $(u, w_{\max})$  and  $(w_{\min}, v)$  and then we applied Claim B.4.  $\square$

The fact that smoothing paths cannot be long then implies that both endpoints of any smoothing path are both in the “bottom half” of at least one layer of  $H_1, H_2$  or  $H_3$ . The following claim is the place where we need three graphs instead of two as in Section 5.2.

**Claim B.6.** *Let  $P$  be any smoothing path from  $u$  to  $v$ . Then, there exists  $i \in \{1, 2, 3\}$  such that*

$$0 \leq \{\tilde{d}(s, u)\}_i \leq \{\tilde{d}(s, v)\}_i < \omega/2.$$

*Proof.* Choose  $i$  to be such that  $0 \leq \{\tilde{d}(s, u)\}_i < \omega/3$ . Then, the  $(\alpha, \delta)$ -smoothness of  $\tilde{d}$  and Claim B.4 yield  $\tilde{d}(s, v) \leq \tilde{d}(s, u) + \alpha \ell_G(P) + \delta < \omega/3 + \alpha(2\delta/(\alpha\varepsilon)) + \delta \leq \omega/3 + 3\delta/\varepsilon$ . We use the definition of  $\omega = 100\delta/\varepsilon$  to get that  $3\delta/\varepsilon < \omega/6$  and thus  $\{\tilde{d}(s, v)\}_i < \omega/2$  as needed.  $\square$

The main difference to the proof of Theorem 5.1 is that we need to argue more carefully about shortest paths in  $H_i$  behaving as in Figure 8.

**Claim B.7.** *Structure of shortest paths in  $H_i$  Given  $s, v \in V(G)$ , let  $P \subseteq H_i$  be either*

1. *a shortest path from  $\sigma_s$  to some  $v \in V(G)$  in  $H_i$ .*
2. *if  $\widehat{d}(\sigma_s, \cdot)$  is tree-like,  $P$  may also be the path that we get by starting at  $v$  and each time going from current node  $u$  to a node  $w$  with  $\widehat{d}_i(s, w) \leq \widehat{d}_i(s, u) - \ell_{H_i}(w, u)$  until we reach  $\sigma_s$ .*

*Let  $\{\sigma_s, u'\}$  be the first edge of  $P$  and let us write  $P = \{\{\sigma_s, u'\}\} \cup P'$ . Then*

1.  $P' \subseteq G$ .
2.  $\lfloor \tilde{d}(s, u') \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$ .
3.  $P'$  is a smoothing path from  $\sigma_s$  to  $v$ .

*Proof.* Assume that  $P$  is the shortest path from  $\sigma_s$  to  $v$ , we discuss the other possibility from the claim statement at the end. We first note that  $P' \subseteq G$ . Otherwise, the length of  $P$  is at least  $3\omega$  as it has to contain at least 3 edges from  $V(G) \times (V(H_i) \setminus V(G))$ , while the edge from  $\sigma_s$  to  $v$  has length only at most  $2\omega$ , contradicting that  $P$  is a shortest path.

Next, we prove that  $\lfloor \tilde{d}(s, u') \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$ . To that end, first assume that  $\lfloor \tilde{d}(s, u') \rfloor_i < \lfloor \tilde{d}(s, v) \rfloor_i$ . In this case, we have on one hand that

$$\ell_{H_i}(P) \leq \ell_{H_i}(\sigma_s, v) = (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w$$

because there is an edge between  $\sigma_s$  and  $v$ . But on the other hand we have

$$\begin{aligned}\ell_{H_i}(P) &= \ell_{H_i}(\sigma_s, u') + \ell_{H_i}(P') \\ &\geq (1 + \varepsilon/10^3)\{\tilde{d}(s, u')\}_i + w + \alpha(1 + \varepsilon)\frac{\tilde{d}(s, v) - \tilde{d}(s, u') - \delta}{\alpha}\end{aligned}$$

where we used  $(\alpha, \delta)$ -smoothness of  $u', v$ . If we put these two bounds together, we get

$$(1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i \geq (1 + \varepsilon/10^3)\{\tilde{d}(s, u')\}_i + (1 + \varepsilon)(\tilde{d}(s, v) - \tilde{d}(s, u') - \delta)$$

which implies that

$$\{\tilde{d}(s, u')\}_i + (1 + \varepsilon)(\lfloor \tilde{d}(s, v) \rfloor_i - \tilde{d}(s, u')) - 2\delta \leq 0$$

and we can simplify it further to

$$\varepsilon(\lfloor \tilde{d}(s, v) \rfloor_i - \lfloor \tilde{d}(s, u') \rfloor_i) \leq 2\delta$$

Finally, our assumption on  $u'$  then implies  $\varepsilon\omega \leq 2\delta$ , a contradiction with the definition of  $\omega$ .

Next, assume that  $\lfloor \tilde{d}(s, u') \rfloor_i > \lfloor \tilde{d}(s, v) \rfloor_i$ . We again derive a contradiction. As a small aside, we note that the following argument is the place where we rely on Line 9 in Algorithm 8, i.e., we set  $\tilde{d}_i(s, v) \leq +\infty$  only for  $0 \leq \{\tilde{d}(s, v)\}_i < \omega/2$ . Also, this is a place where we crucially use the fact  $G$  is undirected.

As in the previous case we start by writing

$$\ell_{H_i}(P) \leq \ell_{H_i}(\sigma_s, v) = (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w$$

and similarly to the previous case we write

$$\begin{aligned}\ell_{H_i}(P) &= \ell_{H_i}(\sigma_s, u') + \ell_{H_i}(P') \\ &\geq (1 + \varepsilon/10^3)\{\tilde{d}(s, u')\}_i + w + \alpha(1 + \varepsilon)\frac{\tilde{d}(s, u') - \tilde{d}(s, v) - \delta}{\alpha}.\end{aligned}$$

Notice that for the second bound we use the fact that as  $G$  is undirected,  $P'$  is also the shortest path from  $v$  to  $u'$ . We again put the two bounds together and simplify, we get

$$(1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i \geq (1 + \varepsilon)(\tilde{d}(s, u') - \tilde{d}(s, v) - \delta)$$

Note that  $\{\tilde{d}(s, v)\}_i < \omega/2$  together with our assumption on  $u'$  implies  $\tilde{d}(s, u') - \tilde{d}(s, v) \geq \omega/2$ , hence we can simplify to

$$(1 + \varepsilon/10^3)\omega/2 \geq (1 + \varepsilon)(\omega/2 - \delta)$$

which is, however, in contradiction with our definition  $\omega = 100\delta/\varepsilon$ . This finishes the proof that  $\lfloor \tilde{d}(s, u') \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$  (a second bullet point in the claim statement).

We finish by showing that  $P'$  is a smoothing path. We again start by writing

$$\ell_{H_i}(P) \leq (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w$$

which in turn implies

$$\ell_{H_i}(P') \leq (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i - (1 + \varepsilon/10^3)\{\tilde{d}(s, u')\}$$

Using the fact that lengths in  $H_i$  are multiplied by the factor of  $\alpha(1 + \varepsilon)$ , we conclude that

$$\{\tilde{d}(s, v)\}_i \geq \{\tilde{d}(s, u')\}_i + \alpha \frac{1 + \varepsilon}{1 + \varepsilon/10^3} \ell_{H_i}(P')$$

which implies the smoothness of  $P'$  as needed.

Finally, consider  $P$  to be not the shortest path but the path defined in the second bullet point of the claim statement. The proof for this case is very similar and we thus omit it. We note that in calculations bounding  $\ell_{H_i}(P)$ , we need to replace the bound

$$\ell_{H_i}(P) \leq (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w$$

by

$$\ell_{H_i}(P) \leq (1 + \varepsilon/10^4) \left( (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + w \right).$$

This leads to a small additional error term of size at most  $\frac{\varepsilon}{10^4} \cdot 3w$  that does not affect the conclusion of above argument.  $\square$

We now follow the single source case and prove an analogue of Proposition 5.11.

**Proposition B.8.** *The output function  $\tilde{d}'(s, \cdot)$  in Algorithm 8 satisfies for any  $s$  that*

$$\min_{u \in V(G)} \left( \tilde{d}(s, u) + \alpha(1 + \varepsilon)d_G(u, v) \right) \leq \tilde{d}'(s, v) \leq \min_{u \in V(G)} \left( \tilde{d}(s, u) + \alpha(1 + \varepsilon)d_G(u, v) \right) + \delta/2, \quad (\text{B.2})$$

*Proof.* Recall we define  $\tilde{d}'(s, v) = \min \left( \tilde{d}(s, v), \tilde{d}_1(s, v), \tilde{d}_2(s, v), \tilde{d}_3(s, v) \right)$ . We start with the first inequality, i.e., we need to prove that the left hand side of (B.2), that we denote as LHS, is upper bounded by  $\tilde{d}(s, v), \tilde{d}_1(s, v), \tilde{d}_2(s, v), \tilde{d}_3(s, v)$ . We certainly have  $\text{LHS} \leq \tilde{d}(s, v)$  as  $v$  is one of the nodes we minimize over in LHS. Hence we need to prove that it is also upper bounded by  $\tilde{d}_i(s, v)$  for each  $i \in \{1, 2, 3\}$ . We recall that whenever  $\tilde{d}_i(s, v)$  is not  $+\infty$  (in which case there is nothing to prove), it is defined as:

$$\tilde{d}_i(s, v) = \lfloor \tilde{d}(s, v) \rfloor_i + \hat{d}_i(\sigma_s, v) - \omega \quad (\text{B.3})$$

$$\geq \lfloor \tilde{d}(s, v) \rfloor_i + d_{H_i}(\sigma_s, v) - \omega \quad (\text{B.4})$$

Recall that by Claim B.7 any shortest path  $P$  from  $\sigma_s$  to  $v$  consists of an edge from  $\sigma_s$  to some  $u'$  with  $\lfloor \tilde{d}(s, u') \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$  and then a path  $P' = P \setminus \{\{\sigma_s, u'\}\} \subseteq G$ . This allows us to rewrite the above inequality as

$$\tilde{d}_i(s, v) \geq \lfloor \tilde{d}(s, u') \rfloor_i + ((1 + \varepsilon/10^3)\{\tilde{d}(s, u')\}_i + \omega) + d_{H_i}(u', v) - \omega \quad (\text{B.5})$$

$$\geq \tilde{d}(s, u') + \alpha(1 + \varepsilon)d_G(u', v) \geq \text{LHS} \quad (\text{B.6})$$

as needed.

We continue with the second inequality in (B.2) whose right hand side we denote as RHS. Let  $u'$  be the minimizer of RHS. The node  $u'$  has to satisfy that  $\tilde{d}(s, u') + \alpha(1 + \varepsilon)d_G(u', v) \leq \tilde{d}(s, v)$

as we can plug in  $v$  to the minimization in RHS. Hence, the shortest path from  $u'$  to  $v$  in  $G$  is smoothing according to (B.1). Applying Claim B.6, we get existence of  $i$  such that  $0 \leq \{\tilde{d}(s, u)\}_i \leq \{\tilde{d}(s, v)\}_i < \omega/2$ . For this  $i$ , we have  $\tilde{d}(s, v)_i < +\infty$  and thus we can write

$$\tilde{d}'(s, v) \leq \tilde{d}(s, v)_i = \lfloor \tilde{d}(s, v) \rfloor_i + \widehat{d}_i(\sigma_s, v) - \omega \quad (\text{B.7})$$

We continue by using the fact that our distance oracle is  $(1+\varepsilon/10^4)$ -approximate to bound  $\widehat{d}_i(\sigma_s, v) \leq (1 + \varepsilon/10^4)d_{H_i}(\sigma_s, v)$ . Plugging this to above inequality and using Claim B.7 that says that the shortest path  $P$  from  $\sigma_s$  to  $v$  can be written as  $\{\{\sigma_s, u\}\} \cup P'$  for  $P' \subseteq G$ , we get

$$\tilde{d}'(s, v) \leq \lfloor \tilde{d}(s, v) \rfloor_i + (1 + \varepsilon/10^4) \cdot \left( \min_{u \in V(G)} ((1 + \varepsilon/10^3)\{\tilde{d}(s, u)\}_i + \omega) + \alpha(1 + \varepsilon)d_G(u, v) \right) - \omega \quad (\text{B.8})$$

We need to simplify the right-hand side of this inequality. First, we get rid of the inner term  $(1 + \varepsilon/10^3)$  by simplifying the inequality as

$$\tilde{d}'(s, v) \leq \lfloor \tilde{d}(s, v) \rfloor_i + (1 + \varepsilon/10^4) \cdot \left( \min_{u \in V(G)} (\{\tilde{d}(s, u)\}_i + \omega) + \alpha(1 + \varepsilon)d_G(u, v) \right) - \omega + \frac{2\varepsilon}{1000}\omega \quad (\text{B.9})$$

Next, let  $u'$  be a minimizer of the expression on the right-hand side. Claim B.7 implies that  $\lfloor \tilde{d}(s, u') \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$ . Hence, we can rewrite the right hand side of (B.8) by isolating the term proportional to  $\varepsilon/10^4$ , to get

$$\text{RHS of (B.8)} \quad (\text{B.10})$$

$$= \min_{u \in V(G)} \left( \tilde{d}(s, u) + \alpha(1 + \varepsilon)d_G(u, v) \right) + \frac{\varepsilon}{10^4} \cdot \min_{u \in V(G)} \left( (\{\tilde{d}(s, u)\}_i + \omega) + \alpha(1 + \varepsilon)d_G(u, v) \right) + \frac{2\varepsilon}{1000}\omega \quad (\text{B.11})$$

Finally, we bound the term proportional to  $\varepsilon/10^4$  by using that

$$\min_{u \in V(G)} \left( (\{\tilde{d}(s, u)\}_i + \omega) + \alpha(1 + \varepsilon)d_G(u, v) \right) \leq \{\tilde{d}(s, v)\}_i + \omega \leq 2\omega \quad (\text{B.12})$$

where we used the fact that the minimization includes  $v$ . Putting (B.8), (B.10), and (B.12) together, we conclude that

$$\tilde{d}'(s, v) \leq \min_{u \in V(G)} \left( \tilde{d}(s, u) + \alpha(1 + \varepsilon)d_G(u, v) \right) + \frac{\varepsilon}{10^4} \cdot 2\omega + \frac{2\varepsilon}{1000}\omega. \quad (\text{B.13})$$

Finally, the fact  $\frac{\varepsilon}{10^4} \cdot 2\omega + \frac{2\varepsilon}{1000}\omega < \delta/2$  concludes the proof. □

**Proposition B.9.** *If Algorithm 8 is called with an  $(\alpha, \delta)$ -smooth distance estimate  $\tilde{d}$ , then the output  $\tilde{d}'$  is  $(\alpha(1 + \varepsilon), \delta/2)$ -smooth.*

*Proof.* The proof for each  $s$  is very similar to the proof of Proposition 5.12, we rewrite it here for completeness.



Let  $v, w$  be any vertices in  $G$ . Letting  $u := \operatorname{argmin}_{u' \in V(G)} \tilde{d}(u') + \alpha(1 + \varepsilon)d_G(u', v)$  it follows by Proposition B.8 that

$$\tilde{d}(u) + \alpha(1 + \varepsilon)d_G(u, v) \leq \tilde{d}'(v),$$

and

$$\tilde{d}'(w) \leq \tilde{d}(u) + \alpha(1 + \varepsilon)d_G(u, w) + \delta/2.$$

Combining these, we get that

$$\begin{aligned} \tilde{d}'(w) - \tilde{d}'(v) &\leq \alpha(1 + \varepsilon)d_G(u, w) + \delta/2 - \alpha(1 + \varepsilon)d_G(u, v) \\ &\leq \alpha(1 + \varepsilon)(d_G(u, v) + d_G(v, w)) + \delta/2 - \alpha(1 + \varepsilon)d_G(u, v) \\ &= \alpha(1 + \varepsilon)d_G(v, w) + \delta/2. \end{aligned}$$

as desired. □

**Proposition B.10.** *The function  $\tilde{d}'(s, \cdot)$  is tree-like for any  $s$ , provided that the approximate shortest path oracle  $\mathcal{O}$  and the distance function  $\tilde{d}$  are tree-like.*

*Proof.* Fix  $s$ . Let  $v \neq s$  be an arbitrary node; recall that we need to prove existence of  $w_0$  with

$$\tilde{d}'(s, w_0) \leq \tilde{d}'(s, v) - \ell_G(w_0, v) \tag{B.14}$$

Also, recall that  $\tilde{d}'(s, v) = \min(\tilde{d}(s, v), \tilde{d}_1(s, v), \tilde{d}_2(s, v), \tilde{d}_3(s, v))$ ; if  $\tilde{d}'(s, v) = \tilde{d}(s, v)$ , (B.14) is implied by the assumption that  $\tilde{d}(s, \cdot)$  is tree-like and the fact that  $\tilde{d}'(s, w) \leq \tilde{d}(s, w)$ . Next, assume  $\tilde{d}'(s, v) < \tilde{d}(s, v)$  and let  $i$  be such that  $\tilde{d}'(s, v) = \tilde{d}_i(s, v)$ ; in particular we have  $\tilde{d}_i(s, v) < +\infty$  in this case. The tree-likeness of our oracle implies that there exists  $w$  such that

$$\hat{d}_i(s, w) \leq \hat{d}_i(s, v) - \ell_{H_i}(v, w). \tag{B.15}$$

We will next go through all possible cases of where the node  $w$  possibly lies and argue that we can always set  $w_0 = w$  in (B.14).

Let us define a path  $P$  as in Claim B.7: we start with the node  $v$  and each time go from the current node  $u$  to its neighbor  $u'$  with  $\hat{d}_i(s, u') \leq \hat{d}_i(s, u) - \ell_{H_i}(u', u)$  until we reach  $\sigma_s$ . We also write  $P = \{\{\sigma_s, u'\}\} \cup P'$  for  $P' \subseteq G$ .

Let us start with the case when  $P' = \{\}$ , i.e.,  $u' = w$ . In that case, we use (B.15) to observe that  $\hat{d}_i(\sigma_s, v) \geq \ell_{H_i}(\sigma_s, v) = (1 + \varepsilon/10^3)\{\tilde{d}(s, v)\}_i + \omega$ . When we plug this in the definition  $\tilde{d}_i(s, v)$ , we get

$$\tilde{d}_i(s, v) = \lfloor \tilde{d}(s, v) \rfloor_i + \hat{d}_i(\sigma_s, v) - \omega \tag{B.16}$$

$$= \tilde{d}(s, v) + \varepsilon\{\tilde{d}(s, v)\}_i/100 \tag{B.17}$$

$$\geq \tilde{d}(s, v), \tag{B.18}$$

a contradiction with our assumption that  $\tilde{d}_i(s, v) = \tilde{d}'(s, v) < \tilde{d}(s, v)$ .

Hence, in the rest of the proof we consider the case when  $w \in P' \subseteq V(G)$ . Here we need to consider several subcases. First, assume that  $\lfloor \tilde{d}(s, w) \rfloor_i < \lfloor \tilde{d}(s, v) \rfloor_i$ . In this case we use (B.15) to conclude that

$$\widehat{d}_i(s, v) \geq \ell_{H_i}(v, w) + \widehat{d}_i(s, w) \quad (\text{B.19})$$

$$\geq \ell_{H_i}(v, w) + d_{H_i}(\sigma_s, w) \quad (\text{B.20})$$

$$\geq \ell_{H_i}(v, w) + \omega \quad (\text{B.21})$$

where we used the tree-likeness of  $\widehat{d}$ , the fact  $\widehat{d}$  is proper, and that any edge from  $\sigma_s$  to  $V(G)$  has length at least  $\omega$ . Hence

$$\tilde{d}_i(s, v) = \lfloor \tilde{d}(s, v) \rfloor_i + \widehat{d}_i(\sigma_s, v) - \omega \quad (\text{B.22})$$

$$\geq \lfloor \tilde{d}(s, v) \rfloor_i + \ell_{H_i}(v, w) \quad (\text{B.23})$$

$$(\text{B.24})$$

and as our assumption on  $w$  implies  $\tilde{d}(s, w) < \lfloor \tilde{d}(s, v) \rfloor_i$ , we conclude that

$$\tilde{d}_i(s, v) - \tilde{d}_i(s, w) \geq \lfloor \tilde{d}(s, v) \rfloor_i + \ell_{H_i}(v, w) - \tilde{d}_i(s, w) \geq \ell_{H_i}(v, w), \quad (\text{B.25})$$

i.e., we can choose  $w_0 = w$  in (B.14) as desired.

Next, consider the case when  $\lfloor \tilde{d}(s, w) \rfloor_i > \lfloor \tilde{d}(s, v) \rfloor_i$ . This case is in direct contradiction with Claim B.5: If we define  $w_{\max} = \operatorname{argmax}_{w \in P} \tilde{d}(w)$  and  $w_{\min} = \operatorname{argmin}_{w \in P} \tilde{d}(w)$ , we would have  $\tilde{d}(s, w_{\max}) \geq \tilde{d}(s, w) \geq \lfloor \tilde{d}(s, w) \rfloor_i$  and  $\tilde{d}(s, w_{\min}) \leq \tilde{d}(s, v) \leq \lfloor \tilde{d}(s, v) \rfloor_i + \omega/2$  which implies  $\tilde{d}(s, w_{\max}) - \tilde{d}(s, w_{\min}) \geq \omega/2$  contradicting the bound from Claim B.5.

So, we have  $\lfloor \tilde{d}(s, w) \rfloor_i = \lfloor \tilde{d}(s, v) \rfloor_i$ . As a final case distinction, we consider the cases when  $0 \leq \{ \tilde{d}(s, w) \}_i < \omega/2$  (and hence  $\tilde{d}(s, w)_i < +\infty$ ) and when  $\omega/2 \leq \{ \tilde{d}(s, w) \}_i < \omega$  (and hence  $\tilde{d}(s, w)_i = +\infty$ ).

In the first case the tree-likeness follows directly from definition:

$$\begin{aligned} \tilde{d}_i(s, v) - \tilde{d}_i(s, w) &= \left( \lfloor \tilde{d}(s, v) \rfloor_i + \widehat{d}(s, v) - \omega \right) - \left( \lfloor \tilde{d}(s, w) \rfloor_i + \widehat{d}(s, w) - \omega \right) \\ &= \widehat{d}(s, v) - \widehat{d}(s, w) \\ &\geq \ell_{H_i}(v, w) \geq \ell_G(v, w) \end{aligned}$$

It remains to consider the case when  $\omega/2 \leq \{ \tilde{d}(s, w) \}_i < \omega$ . We will show that this case cannot happen since for  $j = (i + 1) \bmod 3$  we have  $\tilde{d}_j(s, v) < \tilde{d}_i(s, v)$  which contradicts our choice of  $i$  as the index for which  $\tilde{d}'(s, v) = \tilde{d}_i(s, v)$ . The intuitive reason for this is that the ‘‘slow down’’  $(1 + \varepsilon/10^3)$  of edges from  $V(H_i) \setminus V(G)$  to  $V(G)$  affects more the graph  $H_i$  than  $H_j$ . For the formal proof we first observe that  $\{ \tilde{d}(s, w) \}_i \geq \omega/2$  implies via Claim B.5 that  $\{ \tilde{d}(s, v) \}_i \geq \omega/2 - 6\delta/\varepsilon$ . This in turn implies that  $\{ \tilde{d}(s, v) \}_j \geq \omega/2 - 6\delta/\varepsilon - \omega/3 \geq 0$  and thus  $\tilde{d}'(s, v) < +\infty$ . In particular, we can write

$$\tilde{d}_i(s, v) - \tilde{d}_j(s, v) = \left( \lfloor \tilde{d}(s, v) \rfloor_i + \widehat{d}_i(\sigma_s, v) - \omega \right) - \left( \lfloor \tilde{d}(s, v) \rfloor_j + \widehat{d}_j(\sigma_s, v) - \omega \right) \quad (\text{B.26})$$

$$= -\omega/3 + \widehat{d}_i(\sigma_s, v) - \widehat{d}_j(\sigma_s, v) \quad (\text{B.27})$$

The value  $\widehat{d}_i(\sigma_s, v)$  is bounded as

$$\widehat{d}_i(\sigma_s, v) \geq \ell_{H_i}(P) = \left(1 + \frac{\varepsilon}{10^3}\right) \{\widetilde{d}(s, u')\}_i + \omega + \alpha(1 + \varepsilon)\ell_G(P')$$

Similarly, we bound the value  $\widehat{d}_j(\sigma_s, v)$  by

$$\widehat{d}_j(\sigma_s, v) \leq \left(1 + \frac{\varepsilon}{10^4}\right) \ell_{H_j}(P) = \left(1 + \frac{\varepsilon}{10^4}\right) \left( \left(1 + \frac{\varepsilon}{10^3}\right) \{\widetilde{d}(s, u')\}_j + \omega + \alpha(1 + \varepsilon)\ell_G(P') \right)$$

Note that we have  $\{\widetilde{d}(s, u')\}_i \geq \omega/2 - 6\delta/\varepsilon$  by Claim B.5, hence  $\{\widetilde{d}(s, u')\}_j \geq 0$  and, in particular,  $\{\widetilde{d}(s, u)\}_j = \{\widetilde{d}(s, u)\}_i - \omega/3$ . This allows us to simplify (B.26) as

$$\widetilde{d}_i(s, v) - \widetilde{d}_j(s, v) = -\omega/3 + \left(1 + \frac{\varepsilon}{10^3}\right) \omega/3 - \frac{\varepsilon}{10^4} \left( \left(1 + \frac{\varepsilon}{10^3}\right) \{\widetilde{d}(s, u')\}_j + \omega + \alpha(1 + \varepsilon)\ell_G(P') \right) \quad (\text{B.28})$$

We now only need to bound the right hand side error term as

$$\left(1 + \frac{\varepsilon}{10^3}\right) \{\widetilde{d}(s, u')\}_j + \omega + \alpha(1 + \varepsilon)\ell_G(P') \leq (1 + \varepsilon/10^3)\omega + \omega + \alpha(1 + \varepsilon) \cdot 2\delta/(\alpha\varepsilon) \leq 3\omega$$

and we conclude that  $\widetilde{d}_i(s, v) - \widetilde{d}_j(s, v) > 0$ , a contradiction with our choice of  $i$ . □

## C Applications in Different Computational Models

In this section, we formally prove our main results in the PRAM and CONGEST model. Some of the results are slightly more general compared to their counterparts stated in the introduction. The results more or less straightforwardly follow by combining the reductions given in this paper together with the state-of-the-art parallel and distributed algorithms for computing approximate distance estimates. There are however some low-level technicalities that have to be dealt with and which we now discuss in more detail.

### C.1 Directed Graphs

Algorithm 1 reduces the computation of exact distances in a graph with polynomially bounded integer edge weights to computing smoothly 2-approximate distance estimates in graphs with nonnegative real weights. Also, Algorithm 3 reduces the computation of smoothly 2-approximate distance estimates in graphs with real edge weights in  $[1, \text{poly}(n)]$  to the computation of  $(1 + O(1/\log n))$ -approximate distance estimates.

We would now like to combine these two reductions. That is, answer each oracle call in Algorithm 1 by using Algorithm 3. The issue that has to be dealt with is that Algorithm 1 invokes the smoothly 2-approximate distance oracle on graphs with polynomially bounded maximum edge length, however, the minimum edge weight can be much smaller than one. In fact, the aspect ratio, i.e. the largest edge weight divided by the smallest edge weight, might not be polynomially bounded. However, this is merely a technicality. One can easily adapt Algorithm 1 to ensure that it only invokes the distance oracle on graphs with polynomially bounded aspect ratio by adding  $1/\text{poly}(n)$  to each edge when invoking the distance oracle, which ensure that the aspect ratio is polynomially bounded. Then, by scaling the weights one can additionally ensure that it only invokes the distance oracle on

graphs with real edge weights in  $[1, \text{poly}(n)]$ . By performing these minor modifications, we obtain a reduction from computing exact distances in graphs with nonnegative polynomially bounded integer weights to  $O(\log n)$  computations of  $(1 + O(1/\log n))$ -approximate distance estimates in graphs with nonnegative real weights.

We would now like to compute the  $(1 + O(1/\log n))$ -approximate distance estimates with the state-of-the-art parallel and distributed algorithms. However, they only work in graphs with nonnegative polynomially bounded integer edge lengths. One can slightly modify Algorithm 1 to ensure that it only invokes the approximate distance oracle on such graphs. First, we can ensure that it only invokes the approximate distance estimates on graphs with polynomial aspect ratio by again adding an additive  $1/\text{poly}(n)$  to each edge. Then, by discretizing and scaling the edge weights, we can ensure that the approximate distance oracle is only invoked on graphs with nonnegative polynomially bounded integer edge weights. By implementing these minor modifications, we therefore obtain a reduction from computing exact distances in directed graphs with nonnegative polynomially bounded integer edge weights to  $O(\log n)$  computations of  $(1 + O(1/\log n))$ -approximate distance estimates on directed graphs with nonnegative polynomially bounded integer edge weights. Using the approximate distance algorithm of Cao, Fineman and Russell [CFR20a] and observing that all of the reductions can be efficiently implemented in the PRAM model, we therefore obtain the following result in the PRAM model.

**Corollary 1.8.** *[Directed SSSP, PRAM] There exists a parallel algorithm that takes as input any directed graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\rho \in [1, n^{1/2}]$ . The algorithm computes for every vertex the exact distance to  $s$  and an exact shortest path tree rooted at  $s$ . The algorithm is randomized, requires  $\tilde{O}(m\rho^2 + n\rho^4)$  work and  $n^{1/2+o(1)}/\rho$  depth, and works with high probability.*

By observing that the reductions can also efficiently implemented in the CONGEST model, and that the state-of-the-art CONGEST algorithm for computing approximate distances in directed graphs also works if the source node is a virtual node, we obtain the following result in the CONGEST model.

**Corollary 1.9.** *[Directed SSSP, CONGEST] There exists a distributed algorithm that takes as input any directed graph  $G$  with nonnegative polynomially bounded integer edge lengths and any source vertex  $s \in V(G)$ . The algorithm computes for every vertex the exact distance to  $s$  and an exact shortest path tree rooted at  $s$ . The algorithm is randomized, runs in  $\tilde{O}(n^{1/2} + D + n^{2/5+o(1)}D^{2/5})$  rounds where  $D$  denotes the undirected hop-diameter of  $G$ , and works with high probability.*

## C.2 Undirected Graphs

We already have argued that we can modify Algorithm 3 to compute smoothly approximate distance estimates in graphs with polynomially bounded integer edge weights such that it only relies on computing approximate distance estimates in graphs with positive polynomially bounded integer edge weights. We can additionally modify it in such a way that it still preserves tree-likeness. We can therefore answer each oracle call of Algorithm 3 using Algorithm 7. Moreover, we can also slightly modify Algorithm 7, along the lines we modified the other reductions, such that it only relies on approximate distance estimates on graphs equipped with positive polynomially bounded integer edge weights. By doing these modifications, we obtain a reduction for computing strongly  $(1 + \varepsilon)$ -approximate distance estimates in graphs with nonnegative polynomially bounded integer edge weights to computing  $(1 + \varepsilon)$ -approximate distance estimates in graphs with nonnegative polynomially bounded integer weights.

Using the state-of-the-art deterministic approximate distance algorithm for undirected graphs and observing that all of the reductions can be efficiently implemented in the PRAM model, we therefore obtain the following result in the PRAM model.

**Corollary 1.12.** *[Undirected Strong Distance Estimates, PRAM] There exists a parallel algorithm that takes as input any undirected graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\varepsilon \in (0, 1]$ . The algorithm computes a strongly  $(1 + \varepsilon)$ -approximate distance estimate from  $s$ . The algorithm is deterministic, requires  $m \cdot \left(\frac{\log n}{\varepsilon}\right)^{O(1)}$  work and depth  $\left(\frac{\log n}{\varepsilon}\right)^{O(1)}$ .*

By observing that the reductions can also efficiently implemented in the CONGEST model, and that the state-of-the-art CONGEST algorithm for computing approximate distances in undirected graphs also works if the source node is a virtual node, we obtain the following result in the CONGEST model.

**Corollary 1.13.** *[Directed Strong Distance Estimates, CONGEST] There exists a distributed algorithm that takes as input any undirected graph  $G$  with nonnegative polynomially bounded integer edge lengths, any source vertex  $s \in V(G)$  and a parameter  $\varepsilon \in (0, 1]$ . The algorithm computes a strongly  $(1 + \varepsilon)$ -approximate distance estimate from  $s$ . The algorithm is deterministic and runs in  $(\sqrt{n} + D) \left(\frac{\log n}{\varepsilon}\right)^{O(1)}$  rounds.*