

Improved Distributed Algorithms for Exact Shortest Paths

Mohsen Ghaffari
ETH Zurich
ghaffari@inf.ethz.ch

Jason Li
CMU
jmli@cs.cmu.edu

Abstract

Computing shortest paths is one of the central problems in the theory of distributed computing. For the last few years, substantial progress has been made on the *approximate* single source shortest paths problem, culminating in an algorithm of Becker et al. [DISC'17] which deterministically computes $(1 + o(1))$ -approximate shortest paths in $\tilde{O}(D + \sqrt{n})$ time, where D is the hop-diameter of the graph. Up to logarithmic factors, this time complexity is optimal, matching the lower bound of Elkin [STOC'04].

The question of *exact* shortest paths however saw no algorithmic progress for decades, until the recent breakthrough of Elkin [STOC'17], which established a sublinear-time algorithm for exact single source shortest paths on undirected graphs. Shortly after, Huang et al. [FOCS'17] provided improved algorithms for exact all pairs shortest paths problem on directed graphs.

In this paper, we present a new single-source shortest path algorithm with complexity $\tilde{O}(n^{3/4}D^{1/4})$. For polylogarithmic D , this improves on Elkin's $\tilde{O}(n^{5/6})$ bound and gets closer to the $\tilde{\Omega}(n^{1/2})$ lower bound of Elkin [STOC'04]. For larger values of D , we present an improved variant of our algorithm which achieves complexity $\tilde{O}(n^{3/4+o(1)} + \min\{n^{3/4}D^{1/6}, n^{6/7}\} + D)$, and thus compares favorably with Elkin's bound of $\tilde{O}(n^{5/6} + n^{2/3}D^{1/3} + D)$ in essentially the entire range of parameters. This algorithm provides also a qualitative improvement, because it works for the more challenging case of directed graphs (i.e., graphs where the two directions of an edge can have different weights), constituting the first sublinear-time algorithm for directed graphs. Our algorithm also extends to the case of exact κ -source shortest paths, giving a complexity of $\tilde{O}(\min\{\kappa^{1/2}n^{3/4+o(1)} + \kappa^{1/3}n^{3/4}D^{1/6}, \kappa^{3/7}n^{6/7}\} + D)$. For moderately small κ and D , this improves on the $\tilde{O}(\kappa^{1/2}n^{3/4} + n)$ bound of Huang et al.

1 Introduction & Related Work

Computing shortest paths—either from a single node to all or from multiple or even all nodes to all—are among the most central problems in the theory of distributed computing, with consequential applications; e.g., they are a center-piece of routing in computer networks. As such, shortest path problems and distributed algorithms for them have been studied extensively since the 1950’s, starting with the algorithms of Bellman [Bel58] and Ford [FJ56]. In this paper, we present new algorithms that improve on the state of the art of this decades-old problem. Next, we first overview the background of the problem and the known results, and then we outline our contributions.

1.1 Background, Related Work, and State of the Art

Model We work with the standard synchronous message-passing model of distributed computing, called CONGEST [Pel00], where the network is abstracted as an undirected graph $G = (V, E)$ with one node (processor) at each vertex in V . Each edge has an integer weight in range $[1, \Lambda]$; we assume that the weights can be described in $\Theta(\log n)$ bits and thus $\Lambda \leq \text{poly}(n)$. In general, we allow the two directions of an edge to have asymmetric weights, i.e., the weight of the link $v \rightarrow u$ might be different than that of $u \rightarrow v$. At the beginning, each node v knows only its neighbors and the weights of the corresponding incident edges. The communication network is independent of the weights and is thus bi-directional. Per round, each node $u \in V$ can send an $O(\log n)$ -bit message to each node $v \in V$ iff $\{u, v\} \in E$. Moreover, each node can perform arbitrary computations given the information that it has at that point.

The SSSP Problem The *single-source shortest path (SSSP)* problem is to compute $d(s, t)$ for each $t \in V$, given a source vertex $s \in V$. In the distributed setting, each node $t \in V$ should learn its distance $d(s, t)$ from the source.

State of the Art For Exact SSSP The classic approach of Bellman-Ford [Bel58, FJ56] gives an $O(n)$ round algorithm for SSSP. This complexity is optimal in the worst case, in the sense that in graphs of hop-diameter $D = \Theta(n)$ —e.g., think of a cycle—one cannot do better. But this is an unfortunate and uninformative impossibility. As was argued by Garay, Kutten, and Peleg in their influential work [GKP93], the case of graphs with small diameter D is far more relevant in the real world, and far more interesting theoretically, and the above impossibility does not preclude faster algorithms for those cases. Since then, the case of network graphs with moderately low-diameter has become the focus point of the area of distributed graph algorithms, e.g., see the progress on problems such as minimum spanning tree [GKP93, KP95, DSHK⁺11, Elk04b, PR99, MH16], min-cut [MK13, NS14], etc. However, despite this change of focus, for a long time, there was no progress on the algorithmic side of exact SSSP. The only development was an influential work of Elkin [Elk04b], which proved that computing the SSSP requires $\Omega(n^{1/2})$ rounds, even on graphs with diameter as low as $D = O(\log n)$.

A Turn Towards Approximation Algorithms Due to the lack of progress on the algorithmic side of the exact SSSP, the area turned focus to the approximate version. It can be argued that this was sparked by the visionary article of Elkin [Elk04a] that emphasized approximations. Over the years, this led to a remarkable sequence of work [DSHK⁺11, LPS13, Nan14, HKN16, BKKL27], providing better and better distributed SSSP approximation algorithms. By now, we can say that we have a good understanding of the approximate version of the problem, having the deterministic algorithm of Becker et al. [BKKL27], which computes a $(1 + o(1))$ -approximation¹ of SSSP in $\tilde{O}(D + \sqrt{n})$ rounds, and whose complexity is known to be the best possible up to logarithmic

¹Their algorithm computes a $(1 + \epsilon)$ -approximate SSSP in $\tilde{O}(\epsilon^{-O(1)}(\sqrt{n} + D))$ rounds, so we can take, e.g., $\epsilon := 1/\log n$.

factors, due to a lower bound of Elkin [Elk04b]. More precisely, the latter exhibits graphs of diameter $D = O(\log n)$ in which no SSSP approximation algorithm—even for rather coarse approximations—can run in less than $\tilde{\Omega}(\sqrt{n})$ rounds.

Return of the Exact Algorithms Despite the numerous steps of progress that brought us to an almost complete understanding of the approximate case of SSSP, computing exact SSSP remained open until very recently. But then came the breakthrough of Elkin [Elk17], which provided the first sublinear-time algorithm for exact SSSP. More precisely, Elkin gave an algorithm that computes exact SSSP in $\tilde{O}(n^{5/6})$ rounds if $D = \tilde{O}(\sqrt{n})$, and in $\tilde{O}(D^{1/3}n^{2/3})$ rounds if $D = \tilde{\Omega}(\sqrt{n})$.

Shortly after, another significant step of progress was made on the closely related problem of All-Pairs Shortest Path (APSP), when Huang et al. [HNS17] gave an algorithm with complexity $O(n^{5/4})$ rounds for positive integer-weighted graphs². Similar to SSSP, also for APSP, the $(1+o(1))$ -approximate version of the problem is well-understood by now, having $\tilde{\Theta}(n)$ upper and lower bounds due to Nanongkai [Nan14]. Though, the exact version remained open, with the best previous upper bounds remaining at the trivial extreme of $O(m)$, where m denotes the number of the edges. We note that the result of Huang et al. is more general: for k sources, the complexity grows smoothly as $O(n^{3/4}k^{1/2} + n)$; though unfortunately this bound always remains $\Omega(n)$ even for very small values of k . Moreover, it is important to remark that the result of Huang et al. [HNS17] is more general in that it can handle asymmetric distances along an edge—i.e., where the length of the (s, t) edge may differ from that of (t, s) —while Elkin’s SSSP algorithm assumes the symmetry [Elk17].

1.2 Our Contribution

In this paper, we present algorithms that improve on this state of the art both quantitatively and qualitatively for integer-weighted graphs. We will soon discuss the quantitative aspect, i.e., the improvements in the bounds. The qualitative aspect is the fact that our algorithms extend to networks with asymmetric weights, i.e., when the two directions of an edge can have different weights. Although this asymmetric case is relevant for practical networks, it has gained a far more important motivation recently even when working on graphs with symmetric weights, due to the introduction of the *scaling framework* [HNS17] to distributed computing. The scaling framework massages the weights in a way that simplifies the problem considerably, modulo the potential of making the weights asymmetric. We will discuss this framework in the next section.

We next overview our round complexity improvements. Our basic result gives an SSSP algorithm with complexity $\tilde{O}(n^{3/4})$, for graphs with polylogarithmic diameter D . This improves on the $\tilde{O}(n^{5/6})$ bound of Elkin [Elk17], and gets closer to the $\Omega(n^{1/2})$ lower bound of Peleg and Rubinovich [PR99], which holds for graphs of diameter $D = O(\log n)$. More generally, the result is as follows:

Theorem 1.1. *There is a randomized distributed algorithm that computes the exact distances from the source s to all nodes $t \in V$, in $\tilde{O}(n^{3/4}D^{1/4})$ rounds, with high probability.*

For graphs of large diameter, we present improved versions of this algorithm which achieve complexity $\tilde{O}(n^{3/4+o(1)} + \min\{n^{3/4}D^{1/6}, n^{6/7}\} + D)$. This compares favorably with Elkin’s bound of $\tilde{O}(n^{5/6} + n^{2/3}D^{1/3} + D)$ in essentially the entire range of parameters. Due to the space limitations, these improvements are deferred to [Appendix C](#).

Theorem 1.2. *There is a randomized distributed algorithm that computes the exact distances from the source s to all nodes $t \in V$, in $\tilde{O}(n^{3/4+o(1)} + \min\{n^{3/4}D^{1/6}, n^{6/7}\} + D)$ rounds, with high probability.*

²Elkin pointed out shortly after the work of Huang et al. [HNS17], but independently from them, that the techniques in his work [Elk17] immediately lead to an exact APSP algorithm with complexity $\tilde{O}(n^{5/3})$. It is also worth noting that Elkin’s algorithm works for nonnegative real weights, not just integer weights.

We also show in [Appendix D](#) how to extend our algorithm to the case with κ -sources, providing some improvements on the results of Huang et al. [\[HNS17\]](#) for small κ and D .

Theorem 1.3. *There is a randomized distributed algorithm that computes the exact distances from κ given source nodes $s_1, s_2, \dots, s_\kappa$ to all nodes $t \in V$, in*

$$\tilde{O}\left(\min\left\{\kappa^{1/2}n^{3/4+o(1)} + \kappa^{1/3}n^{3/4}D^{1/6}, \kappa^{3/7}n^{6/7}\right\} + D\right)$$

rounds, with high probability.

2 Preliminaries

Definitions—Paths, Distances, and Neighborhoods For a path $P = (v_0, v_1, \dots, v_h)$ from v_0 to v_h , we define the *length* of the path P to be the summation of the weights of its edges, i.e., $w(P) := \sum_i w(v_{i-1}, v_i)$. Moreover, we say that P has h hops, where h is the number of edges in P . A *shortest path* from node s to t is an $s \rightarrow t$ path P that minimizes $w(P)$. Its length is denoted $d(s, t)$, also called the *distance* from node s to t . Given a source node s , the source-wise *s-radius* of the graph is equal to $\max_t d(s, t)$. Similarly, the *h-hop distance* $d^h(s, t)$ is the minimum length of an $s \rightarrow t$ path P , among all paths that have at most h hops. We define the *h-hop distance neighborhood* from a source s as the distances to all vertices $t \in V$ such that $d^h(s, t) = d(s, t)$, i.e., there exists a shortest $s \rightarrow t$ path with at most h hops; for all other $t \in V$ with $d^h(s, t) \neq d(s, t)$, their distance is undefined.

2.1 Tool 1: The Scaling Framework

The general case of shortest path problem is hard to tackle directly, while the problem becomes more manageable if the maximum distance from the source s to any other node t is somewhat small. To leverage this, we make use of Gabow’s scaling method [\[Gab83\]](#). We note that this method was used for the first time in the distributed setting for shortest paths by [\[HNS17\]](#).

What do we get from scaling? The interface of this method is as follows: Given a graph with integer weights in $[0, \Lambda]$, the scaling framework reduces the problem to $\log_2 \Lambda$ iterations of computing SSSP on a graph with nonnegative integer weights where in each iteration we have $d(s, t) \leq n - 1$ for all nodes $t \in V$. That is, in each iteration, the source-wise *s-radius* is bounded by $n - 1$. We emphasize that the weights in each of the iterations can be zero, and asymmetric, even if the original weights in the graph were all positive and symmetric. This poses a challenge for standard distributed distance computation methods, which we will discuss in the other tools.

To summarize, the scaling framework implies that, modulo this challenge of having to handle zero weights and asymmetric weights and an $O(\log \Lambda)$ factor overhead in the running time, it is sufficient to solve SSSP only on graphs satisfying $d(s, t) \leq n - 1$ for all t .

How does scaling work? Represent each weight $w(u, v)$ in base 2 using $T := \lfloor \log_2 \Lambda \rfloor + 1$ bits, including leading zeroes if necessary. Let $w_i(u, v)$ denote the integer representing the i most significant bits of $w(u, v)$. Let $b_i(u, v)$ denote the i ’th most significant bit of $w(u, v)$. The scaling algorithm is as described in [ScalingFramework](#).

We refer to [\[HNS17\]](#) for the following lemma, which is itself derived from earlier incarnations of the scaling framework in centralized and parallel settings [\[Gab83\]](#).

Lemma 2.1. *For each $i \in [T]$, ℓ_i and w_i satisfy the following: (1) For all (u, v) , $\ell_i(u, v) \geq 0$. (2) For all t , $d_{\ell_i}(s, t) \leq n - 1$. (3) For all t , $d_{w_i}(s, t) = 2d_{w_{i-1}}(s, t) + d_{\ell_i}(s, t)$. Hence, by induction, $\delta_i(v) = d_{\ell_i}(v)$ and $\tilde{d}_i(v) = d_{w_i}(v)$, and thus $d(v) = \tilde{d}_{\lfloor \log_2 \Lambda \rfloor}(v)$ is the correct distance.*

Algorithm 1 `ScalingFramework`($s, G = (V, E, w)$)

Input: Every node $v \in V$ knows the weight and direction of each edge $e \in E$ incident to v .

Output: Every node v learns $d(v)$, which equals the exact distance $d_w(s, v)$.

- 1: $\tilde{d}_0(v) \leftarrow 0$ for all nodes v
 - 2: **for** $i = 1, \dots, T$ **increasing do**
 - 3: Every node v broadcasts $\tilde{d}_{i-1}(v)$ to its neighbors
 - 4: Every node u computes $\ell_i(u, v) \leftarrow 2\tilde{d}_{i-1}(u) + w_i(u, v) - 2\tilde{d}_{i-1}(v)$ for each neighbor v
 - 5: $\delta_i(v) \leftarrow \text{SSSP}(s, \ell_i)$ \triangleright For each $v \in V$, $\text{SSSP}(s, \ell_i)$ informs node v the value of $d_{\ell_i}(s, t)$
 - 6: $\tilde{d}_i(v) \leftarrow 2\tilde{d}_{i-1}(v) + \delta_i(v)$ for all nodes v
 - 7: $d \leftarrow \tilde{d}_{\lceil \log_2 \Lambda \rceil}$
-

Corollary 2.2. After T iterations of *SSSP* on a graph with nonnegative integer weights such that $d(s, t) \leq n - 1$ for all $t \in V$, `ScalingFramework` correctly outputs all distances.

2.2 Tool 2: The ShortRange Algorithm

Another ingredient in our algorithm is to compute the h -hop distances from a given source s' (not necessarily the actual source s). For this, we use an algorithm of [HNS17], called `ShortRange`, which itself is a simple and clever hybrid of the two standard methods, BFS and Bellman-Ford. We next briefly discuss the shortcomings of these methods and then overview `ShortRange`.

The shortcomings of the standard methods We can approach the h -hop distances problem using the standard methods, namely Bellman-Ford and BFS. However, each of these has a shortcoming, as we discuss next. The Bellman-Ford (B-F) algorithm computes h -hop distances in h rounds: in each round, every vertex updates its best known distances and then broadcasts it to each neighbor. However, B-F comes with a downside in edge *congestion*: there may be $\Omega(h)$ messages passed along a single edge, which is the largest possible for an algorithm running in h rounds. This is problematic, especially because we will need to compute h -hop distances for multiple sources at the same time, one for each node in a chosen subset $V' \subseteq V$. In that case, the high edge congestion of this algorithm becomes a bottleneck. Specifically, if we run B-F from each $s' \in V'$, then an edge may need to pass $\Omega(h|V'|)$ messages, immediately giving a lower bound of $\Omega(h|V'|)$ rounds to compute all h -hop distances in the **CONGEST** model.

On the other hand, if the graph was unweighted, then breadth-first search (BFS) for h rounds would compute h -hop distances correctly, while sending at most one message along each edge. Hence, BFS would be a good candidate in terms of edge congestion. However, our graph is weighted and also contains zero-weight edges, which BFS alone cannot solve.

The ShortRange algorithm To address both of the above issues, [HNS17] introduced the algorithm `ShortRange`, which is a hybrid of BFS and B-F. It computes h -hop distances for a source s' in few rounds and with low edge congestion, at a cost of limiting itself only to $t \in V$ satisfying $d^h(s, t) \leq \ell$ for some parameter ℓ . That is, it computes the h -hop distance neighborhood restricted to only the nodes t satisfying $d(s, t) = d^h(s, t) \leq \ell$. For a node t not satisfying this restriction, the algorithm computes a distance that may overestimate $d(s, t)$, but is never an underestimate. The pseudo-code is deferred to **Appendix A**, due to space limitations.

Lemma 2.3. For any parameter $q \geq 1$, the *ShortRange* algorithm runs in $O(\ell q + h)$ rounds, and sends $O(h/q)$ messages along each edge. The distances that it computes satisfy the following properties: (1) $\forall t, d(t) \geq d_w(s, t)$. (2) If the shortest $s \rightarrow t$ path has at most h hops and length at most ℓ , then $d(t) = d_w(s, t)$.

2.3 Tool 3: Virtual Node Sampling

Following the steps of [UY91] and previous distributed shortest path algorithms [LPS13, Nan14, Elk17, HNS17], our algorithm samples a set of *virtual* nodes $V' \subseteq V$ as follows: every node independently joins the set with probability k/n , for a parameter k . We also add s to V' , i.e., s joins V' with probability 1 instead. A simple Chernoff bound shows that the size of V' is $O(k + \log n)$ w.h.p. Furthermore, V' satisfies the following crucial property:

Lemma 2.4. *W.h.p., for every $t \in V$, there is a shortest $s \rightarrow t$ path that has a virtual node inside every consecutive $\Theta(n \log n/k)$ vertices on the path. We call this a good $s \rightarrow t$ shortest path. In particular, the last virtual node on this path is within $O(n \log n/k)$ hops from node t .*

2.4 Tool 4: Distributed Scheduling of Algorithms

At certain points, our main algorithm will be calling a set of distributed subprocedures *in parallel*. That is, the computations of any two distinct subprocedures do not depend on each other. We cannot actually run all of these subprocedures in parallel, since many subprocedures might need to send a message through the same edge $e \in E$ simultaneously, but the CONGEST model only allows an $O(\log n)$ -bit sized message per round through e . On the other hand, we do not want to run these subprocedures sequentially, one after another. Our solution is to *schedule* the distributed algorithms, following [Gha15], based on two parameters called *dilation* and *congestion*.

Formally, consider a set of distributed algorithms A_1, \dots, A_k that run independently of one another. We define *dilation* as the maximum running time of any algorithm A_i , and *congestion* as the maximum number of times a single edge is used over all k algorithms. Formally, define $b_{e,i}$ as the number of times the algorithm A_i sends a message along edge $e \in E$, and let $\text{congestion} := \max_e \sum_{i=1}^k b_{e,i}$. Note that *dilation* and *congestion* are both lower bounds on the time it takes to run all the algorithms. It turns out that the upper bound nearly matches these lower bounds.

Theorem 2.5. [Gha15] *Given a set of independent algorithms with parameters *dilation* and *congestion*, we can run them all in time $\tilde{O}(\text{dilation} + \text{congestion})$, w.h.p.*

2.5 Tool 5: Additive Approximation of SSSP

In our exact SSSP computation, we make use of an *approximate* SSSP \tilde{d} that is accurate up to some *additive error* α ; that is, for all nodes t , we have $d(s, t) \leq \tilde{d}(s, t) \leq d(s, t) + \alpha$. To our knowledge, this is the first time a distributed exact shortest path algorithm relies on an approximate algorithm as a precomputation step. To compute this α -additive approximate SSSP, we use a method very similar to (a part of) the $(1 + \epsilon)$ -multiplicative approximation algorithm of Nanongkai [Nan14]. It is notable that we do not utilize the hop-set idea of Nanongkai’s algorithm, since hop-sets do not translate to directed graphs. Since Nanongkai [Nan14] does not explicitly state the dependencies on the error term ϵ , we present, for completeness, the full algorithm and its analysis in the Appendix. Finally, since the scaling framework guarantees us that the s -radius is at most $n - 1$, we can take $\epsilon := \alpha/n$ in the $(1 + \epsilon)$ -multiplicative approximation to obtain an α -additive approximation.

Lemma 2.6. *Suppose that the network graph has s -radius at most $n - 1$. Then, we can compute α -additive approximate SSSP in $\tilde{O}(\frac{n^2}{\alpha k} + kD)$ time, and with $\tilde{O}(k)$ messages sent along each edge.*

3 Main Algorithm

3.1 Algorithm Overview

In this section, we present our algorithm that proves [Theorem 1.1](#) and computes distances from source s to all nodes $t \in V$, in $\tilde{O}(n^{3/4}D^{1/4})$ rounds. We start with a high-level description of the

algorithm. On the outer level, this algorithm runs the [ScalingFramework](#) for $O(\log \Lambda)$ iterations, solving an SSSP problem in each iteration. Recall that, on every iteration, [ScalingFramework](#) reweights the edges on the graph so that the s -radius, defined to be $\max_t d(s, t)$, is at most $n - 1$. Once SSSP is solved on this reweighted graph, [ScalingFramework](#) proceeds to the next iteration, reweighting the edges again to have s -radius at most $n - 1$. This happens for $O(\log \Lambda)$ iterations, after which each node can locally deduce its correct distance from node s . Fix an iteration of [ScalingFramework](#), and let $G = (V, E)$ denote the reweighted graph with s -radius at most $n - 1$. We next describe our algorithm [SmallWeightSSSP](#), which solves SSSP on this graph.

Virtual Graph Following [Lemma 2.4](#), every node becomes a virtual node with probability k/n , for a parameter $k \leq n$ to be fixed, except for s which becomes virtual with probability 1.

To build intuition for the rest of the algorithm, let us first consider the task of computing $d(s, u')$ for only the virtual nodes $u' \in V'$. Let $h := \Theta(n \log n / k)$, and consider the *virtual graph* $G' = (V', E')$ with vertex set V' and an edge between each $u', v' \in V'$ of weight $d(u', v')$ if the shortest $u' \rightarrow v'$ path has $\leq h$ hops. The graph G' is called virtual because the edges in E' are not necessarily edges in the network, making direct communication between edges in E' impossible. Observe that, by [Lemma 2.4](#), running SSSP from $s \in V'$ in the virtual graph G' will, w.h.p., result in the same distances to each virtual node as running SSSP from node s in G . Therefore, a natural next step for an algorithm is to compute all edges in E' ; specifically, for every edge (u', v') , we want node v' to know its weight and direction.³

However, our algorithm will not compute all edges in E' , but only a subset of the edges. Define an edge $(u', v') \in E'$ to be *light* if $d(u', v') \leq \ell$ for a parameter ℓ , and *heavy* otherwise. Let $G'_L = (V', E'_L)$ and $G'_H = (V', E'_H)$ be G' with only light edges and heavy edges in E' , respectively. The algorithm will compute all light edges in E'_L , so that for each directed edge $(u', v') \in E'_L$, the endpoint v' knows that edge. For the heavy edges in E'_H , computing them can be time-consuming; thus, the algorithm will take another approach. We next discuss these two cases separately.

Light Edges To compute all light edges, we run [ShortRange](#) (u', G, h, ℓ, q) for each virtual node $u' \in V'$, for some parameter q . For each pair $u', v' \in V'$, let $d_{u'}(v')$ be the value of $d(v')$ returned by [ShortRange](#) (u', G, h, ℓ, q) . If $d_{u'}(v') < \infty$, then v' can now learn a virtual edge (u', v') with weight $d_{u'}(v')$. For each light edge (u', v') in E'_L , v' learns the virtual edge (u', v') with correct weight $d(u', v')$, due to the guarantee of [ShortRange](#). There may be other extraneous virtual edges learned by virtual nodes that overestimate their distance, but, as we will see, they will not affect the output.

Heavy Edges Note that the only important virtual edges in E'_H are those that belong to the shortest path tree from node s in G' . Hence, it suffices to limit our attention to the heavy edges $(u', t') \in E'_H$ that belong to a shortest $s \rightarrow t'$ path in G' , for some node t' . In this case, we leverage the fact that $d(s, t') > d(s, u') + \ell$. We partition V into $\lfloor n/\ell \rfloor + 1$ *buckets* such that bucket $i \in [0, \lfloor n/\ell \rfloor]$, denoted V_i , has all nodes v with $d(s, v) \in [i\ell, (i+1)\ell)$. We will discuss later how to compute the buckets.⁴ Observe that the bucket containing u' has a smaller index than the one containing t' . The algorithm will compute the correct distances to each bucket in increasing bucket index i , with the invariant that before processing bucket V_i , all nodes in V_0 through V_{i-1} know their correct distance from node s . To process bucket V_i , the algorithm first executes a single B-F to depth h ,⁵ then runs a *virtual SSSP*, and then runs another B-F to depth h . We will soon explain

³Since communication on the network is bidirectional, it turns out that we can also allow node u' to learn edge (u', v') . However, this will not be necessary.

⁴At first glance, this might seem troublesome, since it requires knowing the distances beforehand.

⁵In the algorithm pseudocode, we run an algorithm [Extend](#) instead. The reader is encouraged to replace the [Extend](#) calls with calls of B-F to depth h ; see Remark 3.1. We use B-F for this discussion to simplify this overview.

the *virtual SSSP* algorithm and the difficulties in it, as it needs to be run on a virtual graph. By a B-F to depth h , we mean initializing the B-F with distances $d(v)$ for all $v \in V_i$, and then running B-F, which broadcasts and updates distances for h rounds. Going back to our consideration of a heavy edge $(u', t') \in E'_H$, note that, right before the algorithm processes the bucket containing t' , the bucket containing u' has already been processed, and u' knows its correct distance. Then, when the algorithm processes the bucket containing t' , in the first B-F to depth h , node u' will propagate its distance to node t' , informing t' of its correct distance from s .

Virtual SSSP We now discuss how we run the virtual SSSP when processing each bucket. Suppose the algorithm is currently processing bucket V_i . By the invariant, all nodes in V_1 through V_{i-1} already know their correct distance, and the objective is now for those of bucket V_i to learn their distance. Note that a B-F to depth h alone is not enough. This is because edges can have weight zero and thus, a virtual node in V_i can be a large number of hops away from any node in V_1 through V_{i-1} , even if the total weighted distance is small.

Consider some virtual node $t' \in V_i$ and the good $s \rightarrow t'$ path $P_{t'}$, guaranteed to exist by [Lemma 2.4](#). Although there may be many virtual nodes on $P_{t'}$ belonging to bucket V_i , the key observation is that, of the virtual nodes in $P_{t'} \cap V_i$ ordered by $P_{t'}$, the *first* node knows its correct distance, and for the remaining nodes, there is an edge in E'_L from each one to the next. For this *first* virtual node $v' \in P_{t'} \cap V_i$, we will show that, after the first B-F in processing bucket V_i —as mentioned above when discussing heavy edges—node v' will learn its correct distance. Afterward, we would like to propagate the distance of v' along the edges in E'_L by solving a SSSP problem on G'_L . We define this SSSP instance as follows: we begin with the subgraph G'_L induced by $\{s\} \cup V_i$, and for each node $v' \in V_i$ that has computed some (possibly incorrect) distance $d(v')$, we add the edge (s, v') with weight $d(v')$. We can think of each additional edge (s, v') as *shortcutting* some path from s to v' . We now run SSSP on this subgraph, but we cannot communicate along the virtual edges in G'_L directly, since they may not exist in the network. Hence, we instead run a *virtual SSSP* on the subgraph of G'_L induced by $V' \cap V_i$. This [VirtualSSSP](#) is described in [Section 3.4](#). Together, [VirtualSSSP](#) and the B-Fs correctly compute all distances from node s to the virtual nodes.

Computing the Buckets We now discuss how to compute the buckets V_i . We will actually compute a set of nodes \tilde{V}_i that contains V_i , so that the subgraph of G'_L induced by $V' \cap \tilde{V}_i$ contains all vertices and edges in the target subgraph. However, we also need that the \tilde{V}_i are, on average, not much larger than the V_i , in order to maintain a fast algorithm.

First, we use the additive approximate distance algorithm of [Section 2.5](#) to compute ℓ -additive approximate distances from node s to every other node. Since the buckets have distance intervals of size ℓ , each node $v' \in V'$ can pinpoint down two buckets, one of which contains v' . Then, define \tilde{V}_i to include every $v' \in V'$ that guesses V_i as one of its two possible buckets. The total size of all $V' \cap \tilde{V}_i$ is at most twice the total size of all $V' \cap V_i$, so the asymptotic running time is unaffected.

Extending the Distance Computation to Non-Virtual Nodes By [Lemma 2.4](#), every node $t \in V$ has a shortest $s \rightarrow t$ path that contains a virtual node $v' \in V'$ within its last h vertices, w.h.p. Consider the iteration when the algorithm processes the bucket containing v' . Node v' learns its correct distance $d_w(s, v')$ right after the virtual SSSP computation. Then, in the second B-F to depth h , node v' propagates its distance to node t , and node t learns its correct distance. Thus, node t learns its correct distance after the algorithm processes the bucket containing v' , which either equals or appears earlier than the bucket containing t . Hence, once the algorithm finishes bucket V_i , all nodes in buckets V_1 through V_i learn their correct distance, satisfying the invariant.

3.2 The Algorithm

In this section, we give a detailed description of our main algorithm, which computes exact shortest paths in $\tilde{O}(n^{3/4}D^{1/4})$ time, proving [Theorem 1.1](#). The main algorithm, called [Main](#), first computes an ℓ -additive approximate SSSP from source s ; recall that this approximate computation allows each node to guess which bucket it belongs to. Then, it runs [ScalingFramework](#), reducing the problem to $O(\log \Lambda)$ computations of SSSP, each time on a graph with s -radius at most $n - 1$. On each iteration, we call the subprocedure [SmallWeightSSSP](#), which is the SSSP algorithm with the promised s -radius bound. We will defer the pseudo-code of [Main](#) to [Appendix B](#).

Parameters The algorithm will maintain the parameters k, h, ℓ, q , as described below. For an initial read, we encourage the reader to assume that $D = \tilde{O}(1)$ and ignore the D factors.

- k is the desired number of virtual nodes. Every node becomes virtual with probability k/n , except for s which is virtual. W.h.p., we sample $O(k + \log n)$ virtual nodes.
- h is a parameter given to the [ShortRange](#) algorithm, a bound on the number of hops. We set $h = \Theta(n \log n/k)$ so that, by [Lemma 2.4](#), w.h.p. for every node t , there is an $s \rightarrow t$ path with a virtual node inside every consecutive $\Theta(n \log n/k)$ vertices on the path.
- ℓ represents four quantities, which we will observe are closely related to one another. First, it is the parameter into the [ShortRange](#) algorithm, a bound on the maximum distance. Second, it is the parameter in the initial approximate shortest paths computation; that is, the algorithm initially computes ℓ -additive approximate shortest path distances from source s . Third, it is the threshold such that an edge in the virtual graph is called light if its weight is at most ℓ , and heavy otherwise. Lastly, it is the width of each bucket: bucket i represents all nodes whose distance from s is within $[i\ell, (i+1)\ell)$.

These parameters are summarized below, along with their values in the $\tilde{O}(n^{3/4}D^{1/4})$ algorithm.

Param.	Value	Description
k	$n^{3/4}D^{-3/4}$	The number of virtual nodes we want to sample.
h	$\Theta(n \log n/k) = \tilde{O}(n^{1/4}D^{3/4})$	The parameter for the number of hops in the ShortRange algorithm.
ℓ	$n^{1/2}D^{1/2}$	The parameter for the maximum distance in the ShortRange algorithm. Also, the error of the initial additive approximation.
q	$n^{1/4}D^{-1/4}$	The optimization parameter for the ShortRange algorithm.

Algorithm Description Below, we present the pseudocodes for [SmallWeightSSSP](#) and [Extend](#).

Algorithm 2 `SmallWeightSSSP`($s, G = (V, E, w), \tilde{d}, k, \ell, q$)

Input: Every node v knows the weight and direction of each edge $e \in E$ incident to v and the ℓ -additive distance estimate $\tilde{d}(v)$. We have the promise that $d_w(s, t) \leq n - 1, \forall t \in V$.

Output: Every node v learns $d(v)$, which equals the exact distance $d_w(s, v)$.

- 1: Every node becomes virtual with probability k/n , except for s , which becomes virtual with probability 1. Let $V' \subseteq V$ be the set of virtual nodes. // We have $s \in V'$ and, w.h.p., we sample $O(k + \log n)$ virtual nodes.
 - 2: Set $h \leftarrow \Theta(n \log n / k)$. For each $s' \in V'$, let $d_{s'}$ be the result of `ShortRange`(s', G, h, ℓ, q). Each $v \in V$ knows $d_{s'}(v)$ for each $s' \in V'$.
 - 3: Build graph $G' = (V', E', w')$ with directed edge $(u', v') \in E'$ having weight $d_{u'}(v')$, for each pair $(u', v') \in V' \times V'$ satisfying $d_{u'}(v') < \infty$. Every $v' \in V'$ knows each directed edge $(u', v') \in E'$ and its weight.
 - 4: Initialize distances $d(s) \leftarrow 0, d(v) \leftarrow \infty$ for all $v \in V \setminus s$.
 - 5: **for** $i = 0, \dots, \lfloor n/\ell \rfloor$ **do** // Iterate over buckets $V_0, \dots, V_{\lfloor n/\ell \rfloor}$.
 - 6: Run `Extend`(G, d, h, ℓ, i), and let d be the updated distance that each node knows. // Equivalently, run a B-F to depth h .
 - 7: Define \tilde{V}'_i as $\{v' \in V' : \tilde{d}(v') \in [i\ell, (i+2)\ell)\}$. // \tilde{V}'_i contains all virtual nodes that guess bucket i as one of their two guesses. In particular, $V' \cap V_i \subseteq \tilde{V}'_i$.
 - 8: Consider the subgraph of G' induced by $\{s\} \cup \tilde{V}'_i$. For each $v' \in \tilde{V}'_i$ with $d(v') < \infty$, add an edge (s, v') of weight $d(v')$ to this subgraph. Call the resulting graph \tilde{G}'_i . // Note that every node $v' \in \tilde{V}'_i$ still knows every edge (u', v') in \tilde{G}'_i .
 - 9: Run `VirtualSSSP`(s, \tilde{G}'_i), and let \tilde{d}'_i be the resulting distance that each node learns.
 - 10: Every $v' \in V'$ updates $d(v') \leftarrow \min\{d(v'), \tilde{d}'_i(v')\}$.
 - 11: Run `Extend`(G, d, h, ℓ, i), and let d be the updated distance that each node knows. // Equivalently, run a B-F to depth h .
-

Algorithm 3 `Extend`(G, d, h, ℓ, i)

Input: Every node v knows its current distance estimate $d(v)$.

Output: Every node v updates its distance estimate $d(v)$.

- 1: **for** h rounds **do** // Run a *restricted* B-F to depth h . See Remark 3.1.
 - 2: Every node u sends $d(u)$ to all neighbors v with $d(u) + w(u, v) \in [i\ell, (i+1)\ell)$
 - 3: Every node v sets $d(v) = \min\{d(v), d(u) + w(u, v)\}$ for each $d(u)$ received from neighbor u
-

Remark 3.1. On line 6 and line 11, we can replace the algorithm `Extend` with a B-F to depth h for the same correctness and running time of `SmallWeightSSSP`.⁶ The reader is encouraged to do so for a simpler algorithm and the same proof of correctness. However, we need `Extend` for the multiple sources case in Appendix D.

3.3 Analysis

First, we prove the aforementioned invariant: after the algorithm processes bucket V_i , all nodes in buckets V_1 through V_i know their correct distance.

Lemma 3.2. At the end of each iteration i of the **for** loop, $d(t) = d_w(s, t)$ for all t such that $d(t) \in [0, (i+1)\ell)$. Hence, at the end of `SmallWeightSSSP`, $d(t) = d_w(s, t)$ for all $t \in V$. In other words, every node knows its correct distance from node s .

Proof. First, consider the base case iteration 0 and a node t with $d(s, t) \in [0, (i+1)\ell) = [0, \ell)$. Consider the shortest $s \rightarrow t$ path P_t satisfying the properties in Lemma 2.4, and let t' be the last

⁶The algorithm `Extend` is actually a *restricted* B-F to depth h , in that only nodes that satisfy a certain property can broadcast their distance.

virtual node on P_t , possibly s itself. Since each virtual node v' on the path satisfies $d(s, v') \in [0, \ell)$, its approximate distance satisfies $\tilde{d}(s, v') \in [0, 2\ell)$, so $v' \in \tilde{V}'_0$. Also, w.h.p., consecutive virtual nodes in P_t are within $h = O(n \log n/k)$ hops away from each other. Since they are also distance $\leq \ell$ apart, they are connected by edges in E' , and **VirtualSSSP**(s, \tilde{G}'_i) (line 9) correctly computes $\tilde{d}'_i(t') = d(t')$. Here, we use the fact that edges in \tilde{G}'_i can only be overestimates of their true distances in G , so that we can never have $\tilde{d}'_i(t') < d(t')$. Finally, since t' and t are $\leq h$ hops apart on P_t , and since every node in between t' and t has distance from node s in the range $[0, \ell)$, the second **Extend** algorithm (line 11) correctly computes $d(v)$ for each v in between t' and t on P_t . In particular, t will know its correct distance $d(t)$.

We now apply induction on i . If node t satisfies $d(t) \in [0, i\ell)$, then the statement $d(t) = d_w(s, t)$ has already been proved on iteration $i - 1$, so assume that $d(t) \in [i\ell, (i + 1)\ell)$. Take the shortest $s \rightarrow t$ path P_t , let u be the last node on P_t satisfying $d(s, u) < i\ell$, let v be the next node after u on P_t , and let v' be the next virtual node after u on P_t . Let t' be the last virtual node on P_t , possibly v' or s itself. Suppose that $d(u) \in [j\ell, (j + 1)\ell)$ for some $j < i$. By induction, $d(u)$ was correctly computed at the end of iteration j . Observe that all nodes on P_t from v to v' have their distance from node s in the range $[i\ell, (i + 1)\ell)$, and moreover, v and v' are $\leq h - 1$ hops apart on P_t w.h.p. Therefore, for the h iterations of the first **Extend** (line 6) on iteration i , u will first send $d(u) + w(u, v)$ to v , and then $h - 1$ rounds later, v' will update $d(v')$ to be its correct distance. The rest of the argument follows similarly to the base case. All virtual nodes in $P_{t'}$ from v' onwards belong to \tilde{V}'_i , and since they are $\leq h = O(n \log n/k)$ hops away from each other, they are connected by edges in \tilde{G}'_i . Therefore, the path consisting of s, v' , and all virtual nodes after v' on $P_{t'}$ has length exactly $d(t')$ in \tilde{G}'_i , and the **VirtualSSSP** (line 9) of iteration i computes $d'(t')$ correctly. Finally, using the same argument as in the base case, the second **Extend** (line 11) correctly computes $d(t)$. \square

Lemma 3.3. **SmallWeightSSSP** runs in $\tilde{O}\left(\ell q + \frac{n}{q} + kD + \frac{n^2}{\ell k}\right) = \tilde{O}(n^{3/4}D^{1/4})$ time.

Proof. Computing an ℓ -additive approximate SSSP takes $\tilde{O}(n^2/(\ell k) + kD)$ time by **Lemma 2.6**. We use the same value of k for ℓ -AdditiveErrorSSSP as we do for **SmallWeightSSSP**.

We run **ShortRange**(\cdot, \cdot, h, ℓ, q) for k nodes, with $h = O(n \log n/k)$, and each node congests a given edge $O(h/q)$ times, so the total congestion is $O(k(n \log n/k)/q) = O(n \log n/q)$. The dilation is $O(\ell q + h) = O(\ell q + n \log n/k)$. We can schedule all the **ShortRange** algorithms in time $\tilde{O}(\text{dilation} + \text{congestion})$ by **Theorem 2.5**, giving time $\tilde{O}(\ell q + n/k + n/q)$.

For $O(n/\ell)$ iterations, we run **VirtualSSSP** on a graph with $|\tilde{V}'_i| + 1$ vertices, which takes $O((|\tilde{V}'_i| + 1)D \log n)$ time by **Lemma 3.4**, and **Extend**($\cdot, \cdot, h, \cdot, \cdot$), which takes time $h + 1 = O(n \log n/k)$. Every $v' \in V'$ belongs to at most 2 sets \tilde{V}'_i , so $\sum_i |\tilde{V}'_i| \leq 2k$ and the total running time for **VirtualSSSP** and B-F is $\tilde{O}(kD + n/\ell \cdot n/k)$.

Setting $k := n^{3/4}D^{-3/4}$, $\ell := n^{1/2}D^{1/2}$, $q := n^{1/4}D^{-1/4}$ gives the desired time. \square

3.4 SSSP on Virtual Graph

From **SmallWeightSSSP**, we have the following setting: there is a virtual graph $G' = (V', E', w')$ such that $d_{w'}(s, v') \leq n - 1$ for all $v' \in V'$, and for each edge $(u', v') \in E'$ directed from u' to v' , only v' knows its weight and direction. Recall that since edges are virtual, they may not exist in the communication network, so we cannot pass information directly along edges in E' . For **VirtualSSSP**, our goal is to inform each virtual node $v' \in V'$ of its exact distance $d_{w'}(s, v')$.

Algorithm VirtualSSSP Let $k := |V'|$. The algorithm **VirtualSSSP** first runs **ScalingFramework** on G' , reducing the problem to $\log n$ iterations of SSSP with the guarantee $d(s, v') \leq k - 1$. After each iteration, instead of local broadcasts from each node, we let every node $v' \in V'$ broadcast its

computed $d(v')$ to all other nodes, which takes $O(k + D)$ time using standard pipelining techniques on a BFS tree, so that weights on the next iteration can be computed.

The algorithm for each iteration of **ScalingFramework** is promised that $d_{w'}(s, v') \leq k - 1, \forall v' \in V'$. We now discuss this algorithm, which we name **SmallWeightVirtualSSSP**; its pseudocode is deferred to **Appendix B**. It is instructive to first consider this algorithm in a parallel computation setting, as follows. For each $i \in [0, k - 1]$, initialize a queue Q_i to be used in the algorithm. To begin, insert the value $(s, 0)$ to queue Q_0 . We process the queues in increasing order Q_0, Q_1, \dots, Q_{k-1} , continuing onto the next queue whenever the current queue is empty. When processing Q_i , we (1) mark each $u' \in Q_i$ as completed, (2) write $d(s, u')$ for each $u' \in Q_i$ in (centralized) memory, and (3) for each unmarked v' with an edge $(u', v') \in E'$, remove v' from its current queue (if any) and insert it to Q_j with $j := i + \min_{(u', v') \in E'} w(u', v')$. Note that we may have to process Q_i again, since there may be zero-weight edges. It is straightforward to argue that this algorithm is correct. Furthermore, since at least one node is marked every iteration, the algorithm runs in $O(k)$ parallel rounds.

Finally, we translate this algorithm to a distributed setting. Instead of writing each completed $u' \in Q_i$ to memory, we broadcast $d(s, u')$ to all other nodes in time $O(|Q_i| + D)$ so that all unmarked nodes can locally compute $\min_{(u', v') \in E'} w(u', v')$. To figure out when to advance to the next queue, we continue to process the current queue Q_i until there is a round with no broadcasts. At that point, all nodes can advance to queue Q_{i+1} in a synchronized manner.

Lemma 3.4. **SmallWeightVirtualSSSP** on a virtual graph $G' = (V', E')$ takes $O(|V'|D)$ rounds, and sends $O(|V'|)$ along each edge. Hence, **VirtualSSSP** on a virtual graph $G' = (V', E')$ takes $O(|V'|D \log n)$ rounds, and sends $O(|V'| \log n)$ messages along each edge.

Proof. Every node broadcasts once, so the total number of broadcasts is $O(|V'|)$. The while loop terminates in $O(|V'|)$ rounds, since in each round, either some node broadcasts (and then never again), or i increases. Every broadcast can be executed in $O(D)$ time through an auxiliary BFS tree, so in total, everything takes $O(|V'|D)$ time. Similarly, every broadcast congests each edge of the BFS tree once, so the total congestion is $O(k)$. \square

Acknowledgement The first author is grateful to Michael Elkin for explaining his result [Elk17] and discussions about potential improvement directions, and to Danupon Nanongkai for sharing an earlier manuscript of [HNS17].

References

- [Bel58] Richard Bellman. On a routing problem. *Quarterly of applied mathematics*, 16(1):87–90, 1958.
- [BKKL27] Ruben Becker, Andreas Karrenbauer, Sebastian Krinninger, and Christoph Lenzen. Approximate undirected transshipment and shortest paths via gradient descent. In *Proc. of the Int’l Symp. on Dist. Comp. (DISC)*, 2017, arXiv preprint arXiv:1607.05127.
- [DSHK⁺11] Atish Das Sarma, Stephan Holzer, Liah Kor, Amos Korman, Danupon Nanongkai, Gopal Pandurangan, David Peleg, and Roger Wattenhofer. Distributed verification and hardness of distributed approximation. In *Proc. of the Symp. on Theory of Comp. (STOC)*, pages 363–372, 2011.
- [Elk04a] Michael Elkin. Distributed approximation: a survey. *ACM SIGACT News*, 35(4):40–57, 2004.

- [Elk04b] Michael Elkin. Unconditional lower bounds on the time-approximation tradeoffs for the distributed minimum spanning tree problem. In *Proc. of the Symp. on Theory of Comp. (STOC)*, pages 331–340, 2004.
- [Elk17] Michael Elkin. Distributed exact shortest paths in sublinear time. In *Proc. of the Symp. on Theory of Comp. (STOC)*, pages 757–770. ACM, 2017.
- [FJ56] Lester R Ford Jr. Network flow theory. Technical report, RAND CORP SANTA MONICA CA, 1956.
- [Gab83] Harold N Gabow. Scaling algorithms for network problems. In *Foundations of Computer Science, 1983., 24th Annual Symposium on*, pages 248–258. IEEE, 1983.
- [Gha15] Mohsen Ghaffari. Near-optimal scheduling of distributed algorithms. In *the Proc. of the Int’l Symp. on Princ. of Dist. Comp. (PODC)*, 2015.
- [GKP93] J.A. Garay, S. Kutten, and D. Peleg. A sub-linear time distributed algorithm for minimum-weight spanning trees. In *Proc. of the Symp. on Found. of Comp. Sci. (FOCS)*, 1993.
- [HKN16] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. A deterministic almost-tight distributed algorithm for approximating single-source shortest paths. In *Proc. of the Symp. on Theory of Comp. (STOC)*, pages 489–498. ACM, 2016.
- [HNS17] Chien-Chung Huang, Danupon Nanongkai, and Thatchaphol Saranurak. Distributed exact weighted all-pairs shortest paths in $\tilde{O}(n^{5/4})$ rounds. In *Proc. of the Symp. on Found. of Comp. Sci. (FOCS)*, 2017.
- [KP95] Shay Kutten and David Peleg. Fast distributed construction of k-dominating sets and applications. In *the Proc. of the Int’l Symp. on Princ. of Dist. Comp. (PODC)*, pages 238–251, 1995.
- [LPS13] Christoph Lenzen and Boaz Patt-Shamir. Fast routing table construction using small messages: Extended abstract. In *Proc. of the Symp. on Theory of Comp. (STOC)*, pages 381–390, 2013.
- [MH16] M. Ghaffari and Bernhard Haeupler. Distributed algorithms for planar networks II: Low-congestion shortcuts, mst, and min-cut. In *Pro. of ACM-SIAM Symp. on Disc. Alg. (SODA)*, 2016.
- [MK13] M. Ghaffari and Fabian Kuhn. Distributed minimum cut approximation. In *Proc. of the Int’l Symp. on Dist. Comp. (DISC)*, pages 1–15, 2013.
- [Nan14] Danupon Nanongkai. Distributed approximation algorithms for weighted shortest paths. In *Proc. of the Symp. on Theory of Comp. (STOC)*, 2014.
- [NS14] Danupon Nanongkai and Hsin-Hao Su. Almost-tight distributed minimum cut algorithms. In *International Symposium on Distributed Computing*, pages 439–453. Springer, 2014.
- [Pel00] David Peleg. *Distributed Computing: A Locality-sensitive Approach*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.

- [PR99] David Peleg and Vitaly Rubinfeld. A near-tight lower bound on the time complexity of distributed MST construction. In *Proc. of the Symp. on Found. of Comp. Sci. (FOCS)*, pages 253–, 1999.
- [UY91] Jeffrey D Ullman and Mihalis Yannakakis. High-probability parallel transitive-closure algorithms. *SIAM Journal on Computing*, 20(1):100–125, 1991.

A Missing Details of Section 2

Algorithm 4 ShortRange($s, G = (V, E, w), h, \ell, q$)

Input: Source s , network graph $G = (V, E, w)$, and parameters h, ℓ, q . Every node v knows the weight and direction of each edge $e \in E$ incident to v .

Output: Every node v learns $d(v)$, which satisfies $d(v) \geq d(s, v)$. If v is within the h -hop distance neighborhood and satisfies $d(s, v) \leq \ell$, then $d(v) = d(s, v)$.

```

1:  $d(s) \leftarrow 0, d(t) \leftarrow \infty$  for all  $t \neq s$ 
2:  $w'(u, v) \leftarrow \max\{w(u, v), 1/q\}$  for each  $(u, v)$   $\triangleright w'(u, v) = 1/q$  if  $w(u, v) = 0$  and  $w(u, v)$  otherwise
3: for  $i = 0, \dots, \ell q + h$  do  $\triangleright$  BFS
4:   Every node  $u$  with  $d(u) = i/q$  broadcasts  $d(u)$  to its neighbors
5:   If node  $v$  receives  $d(u)$ , updates  $d(v) \leftarrow \min\{d(v), d(u) + w'(u, v)\}$ 
6:  $d(v) \leftarrow \lfloor d(v) \rfloor$  for all  $v$ 
7: for  $h$  iterations do  $\triangleright$  Bellman-Ford (B-F)
8:   Every node  $u$  broadcasts  $d(u)$  unless (1)  $d(u)$  has been broadcast previously, or (2)  $u$  has broadcast  $\lfloor h/q \rfloor$  times already
9:   If node  $v$  receives  $d(u)$ , updates  $d(v) \leftarrow \min\{d(v), d(u) + w(u, v)\}$ 
10: return  $d$ 

```

Proof of Lemma 2.3. First, we show that $d(t) \geq d_w(s, t)$ for all t . Let $d'(v)$ be the value of $d(v)$ right after line 6 of ShortRange. Since the weights w' are overestimates of w , we have $d'(t) \geq d_w(s, t)$. By the properties of B-F, if a node t has value $d(t)$ at the end of the algorithm, then there is a path P from some node $v \in V$ to t such that $d(t) = d'(v) + w(P)$. Therefore, $d(t) = d'(v) + w(P) \geq d_w(s, v) + w(P) \geq d_w(s, t)$.

Now suppose that the shortest $s \rightarrow t$ path P has $\leq h$ hops and length $\leq \ell$. We have $d_{w'}(s, t) \leq d_w(s, t) + h/q \leq \ell + h/q$, since this path has $\leq h$ edges of length 0 which add $\leq h/q$ additional length according to w' . Moreover, every node $v \in P$ satisfies $d_{w'}(s, v) \leq d_w(s, v) + h/q \leq \ell + h/q$. The BFS (lines 3–5) computes $d(v) = d_{w'}(s, v)$ for all v satisfying $d_{w'}(s, v) \leq (\ell q + h)/q = \ell + h/q$, so for all $v \in P$, $d(v) = d_{w'}(s, v)$ by line 5. Observe that if the B-F (lines 7–9) did not have restriction (2), then all distances within h hops would be correctly computed. However, after line 6, every node $v \in P$ has $d(v) = \lfloor d_{w'}(s, v) \rfloor \leq d_w(s, v) + \lfloor h/q \rfloor$, and since every broadcast results from decreasing the estimated $d(v)$ by ≥ 1 , each node v will indeed broadcast its correct $d(s, v)$ within $\lfloor h/q \rfloor$ broadcasts. In particular, $d(t)$ will be the correct $d_w(s, t)$ at the end of the algorithm. \square

Proof of Lemma 2.4. For each node $t \in V$, fix a shortest $s \rightarrow t$ path $P_t = (s = v_0, \dots, v_h = t)$, and let $C := cn \ln n/k$ for a constant c . If $h < C$, then the first part of the statement is vacuous, and the second part is immediate since s is a virtual node within C hops from node t .

If $h \geq C$, then for each $i \in [h - C + 1]$, the probability that none of the nodes v_i, \dots, v_{i+C-1} is virtual is exactly

$$(1 - k/n)^C = (1 - k/n)^{cn \ln n/k} < \exp(k/n \cdot cn \ln n/k) = n^{-c}.$$

Taking a union bound over all values of i and all $n - 1$ paths P_t gives the w.h.p. result. \square

Approximate SSSP, proving Lemma 2.6 Next, we prove Lemma 2.6, restated below.

Lemma 2.6. *Suppose that the network graph has s -radius at most $n - 1$. Then, we can compute α -additive approximate SSSP in $\tilde{O}(\frac{n^2}{\alpha k} + kD)$ time, and with $\tilde{O}(k)$ messages sent along each edge.*

To compute an α -additive approximation on a graph with s -radius $\leq n - 1$, it suffices to compute a $(1 + \alpha/n)$ -multiplicative approximation, i.e., $d(s, t) \leq \tilde{d}(s, t) \leq (1 + \alpha/n)d(s, t)$. Let $\epsilon := \alpha/n$, so that we are looking for a $(1 + \epsilon)$ -multiplicative approximation (with running time dependent on ϵ).

First, we use the algorithm of [Nan14] to compute h -hop distances. The following is a restatement of Theorem 3.2 from [Nan14] with the added dependence on ϵ in the running time.

Theorem A.1. **BoundedHopSSSP** computes $(1 + \epsilon)$ -multiplicative approximate h -hop distances in $\tilde{O}(\epsilon^{-1}h)$ time, and $\tilde{O}(1)$ messages are sent along each edge.

Proof. For proof of correctness, we refer the reader to Theorem 3.2 of [Nan14]. Note that since G has s -radius $\leq n - 1$, the range $[\lceil \log(n - 1) \rceil]$ suffices in line 2 of **BoundedHopSSSP**, as is explained in Lemma 3.4 of [Nan14].

For running time and edge congestion, **BoundedDistanceSSSP** clearly runs in $K = (1 + 2/\epsilon)h = O(\epsilon^{-1}h)$ rounds, and every edge (u, v) has at most two messages sent along it, one from u 's broadcast and one from v 's broadcast. **BoundedHopSSSP** runs $O(\log n)$ iterations of **BoundedDistanceSSSP**, so the total running time is $O(\epsilon^{-1}h \log n)$ and the total edge congestion is $O(\log n)$. \square

Algorithm 5 **BoundedHopSSSP**($G = (V, E, w), s, h$)

Input: Weighted directed graph G , source vertex s , and integer h . G has s -radius $\leq n - 1$.

Output: Every node u knows the value of $\tilde{d}^h(s, u)$ such that $d^h(s, u) \leq \tilde{d}^h(s, u) \leq (1 + \epsilon)d^h(s, u)$.

- 1: Let t be the time this algorithm starts. We can assume that all nodes know t .
 - 2: For all $i \in [\lceil \log(n - 1) \rceil]$ and edge (x, y) , let $D'_i \leftarrow 2^i$ and $w'_i(x, y) \leftarrow \lceil \frac{2hw(x, y)}{\epsilon D'_i} \rceil$. Let $K \leftarrow (1 + 2/\epsilon)h$.
 - 3: **for all** i **do**
 - 4: $d'_i \leftarrow \text{BoundedDistanceSSSP}(G, w'_i, s, K)$
 - 5: $\tilde{d}^h(s, u) \leftarrow \frac{\epsilon D'_i}{2h} \min_i d'_i(s, u)$
-

Algorithm 6 **BoundedDistanceSSSP**(G, w, s, K)

Input: Weighted directed graph G , source vertex s , and integer K .

Output: Every node u knows $d'(s, u)$ where $d'(s, u) = d(s, u)$ if $d(s, u) \leq K$ and $d'(s, u) = \infty$ otherwise.

- 1: Let t be the time this algorithm starts. We can assume that all nodes know t .
 - 2: Initially, every node u sets $d'(s, u) \leftarrow \infty$.
 - 3: Source node s sends a message $(s, 0)$ to itself.
 - 4: **if** a node u receives a message (s, ℓ) for some ℓ from node v **then**
 - 5: **if** $(\ell + w(u, v) \leq K)$ and $(\ell + w(u, v) < d'(s, u))$ **then**
 - 6: u sets $d'(s, u) \leftarrow \ell + w(u, v)$.
 - 7: For any $x \leq K$, at time $t + x$, every node u such that $d'(s, u) = x$ broadcasts message (s, x) to all its neighbors to announce that $d'(s, u) = x$.
-

If we sample k virtual nodes (including the source), we can run this algorithm with $h := O(n \log n / k)$ from each virtual node, which is $O(k \log n)$ total congestion. Scheduling these algorithms in $O(\text{dilation} + \text{congestion} \cdot \log n)$ time, we get $O(\epsilon^{-1}h \log n + k \log^2 n) = \tilde{O}(\epsilon^{-1}n/k + k)$.

On the virtual graph, since all edge weights are multiples of $\frac{\epsilon D'_i}{2h}$, we can scale all weights by $\frac{2h}{\epsilon D'_i}$ so that they are integers, and then run **VirtualSSSP** from [Section 3.4](#) in $O(kD \log n)$ time and $O(k \log n)$ congestion along each edge. Every node then scales its distance by $\frac{\epsilon D'_i}{2h}$, giving a $(1 + \epsilon)$ -approximate distance from the source. Finally, extend the $(1 + \epsilon)$ -approximate h -hop SSSP algorithm to all nodes in the same $\tilde{O}(\epsilon^{-1}n/k + k)$ running time and $\tilde{O}(1)$ congestion, obtaining the following:

Lemma A.2. *We can compute $(1 + \epsilon)$ -multiplicative approximate SSSP in $\tilde{O}(\epsilon^{-1}n/k + kD)$ time and $\tilde{O}(k)$ congestion along each edge.*

Applying [Lemma A.2](#) with $\epsilon := \alpha/n$ proves [Lemma 2.6](#).

B Missing Details of [Section 3](#)

Algorithm 7 $\text{Main}(G, s)$

Input: Every node v knows the weight and direction of each edge $e \in E$ incident to v .

Output: Every node v learns $d(v)$, which equals the exact distance $d_w(s, v)$.

- 1: $k \leftarrow n^{3/4}D^{-3/4}$, $\ell \leftarrow n^{1/2}D^{1/2}$, $q \leftarrow n^{1/4}D^{-1/4}$.
 - 2: **for** $\lfloor \log_2 \Lambda \rfloor + 1$ iterations of **ScalingFramework** **do**
 - 3: $\tilde{d} \leftarrow \ell\text{-AdditiveErrorSSSP}(s, G, k)$
 - 4: Run **SmallWeightSSSP** $(s, G, \tilde{d}, k, \ell, q)$
 - 5: Let d be the distances output by **ScalingFramework**.
-

Algorithm 8 $\text{SmallWeightVirtualSSSP}(s, G' = (V', E', w'))$

Input: Every node v' knows the weight and direction of each virtual edge $e' \in E'$ incident to v' . It is promised that $d_{w'}(s, t') \leq |V'| - 1$ for all nodes t' .

Output: Every node v' learns $d(v')$, which equals the exact distance $d_{w'}(s, v')$.

- 1: $d(s) \leftarrow 0$, $d(t') \leftarrow \infty$ for all $t' \neq s$.
 - 2: $i \leftarrow 0$, $k \leftarrow |V'|$
 - 3: **while** $i \leq k - 1$ **do**
 - 4: For each node u' , if $d(u') = i$ and u' has not broadcasted yet, broadcast $d(u')$.
 - 5: If node v' has edge (u', v') and node u' broadcasted, update $d(v') \leftarrow \min\{d(v'), d(u') + w'(u', v')\}$
 - 6: If no node broadcasted, $i \leftarrow i + 1$
-

C Improved Virtual SSSP

In this section, we focus on the virtual SSSP problem, with the goal of developing faster virtual SSSP algorithms for large network diameters D , by reducing the dependency on D .

Let graph G be the network graph with n nodes. Assume that we are in a **ScalingFramework** iteration, so the graph G has s -radius at most $n - 1$. Let $G' = (V', E', w')$ be the virtual graph with $n_{V'}$ nodes containing the source $s \in V'$. Note that D is the diameter of the *original* graph, which could be much larger than $n_{V'}$. The running time $\tilde{O}(n_{V'}D)$ of **VirtualSSSP**, presented in the previous section, is near-optimal for small D , i.e., $D = \tilde{O}(1)$. For larger D , the $\tilde{O}(D)$ factor increases the running time substantially. In this section, we present two alternative algorithms for

computing SSSP on the virtual graph for larger values of D . Recall that our task is for every virtual node $v' \in V'$ to compute a value $d(v')$ that equals $d_{w'}(s, v')$.

C.1 Algorithm 1: Gathering

Algorithm [VirtualSSSPGather](#) Note that the virtual graph G' has size $O(n_{V'}^2)$, so one strategy is to broadcast the entire graph to all nodes in $O(n_{V'}^2 + D)$ time using standard pipelining techniques; we call this procedure *gathering* the graph. Then, all virtual nodes can locally compute their distance from s . We call this algorithm [VirtualSSSPGather](#).

Lemma C.1. *The irtual SSSP algorithm [VirtualSSSPGather](#) runs in $O(n_{V'}^2 + D)$ rounds.*

C.2 Algorithm 2: Virtualizing

The virtual SSSP algorithm of this section mimics certain procedures in [Main](#), except modified to work on a virtual graph. To understand the modifications necessary, we first discuss how to transform generic distributed algorithms to work on a virtual graph.

Distributed Algorithms on Virtual Graph First, we explain how to schedule distributed algorithms on virtual graphs, following a natural extension to [Section 2.4](#).

For a moment, suppose that a virtual graph G' is actually a network graph, in that nodes can communicate directly along edges in G' . Let A_1, \dots, A_k be distributed algorithms running independently on this network graph G' . We place an additional restriction on each algorithm: on every round, a node can either *broadcast* the same message to all of its neighbors, or not send anything. That is, a single node is not allowed to send different messages to different neighbors in a single round. Following [Section 2.4](#), let *dilation* denote the maximum running time of any algorithm A_i , and let *broadcasts* denote the maximum number of times a single node broadcasts, summed up over all k algorithms. Formally, define $b_{v,i}$ as the number of times node v broadcasts in algorithm A_i , and define $\text{broadcasts} := \max_v \sum_i b_{v,i}$. Now, revert back to the normal setting, where G' is a virtual graph of a network graph G ; as usual, we do not assume that edges of G' belong to the network. The following claim relates the performances of A_1, \dots, A_k on the hypothetical network graph G' to their performances on the actual network graph G .

Claim C.2. *Let A_1, \dots, A_k be independent, distributed algorithms on $G' = (V', E')$ in which nodes are only allowed to broadcast. Suppose that G' is a virtual graph of a network graph G with diameter D . Then, we can run algorithms A_1, \dots, A_k in G in time $O(\text{dilation} \cdot D + \text{broadcasts} \cdot n)$.*

Proof. Let $b_{v,i,t} \in \{0, 1\}$ denote whether or not node v broadcasts on step t of algorithm A_i . We can simulate round t of each algorithm A_i in $O(D + \sum_{v,i} b_{v,i,t})$ steps by standard pipelining techniques. Since we need to simulate *dilation* rounds in total, the total running time is $O(\text{dilation} \cdot D + \sum_{v,i,t} b_{v,i,t})$. The latter term is the total number of broadcasts over all nodes and all algorithms, which is at most $n \cdot \text{broadcasts}$. \square

Whenever we transform a set of distributed algorithms to run on a virtual graph, we say we *virtualize* the algorithms.

C.2.1 Virtualizing Previous Tools

Here, we discuss how to virtualize the scaling framework, $(1 + \epsilon)$ -multiplicative approximate SSSP, and the [ShortRange](#) algorithm to work on virtual graphs. The generic process is the same for all three. First, we modify the algorithm so that, at each round, every node can either broadcast a message to all other nodes, or remain silent. Then, we replace each round t of the original algorithm

with a pipelined broadcast, which takes $O(b_t + D)$ time, where b_t is the number of nodes that broadcast on round t . When virtualizing multiple algorithms in parallel, we invoke [Claim C.2](#).

Scaling Framework To update edge weights on each iteration of [ScalingFramework](#), it is enough for every virtual node to broadcast its computed distance, so that all nodes can recompute the weights of their edges. By standard pipelining techniques, this takes $O(n_{V'} + D)$ time per iteration. Since every virtual node has distance at most $n - 1$ from source s , the scaling framework needs just $O(\log n)$ iterations, resulting in $\tilde{O}(n_{V'} + D)$ rounds spent reweighting edges.

ShortRange Observe that, on each round of [ShortRange](#), every node always communicates by sending the same message to its neighbors. Therefore, we can trivially modify the algorithm to run in $O(\ell q + h)$ rounds with $O(h/q)$ broadcasts per node. If we run k [ShortRange](#) algorithms in parallel, each node broadcasts $O(hk/q + k)$ times, so applying [Claim C.2](#) gives the following.

Lemma C.3. *We can simulate k independent runs of the [ShortRange](#) algorithm on a virtual graph with $n_{V'}$ nodes in time $O((\ell q + h) \cdot D + (hk/q + k) \cdot n_{V'})$.*

(1+ ϵ)-Approximate SSSP Consider the problem of computing $(1+\epsilon)$ -multiplicative approximate distances on a virtual graph, for some $\epsilon \leq 1$. The algorithm of [Lemma A.2](#) first runs k independent copies of [BoundedHopSSSP](#) for some parameter k , each of which can be modified to take $\tilde{O}(\epsilon^{-1}n_{V'}/k)$ rounds with $\tilde{O}(1)$ broadcasts per node. The dilation is $\tilde{O}(\epsilon^{-1}n_{V'}/k)$ and each node broadcasts $\tilde{O}(k)$ times total, so by [Claim C.2](#), everything can be pipelined to run in $\tilde{O}(\epsilon^{-1}n_{V'}/k \cdot D + kn_{V'})$ time. Next, the algorithm faces another virtual SSSP instance on k nodes, which it solves using [VirtualSSSPGather](#) in $O(k^2 + D)$ time. Finally, the extension part of the algorithm also runs k independent copies of [BoundedHopSSSP](#), and the analysis is identical. Thus, the total running time to compute $(1 + \epsilon)$ -approximate distances from s is

$$\tilde{O}\left(\frac{n_{V'}D}{\epsilon k} + kn_{V'} + k^2 + D\right).$$

Setting $k := \epsilon^{-1/2}D^{1/2}$ gives the following. We can assume that $k \leq n_{V'}$, since otherwise, the $kn_{V'} + D$ factor makes the running time at least $\Omega(n_{V'}^2 + D)$, and we are better off running [VirtualSSSPGather](#) instead.

Lemma C.4. *We can compute $(1 + \epsilon)$ -multiplicative approximate distances on a virtual graph with $n_{V'}$ nodes in time $\tilde{O}(\epsilon^{-1/2}n_{V'}D^{1/2})$.*

C.2.2 Warm-up: Nonrecursive Virtual SSSP

Here, we describe the virtual SSSP algorithm [VirtualSSSPNonrecursive](#) that uses the virtualized scaling framework and [ShortRange](#) algorithm. It runs in $\tilde{O}(n_{V'}r^{1/2}D^{1/2} + D)$ time on any virtual graph with s -radius at most r . Recall that, after running [ScalingFramework](#), r can be as large as $n_{V'} - 1$, giving a running time of $\tilde{O}(n_{V'}^{3/2}D^{1/2} + D)$. Then, in [Appendix C.2.3](#), we will make this algorithm recursive and further improve the running time.

Virtualized Scaling Framework The algorithm first applies the virtualized [ScalingFramework](#) to the virtual graph G' on $n_{V'}$ nodes, reducing the task to computing $O(\log n)$ iterations of virtual SSSP on a graph with s -radius at most $n_{V'} - 1$. As discussed before, the total time spent reweighting edges in the virtualized [ScalingFramework](#) is $\tilde{O}(n_{V'} + D)$.

Virtualized ShortRange The algorithm first samples a set of nodes $V'' \subseteq V'$, where every virtual node in V' is sampled with probability $k/n_{V'}$, for some parameter k , except that $s \in V''$ with probability 1. We run the virtualized **ShortRange** with parameters $h := O(n_{V'} \log n_{V'}/k)$ and $\ell := r$ from each node in V'' , taking $O((\ell q + h) \cdot D + (hk/q + k) \cdot n_{V'})$ time with a free parameter $q \geq 1$, by **Lemma C.3**.

Then, following notation of line 3 of **SmallWeightSSSP**, we build a graph $G'' = (V'', E'')$, where for each pair $(u'', v'') \in V'' \times V''$ with $d_{u''}(v'') < \infty$, add to E'' a directed edge (u'', v'') of weight $d_{u''}(v'')$ that is known to node v'' . We compute distances in G'' by gathering the graph in $O(k^2 + D)$ time, so that each node in V'' can locally compute its correct distance. Finally, all nodes in V'' broadcast their distances, allowing each remaining node $v' \in V'$ to compute its distance $d(v') = \min_{u'' \in V''} (d(u'') + d_{u''}(v'))$. We next show that, w.h.p., $d(v')$ is the correct distance for each $v' \in V'$.

Consider one iteration of **ScalingFramework**, and fix a node $t' \in V'$ in the above algorithm. By **Lemma 2.4**, there is a shortest $s \rightarrow t'$ path P in G' with a node in V'' within every consecutive h nodes on the path. In other words, consecutive nodes in $V'' \cap P$ are within h hops from each other on the path. Since G' has s -radius at most r , we have $w'(P) \leq r$, so these consecutive nodes are also within distance $r = \ell$. Hence, by the guarantees of **ShortRange**, for every two consecutive nodes $u'', v'' \in V''$, $d_{u''}(v'')$ is the correct distance $d_{w'}(u'', v'')$, and there is an edge $(u'', v'') \in E''$ with weight $d_{w'}(u'', v'')$. Therefore, after computing SSSP on G'' , all nodes in $V'' \cap P$ know their correct distance. In particular, let t'' be the last node on the path that is in V'' . Since t is within h hops and distance ℓ from t'' along P , we also have $d_{t''}(t') = d_{w'}(t'', t')$, so

$$d_{w'}(s, t') = d_{w'}(s, t'') + d_{w'}(t'', t') = d(t'') + d_{t''}(t').$$

Since node t' computes $d(t') = \min_{u'' \in V''} (d(u'') + d_{u''}(t'))$ and the distances $d(v')$ can never be a strict underestimate, we have $d(t') = d_{w'}(s, t')$, as desired.

Finally, we analyze the running time. Overall, the total number of rounds is

$$O\left((\ell q + h) \cdot D + \frac{hkn_{V'}}{q} + kn_{V'} + k^2 + D\right) = \tilde{O}\left(rqD + \frac{n_{V'}D}{k} + \frac{n_{V'}^2}{q} + kn_{V'} + D\right),$$

where we recall that $\ell = r$, $h = \tilde{O}(n_{V'}/k)$, and $kn_{V'} \geq k^2$. Setting $k := D^{1/2}$ and $q := n_{V'}r^{-1/2}D^{-1/2}$ gives running time $\tilde{O}(n_{V'}r^{1/2}D^{1/2} + D)$, as desired. Note that, in the event that $k > n_{V'}$, we have $D > n_{V'}^2$, so we are better off running **VirtualSSSPGather** on G' in the optimal $O(n_{V'}^2 + D) = O(D)$ time. Also, if $q < 1$, the $n_{V'}^2/q$ factor becomes at least $n_{V'}^2$ and again, we are better off running **VirtualSSSPGather**. Thus, we can ignore these two corner cases.

Therefore, we have proved the following.

Lemma C.5. *The algorithm **VirtualSSSPNonrecursive** computes exact distances on a virtual graph with $n_{V'}$ nodes and s -radius at most r in time $\tilde{O}(n_{V'}r^{1/2}D^{1/2} + D)$.*

Algorithm 9 `VirtualSSSPNonrecursive`($s, G' = (V', E', w'), k, q$)

Input: Every node $v' \in V'$ knows the weight and direction of every edge (u', v') in E' . We have the promise that $d_{w'}(s, t') \leq r, \forall t' \in V'$.

Output: Every node v learns $d(v)$, which equals the exact distance $d_w(s, v)$.

- 1: Sample a set of nodes $V'' \subseteq V'$. Every node joins the set with probability k/n , except for s , which joins with probability 1.
 - 2: Set $h \leftarrow \Theta(n_{V'} \log n_{V'}/k)$ and $\ell \leftarrow r$. For each $s'' \in V''$, let $d_{s''}$ be the result of running a virtualized `ShortRange`(s'', G, h, ℓ, q)
 - 3: Build a graph $G'' = (V'', E'')$, where for each pair $(u'', v'') \in V'' \times V''$ with $d_{u''}(v'') < \infty$, add to E'' a directed edge (u'', v'') of weight $d_{u''}(v'')$ that is known to node v'' .
 - 4: Run `VirtualSSSPGather` on the virtual graph G'' , so that each node $v'' \in V''$ knows its distance $d(v'')$ in G'' .
 - 5: Every node $u'' \in V''$ broadcasts its distance $d(u'')$ to all other nodes, and every node $v' \in V' \setminus V''$ computes $d(v') \leftarrow \min_{u'' \in V''} (d(u'') + d_{u''}(v'))$
-

C.2.3 Recursive Virtual SSSP

In this section, we extend `VirtualSSSPNonrecursive` to work recursively, arriving at our final algorithm `VirtualSSSPRecursive`. Our goal is to compute virtual SSSP in $\tilde{O}(n_{V'} r^\epsilon D^{1/2} + r^{1-2\epsilon} D)$ time for any constant $\epsilon \in (0, 1/2]$, on a graph with s -radius at most r . Note that the case $\epsilon = 1/2$ is covered by `VirtualSSSPNonrecursive` and [Lemma C.5](#).

Recursion For ease of notation, define a virtual SSSP algorithm to be ϵ -good if, for any integers $n_{V'}$ and r , and on any virtual graph with $n_{V'}$ nodes and s -radius r , the algorithm runs in $\tilde{O}(n_{V'} r^\epsilon D^{1/2} + r^{1-2\epsilon} D)$ time. Our ϵ -good algorithm will make a recursive call to an ϵ' -good algorithm, for some $\epsilon' > \epsilon$. Eventually, the value of ϵ in the recursion goes up to $1/2$, at which point the algorithm of [Lemma C.5](#) serves as the base case and performs the job.

Additive Approximation Let G' be a virtual graph with $n_{V'}$ nodes and s -radius at most r . The algorithm begins by computing an ℓ -additive approximate SSSP, for some $\ell \leq r$. Note that, since the s -radius is bounded by r this time, a $(1 + \epsilon)$ -multiplicative approximation for $\epsilon := \ell/r$ suffices. By [Lemma C.4](#), this step has round complexity of

$$\tilde{O}\left(\frac{n_{V'} D^{1/2}}{\epsilon^{1/2}}\right) = \tilde{O}\left(\frac{n_{V'} r^{1/2} D^{1/2}}{\ell^{1/2}}\right).$$

Bucketing Similarly to `SmallWeightSSSP`, the algorithm defines buckets so that bucket V'_i contains all virtual nodes whose distance from s is in the range $[i\ell, (i+1)\ell)$. The key difference is that, since the s -radius is at most r this time, we only need $O(r/\ell)$ buckets. Then, use the ℓ -additive approximation \tilde{d} to group each virtual node $v' \in V'$ into at most two buckets, according to the rule $v' \in V'_i \iff \tilde{d}(s, v') \in [i\ell, (i+2)\ell)$. From now on, let $n_i := |V'_i|$ be the number of nodes in bucket n_i , so that $\sum_i n_i \leq 2n_{V'}$.

Graph on Each Bucket The algorithm then processes the buckets V'_i in increasing order of i . For each bucket V'_i , we assume the following invariant right before processing bucket V'_i , which is satisfied vacuously for the first iteration $i = 0$:

- **(Invariant):** Right before we process bucket V'_i , all virtual nodes whose distance from s is less than $i\ell$ has broadcast its correct distance to all other virtual nodes.

Every virtual node $v' \in V'_i$ first computes its best distance $d_i(v')$ based on distance broadcasts from its neighbors in G' on iterations before i . Now build a virtual graph $\tilde{G}'_i = (\tilde{V}'_i, \tilde{E}'_i)$ as follows: take the subgraph of G' induced by V'_i , add s to the graph, and direct an edge from s to each virtual node $v' \in V'_i$ with edge weight $d_i(v') - (i-1)\ell$. Observe that all edge weights are nonnegative, since $d_i(v') \geq d_{G'}(s, v') \geq \tilde{d}_{G'}(s, v') - \ell \geq i\ell - \ell$. We now claim the following:

Corollary C.6. *For all virtual nodes $t' \in V'_i$, $d_{\tilde{G}'_i}(s, t') \geq d_{G'}(s, t') - (i-1)\ell$. If node t' also satisfies $d_{G'}(s, t') \in [i\ell, (i+1)\ell)$, then we have $d_{\tilde{G}'_i}(s, t') = d_{G'}(s, t') - (i-1)\ell$.*

Proof. It is easy to see that $d_{\tilde{G}'_i}(s, t') \geq d_{G'}(s, t') - (i-1)\ell$, since edges do not get shorter except for the added edges from s , so we focus on the other direction.

Fix a virtual node $t' \in V'_i$. We know that $d_{G'}(s, t') \leq \tilde{d}(s, t') < (i+2)\ell$. Let P be the shortest $s \rightarrow t'$ path in G' , and let u' be the last virtual node on P satisfying $d_{G'}(s, u') < i\ell$. Then, u' has already broadcast its correct distance, and its next node v' on P has computed its correct distance $d_{G'}(s, v')$ in G' . In particular, there is an edge from s to v' with weight $d_{G'}(s, v') - (i-1)\ell$ in \tilde{G}'_i . Also, all nodes from v' to t' on P are in V'_i . Therefore, the subpath of P from v' to t' , with s appended at the front, has length exactly $d_{G'}(s, t') - (i-1)\ell$ in \tilde{G}'_i . This proves $d_{\tilde{G}'_i}(s, t') \leq d_{G'}(s, t') - (i-1)\ell$. \square

In particular, all virtual nodes $v' \in V$ satisfying $d_{G'}(s, v') \in [i\ell, (i+1)\ell)$ have distance at most 2ℓ from s in \tilde{G}'_i . Next, the algorithm computes the correct distance from s to these nodes in \tilde{G}'_i through a recursive call.

Recursive SSSP on Each Bucket The ϵ -good algorithm computes a virtual SSSP on \tilde{G}_i recursively by calling an ϵ' -good algorithm, for some $\epsilon' > \epsilon$; the relation between ϵ' and ϵ will be determined later. \tilde{G}_i is a graph with n_i nodes and s -radius at most 2ℓ , so the recursive call takes

$$\tilde{O}(n_i \ell^{\epsilon'} D^{1/2} + \ell^{1-2\epsilon'} D)$$

time. After the recursive call, each virtual node $v' \in V'_i$ takes its computed distance $\tilde{d}_i(v')$ and adds $(i-1)\ell$ to it, and broadcasts as its distance estimate for $d_{G'}(s, v')$ to all other nodes. In particular, all virtual nodes $v' \in V'$ with $d_{G'}(s, v') \in [i\ell, (i+1)\ell)$ broadcast their correct distance, satisfying the Invariant. Note that broadcasting the distances takes $n_i + D$ time, which does not affect the running time asymptotics.

Running Time Summing up the running time of the approximate SSSP and the recursive virtual SSSP calls in each bucket, we obtain:

$$\tilde{O}\left(\frac{n_{V'} r^{1/2} D^{1/2}}{\ell^{1/2}}\right) + \sum_{i=1}^{O(r/\ell)} \tilde{O}\left(n_i \ell^{\epsilon'} D^{1/2} + \ell^{1-2\epsilon'} D\right) = \tilde{O}\left(\frac{n_{V'} r^{1/2} D^{1/2}}{\ell^{1/2}} + n_{V'} \ell^{\epsilon'} D^{1/2} + \frac{rD}{\ell^{2\epsilon'}}\right),$$

using the fact that $\sum_i n_i \leq 2n_{V'}$. Setting $\ell := r^{1/(1+2\epsilon')}$ gives running time

$$\tilde{O}(n_{V'} r^{\epsilon'/(1+2\epsilon')} D^{1/2} + r^{1-2\epsilon'/(1+2\epsilon')} D).$$

In other words, the algorithm is ϵ -good for $\epsilon := \epsilon'/(1+2\epsilon')$.

Moreover, setting $\ell := r^{(1+\delta)/(1+2\epsilon')}$ for $\delta < 2\epsilon'$ gives running time

$$\tilde{O}(n_{V'} r^{(1-(1+\delta)/(1+2\epsilon'))/2} D^{1/2} + r^{1-2\epsilon'(1+\delta)/(1+2\epsilon')} D),$$

and the exponent of r in the first term satisfies $(1 - (1+\delta)/(1+2\epsilon'))/2 \leq \epsilon'(1+\delta)/(1+2\epsilon')$, so the algorithm is $\epsilon'(1+\delta)/(1+2\epsilon')$ -good. Therefore, by choosing an appropriate δ , the algorithm can become ϵ -good for any ϵ satisfying

$$\epsilon'/(1+2\epsilon') \leq \epsilon < \epsilon'.$$

Solving the Recursion Finally, we prove the desired running time of the algorithm.

Lemma C.7. *For any constant $\epsilon \in (0, 1/2]$, there is a virtual SSSP algorithm running in $\tilde{O}(n_{V'} r^\epsilon D^{1/2} + r^{1-2\epsilon} D)$ rounds on a virtual graph with $n_{V'}$ nodes and s -radius at most r .*

Proof. We prove by induction on $\alpha \in \mathbb{N}$ that there is an ϵ -good virtual SSSP algorithm for all $\epsilon \in [1/(2\alpha), 1/2]$. The base case $\alpha = 1$ follows by [Lemma C.5](#).

Now assume that there is an ϵ' -good virtual SSSP algorithm for all $\epsilon' \in [1/(2\alpha), 1/2]$. In particular, the statement is true for $\epsilon' = 1/(2\alpha)$. By our previous observation, in one recursive step, we obtain an ϵ -good algorithm for any ϵ satisfying $\epsilon'/(1 + 2\epsilon') \leq \epsilon < \epsilon'$, which comes out to $\epsilon \in [1/(2\alpha+2), 1/(2\alpha))$. Hence, there is an ϵ -good virtual SSSP algorithm for all $\epsilon \in [1/(2\alpha+2), 1/2]$, completing the induction. \square

Observe that the above claim only holds for constant values of ϵ , since the number of virtual nodes can double upon a recursive call. That is, we need to maintain a constant number of recursion levels.

Final Running Time Finally, due to the virtualized [ScalingFramework](#), we can assume that $r \leq n_{V'} - 1$ initially, giving our main result in this section.

Theorem C.8. *For any constant $\epsilon \in (0, 1/2]$, there is a virtual SSSP algorithm with round complexity of $\tilde{O}(n_{V'}^{1+\epsilon} D^{1/2} + n_{V'}^{1-2\epsilon} D)$.*

Algorithm 10 [VirtualSSSPRecursive](#)($s, G' = (V', E', w'), \epsilon$)

Input: Every node $v' \in V'$ knows the weight and direction of every edge (u', v') in E' . We have the promise that $d_{w'}(s, t') \leq r, \forall t' \in V'$.

Output: Every node v learns $d(v)$, which equals the exact distance $d_w(s, v)$.

- 1: Let ϵ' be the solution to $\epsilon = \epsilon'/(1 + 2\epsilon')$, or $\epsilon' \leftarrow 1/2$ if the solution is greater than $1/2$.
 - 2: Let δ be the solution to $\epsilon = \epsilon'(1 + \delta)/(1 + 2\epsilon')$.
 - 3: $\ell \leftarrow r^{(1+\delta)/(1+2\epsilon')}$, $k \leftarrow (\ell/r)^{-1/2} D^{1/2}$
 - 4: $\tilde{d} \leftarrow \ell$ -AdditiveErrorVirtualSSSP(s, G', k)
 - 5: Define V'_i as $\{v' \in V' : \tilde{d}(s, v') \in [i\ell, (i+2)\ell)\}$.
 - 6: Initialize distances $d(s) \leftarrow 0$, $d(v) \leftarrow \infty$ for all $v \in V \setminus s$.
 - 7: **for** $i = 0, \dots, \lfloor r/\ell \rfloor$ **do** // Iterate over buckets $V_0, \dots, V_{\lfloor r/\ell \rfloor}$.
 - 8: Build the graph $\tilde{G}'_i = (\tilde{V}'_i, \tilde{E}'_i)$ as follows. Start with the graph G' induced by V'_i . Then, add node s to \tilde{V}'_i if necessary, and for each $v' \in V'_i$, add to \tilde{E}'_i a directed edge (s, v') of weight $d(v') - (i-1)\ell$ that is known to node v' .
 - 9: **if** $\epsilon' = 1/2$ **then** // Base case.
 - 10: $k \leftarrow D^{1/2}$, $q \leftarrow |\tilde{V}'_i| r^{-1/2} D^{-1/2}$
 - 11: Run [VirtualSSSPNonrecursive](#)(s, \tilde{G}'_i, k, q), and let \tilde{d}'_i be the resulting distance that each node in \tilde{V}'_i learns.
 - 12: **else** // Recursive step.
 - 13: Run [VirtualSSSPRecursive](#)($s, \tilde{G}'_i, \epsilon'$), and let \tilde{d}'_i be the resulting distance that each node in \tilde{V}'_i learns.
 - 14: Every node $v' \in V'_i$ broadcasts $\tilde{d}'_i(v')$ to all other virtual nodes. // These broadcasts are pipelined along a BFS tree.
 - 15: Every node $v' \in V'$ sets $d(v') = \min\{d(v'), d(u') + w'(u', v')\}$ for each $\tilde{d}'_i(u')$ received from neighbor u' in G' .
-

C.3 Running Time Improvements

From the proof of [Lemma 3.3](#), observe that [SmallWeightSSSP](#) runs in

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k}\right)$$

time, plus the time it takes to solve a virtual SSSP instance on k virtual nodes. If we use [VirtualSSSPGather](#) of [Lemma C.1](#) that runs in $O(k^2 + D)$ rounds, we get running time

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + k^2 + D\right);$$

setting $k := n^{3/7}$, $\ell := n^{5/7}$, and $q := n^{1/7}$ gives total running time $\tilde{O}(n^{6/7} + D)$.

If we use the algorithm of [Theorem C.8](#) that runs in $\tilde{O}(k^{1+\epsilon}D^{1/2} + k^{1-2\epsilon}D)$ rounds, we get running time

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + k^{1+\epsilon}D^{1/2} + k^{1-2\epsilon}D\right),$$

as long as $\epsilon \in [\Omega(1), 1/2]$. To optimize this expression, we first set ϵ to be the solution to the equation $k^{3\epsilon} = D^{1/2}$ so that, for any value of k ,

$$k^{1+\epsilon}D^{1/2} + k^{1-2\epsilon}D = kD^{2/3}.$$

Then, set $k := n^{3/4}D^{-1/2}$, $\ell := n^{1/2}D^{1/3}$, and $q := n^{1/4}D^{-1/6}$ to obtain running time $\tilde{O}(n^{3/4}D^{1/6})$, as long as $\epsilon \in [\Omega(1), 1/2]$.

We first restrict ourselves to the case $D = n^{\Omega(1)}$, which implies $\epsilon = \Omega(1)$. As for the constraint $\epsilon \leq 1/2$, that happens precisely when

$$D^{1/2} = k^{3\epsilon} \leq k^{3/2} = (n^{3/4}D^{-1/2})^{3/2} \iff D \leq n^{9/10}.$$

Thus, we have an $\tilde{O}(n^{3/4}D^{1/6})$ algorithm when $n^{\Omega(1)} \leq D \leq n^{9/10}$, and an $\tilde{O}(n^{6/7} + D)$ algorithm for any D . Note that when $D \geq n^{6/7}$, we are always better off running the $\tilde{O}(n^{6/7} + D)$ algorithm, so the upper bound of $D \leq n^{9/10}$ for the virtualizing SSSP case does not hinder us overall. Hence, the optimal algorithm of the two has a running time expressible as

$$\tilde{O}\left(\min\{n^{3/4}D^{1/6}, n^{6/7}\} + D\right).$$

For the case when $D = n^{o(1)}$, the $\tilde{O}(n^{3/4}D^{1/4})$ algorithm of [Theorem 1.1](#) runs in $\tilde{O}(n^{3/4+o(1)})$ time. Altogether, by choosing the best algorithm of the three, we obtain the running time

$$\tilde{O}\left(n^{3/4+o(1)} + \min\{n^{3/4}D^{1/6}, n^{6/7}\} + D\right),$$

proving [Theorem 1.2](#).

D Multiple sources

Consider now the multiple source SSSP problem: given κ sources, we want each node $v \in V$ to know its exact distance from each source. In this section, we modify our single-source SSSP algorithm to obtain efficient algorithms for the κ -sources problem.

Our high-level ideas are the same: we run the scaling framework $O(\log \Lambda)$ times, and we solve each iteration for all κ sources in parallel. To elaborate, we first construct κ graphs G_s , one for

each source s , so that every graph G_s has s -radius at most $n - 1$. In κ rounds, every node can know its incident edges in each graph G_s . We then run `SmallWeightSSSP`(s, G_s, k, ℓ, q) from each source s in parallel, once again using [Gha15] to schedule our algorithms, so that the entire algorithm runs in $\tilde{O}((\text{dilation} + \kappa \cdot \text{congestion}) \cdot \log \Lambda)$ time. where dilation and congestion are parameters of `SmallWeightSSSP`(s, G_s, k, ℓ, q) for a single source s .

Since we have multiple virtual SSSP algorithms, we will break the running time inside and outside the virtual SSSP computation.

Outside Virtual SSSP The dilation $\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k}\right)$ of `SmallWeightSSSP` outside of the virtual SSSP algorithm follows from Lemma 3.3. The following lemma analyzes the congestion.

Lemma D.1. *In one execution of `SmallWeightSSSP` outside of virtual SSSP, $\tilde{O}(n/q + k + n/k)$ messages are passed along each edge.*

Proof. By Lemma 2.6, ℓ -AdditiveErrorSSSP has congestion $\tilde{O}(k)$. Again, we use the same value of k for ℓ -AdditiveErrorSSSP. Also, `ShortRange` from all virtual nodes takes $\tilde{O}(n/q)$ congestion, following the proof of Lemma 3.3.

We now argue about the congestion of `Extend`. For an edge (u, v) , let i be the first iteration when $d(u)$ is sent from u to v . It follows that the current estimate $d(u)$ satisfies $d(u) + w(u, v) \in [i\ell, (i+1)\ell)$. Since estimates only become lower over time, we still have $d(u) < (i+1)\ell$ at the end of iteration i . Therefore, no message is sent from u to v in future iterations of `Extend`. It follows that u only sends messages to v in one iteration, which is $O(h)$ messages. We can repeat the same argument to bound the number of messages from v to u , so the congestion along each edge is $O(h) = \tilde{O}(n/k)$. \square

Hence, by the $\tilde{O}(\text{dilation} + \kappa \cdot \text{congestion})$ formula, the total running time of `SmallWeightSSSP` outside virtual SSSP is

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + \kappa \left(\frac{n}{q} + k + \frac{n}{k}\right)\right).$$

Inside Virtual SSSP Now we consider the changes to virtual SSSP algorithms. Again, let $n_{V'}$ be the number of nodes in each of the κ virtual graphs.

First, consider the original `VirtualSSSP` algorithm of Lemma 3.4. In `VirtualSSSP`, the total number of broadcasts is $O(\kappa n_{V'})$ over all κ algorithms, so the congestion is $\tilde{O}(n_{V'})$. Adding this onto the dilation, we obtain running time

$$\tilde{O}(\kappa n_{V'} + n_{V'} D).$$

Now consider the two improved algorithms of Appendix C. The gathering algorithm gathers κ graphs of size $O(n_{V'}^2)$, so it takes

$$O(\kappa n_{V'}^2 + D)$$

rounds total.

For the virtualizing algorithm, we first modify the virtualized procedures to run κ iterations in parallel. The only difference in the running time is the value of `broadcasts`, which increases by a factor of κ . We again use Claim C.2 and optimize the parameters.

1. Reweighting edges in the scaling framework now takes $\tilde{O}(\kappa n_{V'} + D)$ rounds.
2. The virtualized $(1 + \epsilon)$ -multiplicative approximate SSSP now takes $\tilde{O}(n_{V'} D / (\epsilon k) + \kappa \kappa n_{V'} + k^2 + D)$ rounds. Setting $k := \kappa^{-1/2} \epsilon^{-1/2} D^{1/2}$ gives running time $\tilde{O}(\kappa^{1/2} \epsilon^{-1/2} n_{V'} D^{1/2})$. Here, to ensure that $k \geq 1$, we impose the restriction $\kappa < D$. Also, if $k > n_{V'}$, then the $\kappa \kappa n_{V'} + D$ factor makes the running time at least $\Omega(\kappa n_{V'}^2 + D)$, and we are better off gathering the graph.

3. The virtualized **ShortRange** now takes $\tilde{O}(rqD + n_{V'}D/k + \kappa n_{V'}^2/q + \kappa n_{V'}k + D)$ rounds. We can optimize $k := \kappa^{-1/2}D^{1/2}$ and $q := \kappa^{1/2}n_{V'}r^{-1/2}D^{-1/2}$ to get running time $\tilde{O}(\kappa^{1/2}n_{V'}r^{1/2}D^{1/2} + D)$. Again, we assume that $\kappa < D$. Also, if $q < 1$, then the $\kappa n_{V'}^2/q + D$ factor is at least $\Omega(\kappa n_{V'}^2 + D)$.

That is, both running times increase by a factor of $\kappa^{1/2}$. We can work through the recursion in **Lemma C.7** with these updated values in the same manner and obtain the following.

Lemma D.2. *For any $\epsilon \in (0, 1/2]$, there is an algorithm that computes virtual SSSP on κ graphs in parallel, each with $n_{V'}$ nodes and s -radius at most r . It runs in $\tilde{O}(\kappa^{1/2}n_{V'}r^\epsilon D^{1/2} + r^{1-2\epsilon}D)$ rounds.*

Since $r = n_{V'} - 1$ by the scaling framework, the virtualizing algorithm solves κ virtual SSSPs in

$$\tilde{O}(\kappa^{1/2}n_{V'}^{1+\epsilon}D^{1/2} + n_{V'}^{1-2\epsilon}D)$$

rounds.

Lastly, we optimize for the parameters in each of the virtual SSSP algorithms. We have three cases, some of which impose constraints on κ and D .

1. $\kappa \geq Dn^{-o(1)}$. In this case, we run the original **VirtualSSSP** in $\tilde{O}(\kappa k + kD) = \tilde{O}(\kappa k)$ rounds, for a total running time of

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + \kappa\left(\frac{n}{q} + k + \frac{n}{k}\right) + kD\right).$$

Setting $k := n^{3/4}\kappa^{-1/2}$, $\ell := n^{1/2}$, and $q := n^{1/2}\kappa^{-1/2}$ gives a final time of

$$\tilde{O}(\kappa^{1/2}n^{3/4+o(1)}).$$

2. The gathering virtual SSSP algorithm in $\tilde{O}(\kappa k^2 + D)$ rounds, for a total running time of

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + \kappa\left(\frac{n}{q} + k + \frac{n}{k}\right) + \kappa k^2 + D\right).$$

Setting $k := \kappa^{-2/7}n^{3/7}$, $\ell := \kappa^{-1/2}n^{5/7}$, and $q := \kappa^{4/7}n^{1/7}$ gives a final time of

$$\tilde{O}(\kappa^{3/7}n^{6/7} + D).$$

3. $\kappa \leq Dn^{-\Omega(1)}$ and $D \leq \kappa^{2/5}n^{9/10}$. In this case, we use the virtualizing algorithm in $\tilde{O}(\kappa^{1/2}k^{1+\epsilon}D^{1/2} + k^{1-2\epsilon}D)$ rounds. We first set ϵ to satisfy $\kappa^{1/2}k^{1+\epsilon}D^{1/2} = k^{1-2\epsilon}D \iff k^{3\epsilon} = \kappa^{-1/2}D^{1/2}$, so that the algorithm runs in $\tilde{O}(\kappa^{1/3}n_{V'}D^{2/3})$ rounds. Note that $\epsilon \in [\Omega(1), 1/2]$ in this case. The total running time becomes

$$\tilde{O}\left(\ell q + \frac{n}{q} + \frac{n^2}{\ell k} + \kappa\left(\frac{n}{q} + k + \frac{n}{k}\right) + \kappa^{1/3}kD^{2/3}\right).$$

Setting $k := n^{3/4}D^{-1/2}$, $\ell := \kappa^{-1/3}n^{1/2}D^{1/3}$, and $q := \kappa^{2/3}n^{1/4}D^{-1/6}$ gives a final time of

$$\tilde{O}(\kappa^{1/3}n^{3/4}D^{1/6}).$$

We now verify the constraints on ϵ . Since $k^{3\epsilon} = \kappa^{-1/2}D^{1/2} = n^{\Omega(1)}$ by assumption, we have $\epsilon = \Omega(1)$. Also,

$$\epsilon \leq 1/2 \iff \kappa^{-1/2}D^{1/2} = k^{3\epsilon} \leq k^{3/2} = (n^{3/4}D^{-1/2})^{3/2} \iff D \leq \kappa^{2/5}n^{9/10},$$

which is precisely our constraint.

Note that when $D > \kappa^{2/5}n^{9/10}$, we also have $D > \kappa^{3/7}n^{6/7}$, so we are better off running the gathering virtual SSSP instead, giving an optimal $\tilde{O}(\kappa^{3/7}n^{6/7} + D) = \tilde{O}(D)$ rounds total. Therefore, the $D \leq \kappa^{2/5}n^{9/10}$ restriction does not affect us overall.

Altogether, the three cases combined produce an algorithm that runs in time

$$\tilde{O}\left(\min\left\{\kappa^{1/2}n^{3/4+o(1)} + \kappa^{1/3}n^{3/4}D^{1/6}, \kappa^{3/7}n^{6/7}\right\} + D\right),$$

proving [Theorem 1.3](#).