Fully dynamic all-pairs shortest paths with worst-case update-time revisited*

Ittai Abraham† Shiri Chechik‡ Sebastian Krinninger§

Abstract

We revisit the classic problem of dynamically maintaining shortest paths between all pairs of nodes of a directed weighted graph. The allowed updates are insertions and deletions of nodes and their incident edges. We give worst-case guarantees on the time needed to process a single update (in contrast to related results, the update time is *not* amortized over a sequence of updates).

Our main result is a simple randomized algorithm that for any parameter $c > 1$ has a worst-case update time of $O(cn^{2+2/3} \log^{4/3} n)$ and answers distance queries correctly with probability $1 - 1/n^c$, against an adaptive online adversary if the graph contains no negative cycle. The best deterministic algorithm is by Thorup [STOC 2005] with a worst-case update time of $\tilde{O}(n^{2+3/4})$ and assumes non-negative weights. This is the first improvement for this problem for more than a decade. Conceptually, our algorithm shows that randomization along with a more direct approach can provide better bounds.

*To be presented at the Symposium on Discrete Algorithms (SODA) 2017. This work was done in part while the authors were at Microsoft Research Silicon Valley Lab, Mountain View, USA.

†Hebrew University of Jerusalem, Israel.

‡Tel Aviv University, Israel. This research was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1528/15).

§Max Planck Institute for Informatics, Saarland Informatics Campus, Germany. Work done in part while at the University of Vienna, Faculty of Computer Science, Austria, and while at the Simons Institute for the Theory of Computing, Berkeley, USA.

1 Introduction

In the all-pairs shortest paths (APSP) problem we are interested in computing the distance matrix of a given graph. In the *fully dynamic* version of this problem the graph might undergo updates in the form of insertions and deletions of nodes and their incident edges. The goal is to refresh the distance matrix after each such update as quickly as possible. In particular we want algorithms that are more efficient than recomputing the distance matrix from scratch after every update in the graph. The time needed to perform the operations for refreshing the matrix is called update time. Our main result is a fully dynamic APSP algorithm for weighted directed graphs with n nodes. Our algorithm is randomized and answers queries correctly with *high probability* (the probability of error is polynomially small (n^{-c})) against an *adaptive online adversary* [BBK+94]. This type of adversary cannot see the algorithm's random coins and internal data structure, but it may choose arbitrary updates and path queries based on all previous responses of the algorithm.

Theorem 1.1. *For every $c \geq 1$, there is a randomized fully dynamic algorithm for maintaining the distance matrix of a weighted directed graph containing no negative cycles that, with probability at least $1 - 1/n^c$, has a worst-case update time of $O(cn^{2+2/3} \log^{4/3} n)$ and is correct against an adaptive online adversary.*

In the past years (see Section 1.3) there has been significant progress on dynamic shortest path problems. However, obtaining worst-case bounds for the most general setting of arbitrary graphs with weights seems challenging. Despite considerable recent attention to dynamic graph problems, the only result in this model is the deterministic algorithm of Thorup [Tho05] from STOC 2005 that obtained worst-case update time of $\tilde{O}(n^{2+3/4})$ for non-negative weights.[1]

We present the first solution that takes advantage of randomization to improve the worst-case update time from $\tilde{O}(n^{2+3/4})$ to $\tilde{O}(n^{2+2/3})$ and extend it to negative weights. We believe this trade-off (significantly better update time at the cost of negligible (n^{-c}) probability of being wrong) is an important point on the design space of fully dynamic shortest path algorithms. In addition, our solution is arguably simpler than the one of Thorup [Tho05]. The algorithm of Thorup [Tho05] relies on the algorithm of Demetrescu and Italiano [DI04] and it is essentially a sophisticated de-amortization of Demetrescu and Italiano [DI04]. In contrast, both our algorithm and analysis are pretty simple and independent of other sophisticated algorithms. Although our algorithm is not deterministic, its guarantees are stronger than those of many other randomized shortest paths algorithms. Namely, it supports updates by an adaptive online adversary and not just an oblivious adversary. This means that the adversary is allowed to base its updates on to the shortest paths previously returned by the algorithm (but does not directly see the algorithm's internal randomness). The weaker oblivious adversary has to fix its sequence of updates before the algorithm starts.

Our algorithms compute the distance matrix explicitly after every update. In general, this is not required for a fully dynamic algorithm as long as it is able to answer queries for the distances between nodes after each update step. The time needed to perform a single query

[1] We use $\tilde{O}(\cdot)$ to hide polylogarithmic factors in n.

is called query time. Our algorithms have constant query time when asked for the distance between two nodes. They can easily be extended to also output the shortest path connecting two nodes in time proportional to the length of the path at the cost of an additional factor of $\log n$ in the update time (see Section 4.4). Restricting the allowed updates to insertions and deletions of nodes is no loss of generality: insertions and deletions of edges as well as edge weight increases can be simulated by at most two node updates.

1.1 Additional results

We believe that Sankowski's framework for maintaining the matrix adjoint [San04, San05] gives a randomized fully dynamic algorithm for maintaining the distance matrix of an *unweighted* directed graph with a worst-case update time of $\tilde{O}(n^{2+1/2})$. However it seems that this extension is inherently limited to maintaining distances and cannot efficiently be extended to output also the shortest path connecting two nodes in time proportional to the length of the path.

In Section 4.2 we resolve this shortcoming. We show how to extend our scheme to *unweighted* graphs with a worst-case update time of $\tilde{O}(n^{2+1/2})$ and allow to output also the shortest path connecting two nodes in time proportional to the length of the path (see also Section 4.4).

Finally in Section 4.3 we show that our scheme can be extended to a *deterministic* version with *negative weights* and obtain a worst-case update time of $O(n^{2+3/4}\log^{2/3} n)$. This result improves on the best known deterministic result by reducing the logarithmic factors.

1.2 Recomputing from scratch

An alternative approach is to recompute all-pairs shortest paths from scratch on each update. Fully dynamic algorithms (like ours and Thorup's) improve on this approach when the edge weights can be relatively large or when the graph is unweighted (like ours and Sankowski's). In the static setting, Zwick's pseudopolynomial all-pairs shortest paths algorithm [Zwi02] has a running time of $O(n^{2.5302})$ [Gal12] if its input is a directed graph with integer edge weights from $\{-W, \ldots, 0, \ldots, W\}$ such that $W \leq n^{3-\omega}$. However, Zwick's algorithm achieves its superior running time in this regime by using a fast rectangular matrix multiplication algorithm as a subroutine. Large constants in the running times of these algorithms make it worthwhile to find solutions that do not rely on fast matrix multiplication as a subroutine. For arbitrary edge weights, the current fastest algorithm has a running time of $O(n^3/2^{\Omega(\log^{1/2} n)})$ [Wil14, CW16].

1.3 Related work

Unless noted otherwise, the algorithms cited in this section are deterministic and allow insertions and deletions of nodes and their incident edges in directed graphs with non-negative edge weights.

Fully dynamic algorithms. The study of fully dynamic APSP algorithms for general directed graphs was initiated by King [Kin99]. She obtained a fully dynamic algorithm with a pseudopolynomial amortized update time of $O(n^{2+1/2}\sqrt{W \log n})$, where the edge weights

have to be positive integers and W is the largest among the weights. She also presented $(1+\epsilon)$- and $(2+\epsilon)$-approximations (for any positive constant ϵ) with amortized update times of $O(n^2 \log^3(Wn)/\epsilon^2)$ and $O(n^2 \log^2 n / \log\log n)$, respectively.

Later, Demetrescu and Italiano [DI06] obtained an algorithm allowing arbitrary edge weight updates with an amortized update time of $O(n^{2+1/2}\sqrt{S \log^3 n})$, where each edge can assume at most S different real values. Using a new framework for exploiting local properties of shortest paths Demetrescu and Italiano [DI04] obtained an algorithm with an amortized update time of $O(n^2 \log^3 n)$. This result is essentially optimal (up to logarithmic factors) if we demand that the algorithm has to maintain the distance matrix explicitly. Thorup [Tho04] slightly improved this update time to $O(n^2(\log n + \log^2((n+m)/n)))$, thus subsuming previous results on this problem in terms of running time. Subsequently, Thorup [Tho05] developed an algorithm with a worst-case update time of $\tilde{O}(n^{2+3/4})$ by deamortizing [DI06]. There are also two randomized algorithms for unweighted directed graphs with non-constant query time. The first one by Roditty and Zwick has an amortized update time of $\tilde{O}(m\sqrt{n})$ and a query time of $O(n^{3/4})$. The second one by Sankowski [San05] uses fast matrix multiplication as a subroutine and has a worst-case update time of $O(n^{1.932})$ and a query time of $O(n^{1.288})$.

Further results include fully dynamic algorithms for planar [KS98, HKR+97, FR06, ACG12] and undirected graphs [RZ12, Ber09, Ber16, ACT14].

Partially dynamic algorithms. In the partially dynamic model only one type of updates is allowed. The *incremental* model is restricted to insertions and the *decremental* model is restricted to deletions. The amortized update time bounds of partially dynamic algorithms in all known algorithms do not depend on the number of updates performed. Thus it is often convenient to report the *total update time*, which is the sum of the individual update times.

A simple incremental algorithm for inserting a single node can be obtained by running one iteration of the Floyd-Warshall algorithm, which takes time $O(n^2)$, or, total update time $O(n^3)$ for inserting n nodes. Ausiello et al. [AIMS+91] presented an incremental algorithm for inserting edges or decreasing edge weights in integer weighted graphs with a total update time of $O(n^3 W \log(nW))$, where W is the largest edge weight.

In terms of node deletions, the algorithm of Demetrescu and Italiano [DI04] has a total update time of $O(n^3 \log n)$. The fastest decremental algorithms for edge deletions have total update times of randomized $O(n^3 \log^2 n)$ in unweighted graphs [BHS07], or $O(n^3 S \log^3 n)$ in weighted graphs [DI06], where S is the number of different values each edge assumes (edge deletions are implemented by setting the weight of the edge to ∞). If we allow approximate answers, the state of the art is a randomized algorithm by Bernstein [Ber16]: in directed graphs with integer edge weights decremental $(1+\epsilon)$-approximate APSP can be maintained with a total update time of $\tilde{O}(mn \log W/\epsilon)$, where W is the largest edge weight.[2] Related work includes various approximation algorithms for undirected graphs [BHS03, RZ12, BR11, HKN16, AC13, HKN14a, ACT14], in particular also for the single-source shortest paths problem [ES81, BR11, HKN13, HKN14a, HKN14b, BC16].

[2]Equivalently, if the graph has rational edge weights, the total update time is $\tilde{O}(mn \log R/\epsilon)$, where R is the ratio of the largest to the smallest edge weight.

2 Preliminaries

In the rest of this paper we consider a weighted directed graph G undergoing insertions and deletions of nodes and their incident edges. At every insertion of a node its incoming and outgoing edges and their respective weights are specified. We define V and E to be the sets of nodes and edges of G, respectively. We set $n = |V|$ and $m = |E|$. Given a subset $S \subseteq V$ of nodes we denote by $G \setminus S$ the subgraph of G induced by $V \setminus S$. The weight of an edge $(u, v) \in E$ is denoted by $w(u, v)$. We define the length of a path to be the sum of the weight of its edges. The shortest path from s to t is the minimum length path from s to t (or ∞ if no such path exists). The distance from s to t in a graph G is the length of the shortest path from s to t and is denoted by $d_G(s, t)$. A $\leq h$ hop path is a path consisting of at most h edges. The shortest $\leq h$ hop path from s to t is a path with minimum length among all $\leq h$ hop paths from s to t. We denote by $d^h_G(s, t)$ the length of the shortest $\leq h$ hop path from s to t (or ∞ if no such path exists). Note that shortest $\leq h$ hop paths may be different from shortest paths in the case where the shortest paths contain more than h edges.

In our algorithm we use the well-known fact that good hitting sets can be obtained by random sampling. This technique was first used in the context of shortest paths by Ullman and Yannakakis [UY91]. A general lemma can be formulated as follows.

Lemma 2.1. *Let $a \geq 1$, let T be a set of size t and let $S_1, S_2, \ldots, S_k$ be subsets of T of size at least q. Let U be a subset of T that was obtained by choosing each element of T independently with probability $p = \min(x/q, 1)$ where $x = a \ln(kt) + 1$. Then, with high probability (whp), i.e., probability at least $1 - 1/t^a$, the following two properties hold:*

1. *For every $1 \leq i \leq k$, the set S_i contains a node of U, i.e., $S_i \cap U \neq \emptyset$.*
2. *$|U| \leq 3xt/q = O(at \ln(kt)/q)$.*

A second ingredient of our algorithm is a simple technique for handling adversarial distribution of loads. The lemma below was observed by Levcopoulos and Overmars [LO88] in the context of balanced search trees. Thorup later applied it to bounding the number of precomputed paths through certain nodes in his fully dynamic APSP algorithm [Tho05].

Lemma 2.2 ([LO88])**.** *Consider the following process for repeatedly distributing L stones on k piles. In each round, remove the pile with the maximum number of stones (together with the stones it contains) and let an adversary distribute a total of at most L stones on the remaining piles. Then, at any time, the maximum number of stones on any pile is $O(L \log k)$.*

Our last ingredient is a reduction for obtaining a fully dynamic algorithm from a decremental algorithm. This approach was first taken by Henzinger and King for the dynamic minimum spanning tree (MST) problem [HK01] and a later variant was used by Thorup [Tho05]. Consider a decremental algorithm that, after a preprocessing stage, can handle a single batch of up to 2Δ deletions. First we reduce the cost of the pre-processing by a factor of Δ at the cost of supporting at most Δ deletions (instead of 2Δ). This is done by keeping two copies: at every interval of Δ operations, one copy is used to answer queries and the other copy is being

gradually built (in each round a $1/\Delta$ fraction of the pre-computation is executed). When the gradually built copy is ready to serve queries, it is Δ rounds behind, but this is okay because it can handle 2Δ deletions and we simply add the at most Δ deletions of the previous interval to the at most Δ deletions on the current interval.

Second, reducing this decremental-only algorithm to a fully dynamic one at the cost of $O(\Delta n^2)$ time per update is done by running a modification of the Floyd-Warshall algorithm. Recall that standard Floyd-Warshall algorithm operates in iterations. In each iteration the algorithm selects a new vertex and updates the shortest paths of all pairs by allowing them to use the new selected node together will all previously selected nodes. Notice that for our needs, this means that we can start from the result of the decremental algorithm and do only 2Δ iterations, one for each new inserted node. Each iteration takes time $O(n^2)$ so the entire process takes time $O(\Delta n^2)$. Note that we do this operation from scratch after every update. That is, after every update (deletion or insertion) the algorithms invokes these $O(\Delta)$ Floyd-Warshall iterations. Therefore, we only need to maintain the decremental data structure dynamically. In Section 4.1 we extend this reduction such that the graph may have negative weights, but no negative cycles, while the decremental algorithm only needs to work with non-negative edge weights.

Lemma 2.3 ([HK01],[Tho05])**.** *If there is a decremental APSP algorithm supporting any sequence of up to 2Δ deletions that spends time t_{pre} for preprocessing the initial graph and worst-case time t_{del} per deletion, then there is also a fully dynamic APSP algorithm with a worst-case update time of $O(t_{pre}/\Delta + t_{del} + \Delta n^2)$. The query time of the fully dynamic algorithm is proportional to the query time of the decremental algorithm.*

3 Decremental shortest paths for a batch of deletions

In this section we design a randomized algorithm with the following properties. We are given a directed graph with non-negative edge weights and preprocess it in time $\tilde{O}(n^3)$. After the preprocessing phase, a batch of nodes D to be deleted from the graph is given to us and we have to compute the all-pairs distances in $G \setminus D$. We will perform this task in time $\tilde{O}(n^{2.5}\sqrt{|D|})$.

Theorem 3.1 (Batch deletion algorithm)**.** *Given a graph $G = (V, E)$ and a parameter $c \geq 1$, Algorithm 1 computes a data structure $\mathcal{D}$ in time $O(n^3 \log^2 n)$ such that given $\mathcal{D}$ and a single set of nodes $D \subseteq V$, Algorithm 2 computes the all-pairs distances of $G \setminus D$ in time $O(n^{2.5}\sqrt{|D|}\log n)$. The running time bounds and the correctness each hold with probability at least $1 - 1/n^c$.*

Using the reduction of Lemma 2.3 this immediately implies our main result (Theorem 1.1) by setting $\Delta = \lceil n^{1/3} \log^{2/3} n \rceil$. We now describe the decremental algorithm and then analyze its correctness and its running time.

3.1 Algorithm description

Our algorithm follows a hierarchical approach where, for every $1 \leq i \leq \lceil \log n \rceil$, layer i is used to obtain the shortest paths whose number of edges is between $h_i/2$ and h_i where $h_i = 2^i$.

Preprocessing. For every layer i (with $1 \leq i \leq \lceil \log n \rceil$), in the preprocessing phase, we first randomly sample a set C_i of $\tilde{O}(n/h_i)$ nodes, called centers, which with high probability will 'hit' every shortest path in the graph that has at least $h_i/2$ edges. The bound on the size of C_i is guaranteed with high probability. After the sampling we visit a subset R_i of the nodes in a specific order that is determined by the algorithm 'on-the-fly'. Every time we visit a node v, we perform the following operations in the graph G_i^v from which all edges incident to previously visited nodes have been removed: for every node x we compute the shortest $\leq h_i$ hop path $\pi_i^v(v, x)$ from v to x and the shortest $\leq h_i$ hop path $\pi_i^v(x, v)$ from x to v by running h_i iterations of the Bellman-Ford algorithm in G_i^v and its reverse graph. To be precise, we compute the shortest $\leq h_i$ hop paths with the minimum number of edges. We keep a counter for every node u to count the 'congestion' of u, which for us is the total number of shortest $\leq h_i$ hop paths containing u computed in the preprocessing.

The order in which we visit the nodes is determined as follows. Until all centers have been visited, we alternate between visiting the center with the largest congestion and the non-center node with the largest congestion. We will show that this greedy strategy limits the maximum congestion of each node, which in turn bounds the work we have to do for updating the shortest paths if this node is deleted later on. Additionally, we compute for every pair of nodes s and t and every visited node v the distance $\delta_i(s, v, t)$ from s to t through v using the in- and out-trees computed for v; we then sort all nodes $v \in R_i$ according to their value $\delta_i(s, v, t)$ and store them in a list $L_i(s, t)$. Algorithm 1 shows the pseudocode for the preprocessing. Observe that a shortest path π from s to t can by reconstructed from the stored paths if, for the hop range $h_i/2 \dots h_i$ of π, we identify the node v on π that was visited first, which means that π is contained in G^v, and concatenate $\pi_i^v(s, v)$ and $\pi_i^v(v, t)$. Thus, the minimum of all $\delta_i(s, v, t)$ gives the distance from s to t (and we can construct the corresponding shortest path). Whenever a node u is deleted from the graph we destroy some of the paths $\pi_i^v(s, v)$ or $\pi_i^v(v, t)$ and for such nodes s and t we will recompute the shortest paths from and to v. The number of destroyed paths is equal to the congestion of u, which is our motivation for limiting the maximum congestion of each node.

Batch deletions. A batch of deletions given by a set D is processed by first recomputing the shortest paths consisting of up to $h = \sqrt{n/|D|}$ edges and then recomputing the shortest paths consisting of more than h edges. To find the shortest paths with at most h edges, we perform the following steps for every $1 \leq i \leq \lceil \log h \rceil$. For every $v \in R_i$, we first iterate over every path $\pi_i^v(x, v)$ from the preprocessing stage to determine every node x for which the path *to* v has been destroyed by one of the deletions in D and store these nodes in the set U_i^v of *affected* nodes. Similarly, we add to U_i^v all nodes whose shortest path *from* v has been destroyed. Note that for all non-affected nodes the shortest $\leq h_i$ path from and to v did not change since the preprocessing.

We next explain how to compute the new shortest paths from and to v for nodes in U_i^v as follows. We compute a sparse sketch graph H_i^v consisting of the following edges. For each affected node $x \in U_i^v$ we add all its incident edges (both incoming and outgoing) to H_i^v. For each non-affected node $x \notin U_i^v$ we include two edges: one to the successor of x on $\pi_i^v(x, v)$

Algorithm 1: Preprocessing a graph $G = (V, E)$

Procedure PREPROCESS(G)
  for $i = 1$ **to** $\lceil \log n \rceil$ **do**
    $h_i \leftarrow 2^i$
    $x \leftarrow 1 + 3(c+1)\ln n$
    $C_i \leftarrow \emptyset$ `// Set of randomly chosen centers`
    foreach $v \in V$ **do** $C_i \leftarrow C_i \cup \{v\}$ with probability $\min(x/h_i, 1)$
    $R_i \leftarrow \emptyset$ `// Set of visited nodes`
    foreach $v \in V$ **do**
      $c_i(v) \leftarrow 0$ `// Initialize congestion counter`
    while $C_i \setminus R_i \neq \emptyset$ **do** `// Repeat until all centers have been visited`
      $v \leftarrow \arg\max_{v \in C_i \setminus R_i} c_i(v)$ `// Visit center with largest congestion`
      VISIT(v, i)
      if $V \setminus (C_i \cup R_i) \neq \emptyset$ **then** `// If there are unvisited non-center nodes`
        `// Visit non-center node with largest congestion`
        $v \leftarrow \arg\max_{v \in V \setminus (C_i \cup R_i)} c(v)$
        VISIT(v, i)
    foreach *pair of nodes* $s, t \in V$ **do**
      Construct list $L_i(s,t)$ containing all nodes $v \in R_i$ sorted by their value $\delta_i(s,v,t)$

Procedure VISIT(v, i)
  Construct $G_i^v = (V, E \setminus (V \times R_i \cup R_i \times V))$ (in which all edges incident to previously visited nodes are removed)
  Run h_i iterations of Bellman-Ford from v in G_i^v and its reverse graph to compute for every node $x \in V \setminus R_i$ the shortest $\leq h_i$ hop path $\pi_i^v(v,x)$ from v to x and the shortest $\leq h_i$ hop path $\pi_i^v(x,v)$ from x to v in G_i^v
  $R_i \leftarrow R_i \cup \{v\}$
  foreach $u \in V \setminus R_i$ **do**
    `// Increase congestion counter of` u `by number of` $\leq h_i$ `hop shortest paths containing` u
    $c_i(u) \leftarrow c_i(u) + |\{x \in V \mid u \in \pi_i^v(x,v) \text{ or } \pi_i^v(v,x)\}|$
  foreach *pair of nodes* $s, t \in V$ **do**
    $\delta_i(s,v,t) \leftarrow d^{h_i}_{G_i^v}(s,v) + d^{h_i}_{G_i^v}(v,t)$

and one to the predecessor of x on $\pi_i^v(v,x)$. We will show that by this rule all shortest paths from and to v in $G_i^v \setminus D$ with at most h_i edges are present in the sketch graph. We then run Dijkstra's algorithm to determine the shortest paths to and from v in H_i^v. Furthermore, for each affected node $x \in U_i^v$ and every possible start or endpoint y, we recompute $\delta_i(x,v,y)$, the distance from x to y through v in H_i^v and $\delta_i(y,v,x)$, the distance from y to x through v in H_i^v, possibly replacing the corresponding value computed in the preprocessing.

To find the shortest paths consisting of more than h edges, we first run Dijkstra's algorithm to compute the shortest paths to and from every center in $C_{\lceil \log h \rceil}$ and then compute the shortest paths through these centers. Finally, we determine, for every pair of nodes s and t, the shortest path distance $\delta(s,t)$ as the length of the shortest path through any of the centers used in this algorithm. Algorithm 2 shows the pseudocode for batch deletions.

Algorithm 2: Deletion procedure for a single batch of nodes D

$h \leftarrow \sqrt{n/|D|}$
for $i = 1$ **to** $\lfloor \log h \rfloor$ **do**
  foreach $v \in R_i$ **do**
    `// Iterate over all paths from preprocessing to determine affected nodes whose shortest` $\leq h_i$ `hop paths to or from` v `was destroyed`
    $U_i^v \leftarrow \{x \in V \mid \pi_i^v(x,v) \cap D \neq \emptyset \text{ or } \pi_i^v(v,x) \cap D \neq \emptyset\}$
    `// Construct sketch graph` $H_i^v = (V, E_i^v)$
    $E_i^v \leftarrow \emptyset$
    `// Add edges to and from all neighbors for affected nodes`
    foreach $y \in U_i^v$ **do** $E_i^v \leftarrow E_i^v \cup \{(y,z) \mid (y,z) \in E\} \cup \{(z,y) \mid (z,y) \in E\}$
    `// Add edges of paths from the preprocessing stage for unaffected nodes`
    foreach $y \notin U_i^v$ **do**
      Determine predecessor x of y on $\pi_i^v(v,y)$ and successor z of y on $\pi_i^v(y,v)$
      $E_i^v \leftarrow E_i^v \cup \{(x,y),(y,z)\}$
    Compute SSSP from and to v in $H_i^v = (V, E_i^v)$ using Dijkstra's algorithm
    `// Update shortest paths through centers for pairs with at least one affected node`
    foreach $(s,t) \in U_i^v \times V$ *and* $(s,t) \in V \times U_i^v$ **do** $\delta_i(s,v,t) \leftarrow d_{H_i^v}(s,v) + d_{H_i^v}(v,t)$
  foreach *pair of nodes* $s,t \in V$ **do**
    $\delta_i(s,t) \leftarrow \min_{v \in R_i} \delta_i(s,v,t)$ `// see 3.8 for implementation using` $L_i(s,t)$
foreach $v \in C_{\lceil \log h \rceil}$ **do** Compute SSSP from and to v in $G \setminus D$ using Dijkstra's algorithm
foreach *pair of nodes* $s,t \in V$ **do**
  $\delta_{\lceil \log h \rceil}(s,t) \leftarrow \min_{v \in C_{\lceil \log h \rceil}} (d_{G \setminus D}(s,v) + d_{G \setminus D}(v,t))$
  $\delta(s,t) \leftarrow \min_{1 \leq i \leq \lceil \log h \rceil} \delta_i(s,t)$

3.2 Correctness

We have to show that our algorithm can serve a batch deletion correctly with probability at least $1 - 1/n^c$. Just for the following analysis we assume that among all shortest paths for a fixed pair of nodes there is a single *designated* shortest path (e.g., the smallest according to some order). In order to prove this statement we first show that the sketch graphs used in our algorithm contain the shortest paths relevant to us.

Claim 3.2. *Let $1 \leq i \leq \lceil \log n \rceil$. For every $v \in R_i$ and all pairs of nodes $s, t \in G_i^v \setminus D$,*

- *if there is a shortest path from s to v in $G_i^v \setminus D$ with at most h_i edges, then $d_{H_i^v}(s, v) = d_{G_i^v \setminus D}(s, v)$*
- *if there is a shortest path from v to t in $G_i^v \setminus D$ with at most h_i edges, then $d_{H_i^v}(v, t) = d_{G_i^v \setminus D}(v, t)$.*

Proof. We only give a proof of the first item as the proof of the second item is symmetric. Let π denote the designated shortest path from s to v in $G_i^v \setminus D$ with at most h_i edges. The proof is by induction on the number of edges of π. Remember that $\pi_i^v(s, v)$ denotes the shortest $\leq h_i$ hop path from s to v in G_i^v determined in the preprocessing. Let x be the successor of s on π and let y be the successor of s on $\pi_i^v(s, v)$.

If $s \in U_i^v$, then the edge (s, x) is contained in H_i^v by the definition of H_i^v and by applying the induction hypothesis on x we get $d_{H_i^v}(s, v) = d_{G_i^v \setminus D}(s, v)$. If $s \notin U_i^v$, then $\pi_i^v(s, v)$ does not contain any deleted nodes and the edge (s, y) is contained in H_i^v. The weight of $\pi_i^v(s, v)$ equal the weight of π because otherwise π would have been a shorter $\leq h_i$ hop path than $\pi_i^v(s, v)$ in G_i^v during the preprocessing. Thus, $\pi_i^v(s, v)$ is a shortest path in $G_i^v \setminus D$. Furthermore the number of edges of $\pi_i^v(s, v)$ is at most the number of edges of π as $\pi_i^v(s, v)$ is the shortest $\leq h_i$ hop path with the *minimum* number of edges in G_i^v. Therefore we may apply the induction hypothesis on y and conclude that $d_{H_i^v}(s, v) = d_{G_i^v \setminus D}(s, v)$. □

Claim 3.3. *With probability at least $1 - 1/n^c$, for all pairs of nodes $s, t \in V$, if the designated shortest path from s to t has at least $h_i/2$ hops, then it contains a center $v \in C_i$.*

Proof. We argue that all lexicographic shortest paths with at least 2^{i-1} edges contain a node of C_i with probability at least $1 - 1/n^c$. We apply Lemma 2.1 with $a = c$, $T = V$, $U = C_i$, $t = n$, $q = h_i/2$ and by defining $S_1, \ldots, S_k$ with $k \leq n^2$ as the sets of nodes on the at most n^2 lexicographic shortest paths of pairs of nodes in $G \setminus D$ with at least $(h_i/2)$ edges. Remember that the algorithm performs the sampling by picking each node with probability $\min(x/q, 1)$ where $x \leq c \ln n^3 + 1$. Thus, all lexicographic shortest paths with at least 2^{i-1} edges contain a center from C_i with probability at least $1 - 1/n^c$. □

Claim 3.4. *With probability at least $1 - 1/n^c$, $\delta(s, t) = d_{G \setminus D}(s, t)$ for all pairs of nodes s and t.*

Proof. We prove the claim by assuming that the statement of 3.3 holds. Thus, the claim we prove also holds with probability at least $1 - 1/n^c$.

First, we argue that $\delta(s,t) \geq d_{G\setminus D}(s,t)$. Observe that whenever we set the value of $\delta_i(s,v,t)$ (for some $1 \leq i \leq \lceil \log h \rceil$ and some $v \in R_i$) during the preprocessing or the deletion procedure, this value corresponds to the length of a path in a subgraph of G. Furthermore, if $\pi_i^v(s,v)$ or $\pi_i^v(v,t)$ contains a deleted node from D, then the value $\delta_i(s,v,t)$ is updated to the length of a path in $G \setminus D$. It is not hard to verify now that $\delta(s,t) \geq d_{G\setminus D}(s,t)$.

We now argue that $\delta(s,t) \leq d_{G\setminus D}(s,t)$. Let π denote the designated shortest path from s to t in $G \setminus D$. Consider first the case that π consists of at most h edges (where $h = \sqrt{n/|D|}$, as set in the algorithm). Let $1 \leq i \leq \lfloor \log h \rfloor$ be the index such that the number of edges of π is between 2^{i-1} and 2^i. By 3.3, π contains a center of C_i with probability at least $1 - 1/n^c$. As $C_i \subseteq R_i$, π contains at least one node of R_i. Now let v be the node that, among all nodes of R_i contained in π, has been visited first in the preprocessing. Then all edges of π are contained in $G_i^v \setminus D$. If either $s \in U_i^v$ or $t \in U_i^v$, then the update algorithm has set $\delta_i(s,v,t) = d_{H_i^v}(s,v) + d_{H_i^v}(v,t)$. By 3.2, we have $d_{H_i^v}(s,v) = d_{G_i^v\setminus D}(s,v)$ and $d_{H_i^v}(v,t) = d_{G_i^v\setminus D}(v,t)$. It follows that

$$\delta(s,t) \leq \delta_i(s,v,t) = d_{H_i^v}(s,v) + d_{H_i^v}(v,t) = d_{G_i^v\setminus D}(s,v) + d_{G_i^v\setminus D}(v,t) \leq d_{G\setminus D}(s,v) + d_{G\setminus D}(v,t) = d_{G\setminus D}(s,t)\,.$$

If both $s \notin U_i^v$ and $t \notin U_i^v$, then we have set $\delta_i(s,v,t) = d_G^{h_i}(s,v) + d_G^{h_i}(v,t)$ in the preprocessing. As $s \notin U_i^v$ and $t \notin U_i^v$, the paths $\pi_i^v(s,v)$ and $\pi_i^v(v,t)$ both are contained in $G \setminus D$ and thus

$$\delta(s,t) \leq \delta_i(s,v,t) = d_G^{h_i}(s,v) + d_G^{h_i}(v,t) \leq d_{G\setminus D}^{h_i}(s,v) + d_{G\setminus D}^{h_i}(v,t) = d_{G\setminus D}(s,v) + d_{G\setminus D}(v,t) = d_{G\setminus D}(s,v)\,.$$

Finally, consider the case that π consists of at least h edges. Then, by another application of 3.3, π contains a center $v \in C_{\lceil \log h \rceil}$ and thus $\delta(s,t) \leq \delta_{\lceil \log h \rceil}(s,t) \leq d_{G\setminus D}(s,v) + d_{G\setminus D}(v,t) = d_{G\setminus D}(s,t)$. □

3.3 Running time

In the following we analyze the time our algorithm needs for performing the preprocessing and for recomputing the all-pairs distances after a batch of deletions. The running time guarantees hold with high probability as they depend on the size of the randomly chosen sets in the preprocessing.

Claim 3.5. *With probability at least $1 - 1/n^c$, the size of C_i is $O((n \log n)/h_i)$ for all $1 \leq i \leq \lceil \log n \rceil$.*

The claim follows in a straightforward way from Lemma 2.1 and its proof is thus omitted. In the remainder of this section we omit the c in the O-notation for readability.

Claim 3.6. *The preprocessing takes time $O(n^3 \log^2 n)$ with probability at least $1 - 1/n^c$.*

Proof. We analyze the running time of each iteration i, where $1 \leq i \leq \lceil \log n \rceil$. We obtain the set C_i of sampled nodes in time $O(n)$. Running h_i iterations of Bellman-Ford on a graph with

at most n nodes takes time $O(h_i n^2)$. We perform one such computation for every node in R_i and $|R_i| \leq 2|C_i| = O((n \log n)/h_i)$ (by 3.5). Thus, the total time spent for the Bellman-Ford computations in iteration i is $O(|R_i| n^2 h_i)) = O(n^3 \log n)$. For updating the counters we iterate over all nodes in the $\leq h_i$ shortest paths in total time $O(|R_i| n h_i) = O(n^3)$. For constructing the list $L_i(s,t)$ for each pair of nodes s and t we sort at most n nodes and thus constructing all the lists takes time $O(n^3 \log n)$. It follows that the running time in each iteration is dominated by the term $O(n^3 \log n)$. As there are $O(\log n)$ iterations, the total time spent in the preprocessing is $O(n^3 \log^2 n)$. □

To bound the time for processing a batch of deletions we first show that the special order in which we have visited the nodes in the preprocessing stage guarantees that only few nodes are affected by each deletion. The fewer nodes are affected, the faster our algorithm runs.

Claim 3.7. *For each $1 \leq i \leq \lceil \log n \rceil$, every node is contained in at most $O(h_i n \log n)$ of the stored shortest $\leq h_i$ hop paths after the preprocessing, i.e., $|\{(x,v) \in V \times R_i \mid u \in \pi_i^v(x,v) \text{ or } u \in \pi_i^v(v,x)\}| = O(h_i n \log n)$.*

Proof. Observe that we can prove the claim by showing that for every node u the counter $c_i(u)$ is $O(h_i n \log n)$ at any time. In the preprocessing we alternate between (a) visiting the center node with maximum counter and (b) visiting the non-center node with maximum counter (if such a node exists). After visiting a node v, the sum of the counters increases by at most $2h_i n$ as for every node x the number of nodes on $\pi_i^v(x,v)$ and $\pi_i^v(x,v)$ (except for v itself) is at most h_i, respectively. A visited node is removed from the graph and thus not visited again in iteration i.

Consider the following two processes for distributing stones on piles. In process 1 we have $k_1 = |C_i|$ piles, one for each center, and in process 2 we have $k_2 = |V \setminus C_i|$ piles, one for each non-center node. Every time our algorithm visits a node v we do the following: If v is a center node, we remove the corresponding pile (and the nodes it contains) from process 1. Similarly, if v is a non-center node, we remove the corresponding pile (and the nodes it contains) from process 2. After visiting v, the algorithm increases the counters of certain nodes and we add the corresponding number of stones to the piles in processes 1 and 2.

Observe that whenever we remove a pile from process 1 this pile always carries the maximum number of stones and between any two removals of piles we have added at most $4h_i n$ stones to the piles. Therefore, the total number of stones on any pile of process 1 (and thus the maximum counter of any center) is $O(h_i n \log k_1) = O(h_i n \log n)$ by Lemma 2.2. By the same reasoning the maximum counter of any non-center node is $O(h_i n \log n)$ as well. □

Claim 3.8. *The time for processing a single batch of deletions given by a set D is $O(n^{2.5}\sqrt{|D|} \log n)$.*

Proof. For every $1 \leq i \leq \lfloor \log h \rfloor$ and every $v \in R_i$, we compute the set of affected nodes U_i^v by iterating over all shortest $\leq h_i$ hop paths to and from v determined in the preprocessing. This takes time $O(h_i n)$ for fixed v and i and thus time $O(\sum_{i=1}^{\lfloor \log h \rfloor} |R_i| h_i n) = O(n^2 \log^2 n)$ in total.

Constructing all sketch graphs and then running Dijkstra's algorithm on them takes time $O(\sum_{i=1}^{\lfloor \log h \rfloor} \sum_{v \in R_i} (|E_i^v| + n \log n))$. For every $1 \leq i \leq \lfloor \log h \rfloor$ we bound the size of the sketch

graph H_i^v by $|E_i^v| \leq |U_i^v|n + 2n$ as each node in U_i^v contributes at most n edges to all its neighbors and each other nodes contributes at most 2 edges. By 3.7, each deleted node of D is contained in $O(h_i n \log n)$ of the $\leq h_i$ hop shortest paths determined in the preprocessing. Therefore the total number of nodes affected by the deletions is $\sum_{v \in R_i} |U_i^v| = O(|D| h_i n \log n)$ and thus $\sum_{v \in R_i} |E_i^v| = O(|D| h_i n^2 \log n + |R_i| n)$. Dijkstra's algorithm in all sketch graphs of layers 1 to $\lfloor \log h \rfloor$ then takes total time

$$O\left(\sum_{i=1}^{\lfloor \log h \rfloor} \sum_{v \in R_i} (|E_i^v| + n \log n)\right) \leq O\left(\sum_{i=1}^{\lfloor \log h \rfloor} (|D| h_i n^2 \log n + |R_i| n \log n)\right) \leq$$
$$O\left(\sum_{i=1}^{\lfloor \log h \rfloor} (|D| 2^i n^2 \log n + n^2 \log n)\right) \leq O(|D| h n^2 \log n + n^2 \log^2 n) \leq O(n^{2.5} \sqrt{|D|} \log n)\,.$$

To compute the minimum $\delta(s,t) = \min_{v \in R_i} \delta_i(s,v,t)$ we do the following. We first compute the minimum over all values for which $\delta_i(s,v,t)$ has changed since the preprocessing. Then, we find the first value of the sorted list $L_i(s,t)$ for which the value $\delta_i(s,v,t)$ computed in the preprocessing has *not* changed. Clearly, the minimum of both values gives $\min_{v \in R_i} \delta_i(s,v,t)$ and both steps takes time proportional to the number of changed values, which is $O(|D| h_i n^2 \log n)$.

Finally, we bound the time spent on running Dijkstra's algorithm for every node $v \in C_{\lceil \log h \rceil}$ and computing $\delta_{\lceil \log h \rceil}(s,t)$ for every pair of nodes s and t. By the sampling procedure, the size of $C_{\lceil \log h \rceil}$ is $O((n \log n)/h)$ and thus both of these steps take time $O((n^3 \log n)/h) = O(n^{2.5} \sqrt{|D|} \log n)$. □

4 Extensions and Additional Results

4.1 Negative edge weights

We extend the reduction of Lemma 2.3 such that the graph may have negative weights, but no negative cycles, while the decremental algorithm only needs to work with non-negative edge weights. A potential function p is a function that assigns a value $p(v)$ to every node v such that, for every edge $(u,v) \in E$, $w(u,v) + p(u) - p(v) \geq 0$. Such a potential exists if and only if the graph contains no negative cycle. We call $w_p(u,v) = w(u,v) + p(u) - p(v)$ the modified weight of the edge (u,v) under the potential function p. For any valid potential function the modified edge weights are obviously non-negative and it is well-known that the shortest paths under the original edge weights are the same as under the modified edge weights. This is known as the "Johnson transformation" [Joh77].

The reduction now has the following additional steps. In the preprocessing we construct a graph G_q that contains an additional node q and an edge (q,v) of weight 0 to every other node v. On this graph we run the Bellman-Ford algorithm in time $O(mn) = O(n^3)$ which either detects a negative cycle or computes $d_{G_q}(q,v)$ for every node v. It is well-known that $p(v) = d_{G_q}(q,v)$ is a valid potential function. Observe that this potential function remains valid even when we delete edges or nodes from the graph. Thus, the update procedure of our decremental algorithm computes the shortest paths of the graph correctly. After running

the updates of the decremental algorithm, we undo the potential transformation. In the fully dynamic algorithm, we then deal with the up to 2Δ inserted nodes by running 2Δ iterations of the Floyd-Warshall algorithm, which by default can deal with negative edge weights.

4.2 Unweighted graphs

In unweighted graphs the shortest $\leq h$ hop paths are identical to the shortest paths with at most h edges. Thus, in the preprocessing, we can determine the shortest $\leq h_i$ hop paths by performing breadth-first search up to depth h_i in time $O(n^2)$. The total time spent for layer i in the preprocessing is therefore $O(n^3/h_i)$.

We slightly alter the reduction of Lemma 2.3 to obtain the fully dynamic algorithm by using different value of Δ for each layer i. Specifically, we use layer i of the decremental algorithm for $2\Delta_i = 2\sqrt{n}/(h_i\sqrt{\log n})$ updates before we rebuild it. At every update, besides running each layer of the decremental algorithm we reconstruct shortest paths consisting of more than $h = \sqrt{n}$ edges in time $O((n^3 \log n)/h)$ and run the Floyd-Warshall algorithm for at most $2\Delta = 2\sqrt{n}$ iterations in time $O(\Delta n^2)$ to handle the latest 2Δ insertions (note that $\Delta \geq \Delta_i$ for all $1 \leq i \leq \lfloor \log h \rfloor$). Thus, the total update time is

$$O\left(\sum_{1\leq i\leq \lfloor \log h\rfloor}\left(\frac{n^3}{h_i\Delta_i} + \Delta_i h_i n^2 \log n + \Delta n^2\right) + \frac{n^3 \log n}{h}\right) = O(n^{2+1/2}\log^{3/2} n)\,.$$

4.3 Deterministic algorithm

In the following we sketch a deterministic fully dynamic APSP algorithm with a worst-case update time of $O(n^{2+3/4}\log^{3/4} n)$. We again design an algorithm that, after some preprocessing, can handle a single batch of up to Δ deletions and turn this into a fully dynamic algorithm using the reduction of Lemma 2.3. The parameters of our algorithm are Δ and h and we will explain how to set them optimally in the running time analysis.

In the preprocessing we prepare a data structure from which the shortest paths with at most h edges can be reconstructed after any batch of at most Δ deletions. We again keep a congestion counter for each node and in each round visit the node with the maximum counter until *all* nodes have been visited. When we visit a node v, we compute the shortest $\leq h$ hop paths to and from v, denoted by $\pi^v(x, v)$ and $\pi^v(v, x)$, by running h iterations of the Bellman-Ford algorithm in the graph G^v from which all edges incident to previously visited nodes have been removed. We then update the congestion counters according to the number of appearances of each node in these paths. By Lemma 2.2, this order of visiting nodes guarantees that after the preprocessing, for every node u, there are at most $O(hn \log n)$ pairs (x, v) or (v, x) such that v is contained in $\pi^v(x, v)$ or $\pi^v(v, x)$, respectively. Again, the idea is that if we later on delete u we only have to repair the shortest paths for such pairs.

When we delete a set of nodes D, we first determine, for every node v, the set of affected nodes U^v consisting of all nodes x such that $\pi^v(s, v)$ or $\pi^v(v, t)$ was destroyed by one of the deletions. We then build a sketch graph H^v as follows: For every affected node $x \in U^v$ we add edges to all neighbors of x to H^v and for every non-affected node $x \notin U^v$ we add the

edge to the successor on $\pi^v(x,v)$ as well as the edge to the predecessor on $\pi^v(x,v)$ to H^v. We then run Dijkstra's algorithm in H^v to recompute the shortest paths to and from v with at most h edges in $G^v \setminus D$. Finally, for every pair of nodes s and t we find the "middle node" v connecting s and t in the shortest way, either by using paths from the preprocessing stage or new paths computed during the update procedure. This gives us all shortest paths that have at most h edges. Using this information we can then compute all shortest paths with more edges deterministically by finding a set of centers that "hits" all the shortest paths with at most h edges using a greedy heuristic. To the best of our knowledge this method was first described in [Zwi02].

Lemma 4.1 ([Zwi02])**.** *For every (weighted directed) graph G and every $h \geq 1$, given all shortest paths in G that have at most h hops, one can compute $d_G(s,t)$ for all pairs of nodes s and t (and the corresponding shortest paths) in time $O(hn^2 + n^3 \log n/h)$.*

The correctness and the running time follow the same arguments as for the randomized algorithm in Section 3. However, the deterministic algorithm is slower because we do not know how to extend it to the multilayer approach of the randomized algorithm. The preprocessing time is dominated by the term $O(n^3 h)$, which is the time needed for running h iterations Bellman-Ford algorithm for all nodes. As in the randomized algorithm, the running time for processing a single batch of deletions is dominated by the time needed to run Dijkstra's algorithm in each sketch graph where the total size of all sketch graphs is $O(\Delta h n^2 \log n)$. Thus, running Dijkstra's algorithm in all sketch graphs takes total time $O(\Delta h n^2 \log n + n^2 \log n) = O(\Delta h n^2 \log n)$. Overall, the update time of this deterministic fully dynamic algorithm is $O(n^3 h/\Delta + \Delta h n^2 \log n + h n^2 + n^3 \log n/h + \Delta n^2)$. By setting $\Delta = n^{1/2}/\log^{1/2} n$ and $h = n^{1/4}/\log^{1/4} n$, this is $O(n^{2+3/4} \log^{3/4} n)$.

4.4 Returning shortest paths

The algorithm of Theorem 1.1 we have formulated above only returns the distance matrix for all pairs of nodes. Our decremental algorithm be extended in a straightforward way to return a shortest path matrix that contains, for every pair of nodes s and t the first edge (s,x) on a shortest path from s to t by additionally storing with every intermediate distance estimate the first edge on the path yielding the corresponding value. The Floyd-Warshall algorithm can also compute the shortest path matrix along with the distances. This matrix provides sufficient information to compute a shortest path between a source and a target in time proportional to the number of edges of this path. Without additional effort however, we can only show that such an algorithm is correct with high probability against an oblivious adversary who chooses its sequence of updates and queries in advance. In the following we sketch how to modify our algorithm such that it is correct with high probability against an adaptive online adversary that may adapt its sequence of updates according to the shortest path matrix returned by the algorithm. Intuitively, this means that we have to avoid that the adversary chooses a clever sequence of updates and queries such that it can identify the random centers picked by our algorithm and alter the graph such that the properties we gained from the random choice of centers will not hold anymore. Note that this is not an issue for algorithms that merely

compute the distance matrix since the exact distances are uniquely defined in every graph and thus the adversary cannot learn anything about our algorithm (and its random choices) by observing its outputs. If we allow path queries, where the adversary may pick any pair of nodes s and t and ask for a shortest path from s to t, then we do not have this uniqueness property because there might be several paths from s to t of minimum weight.

We counter this problem by modifying our algorithm such that it computes lexicographic shortest paths at the cost of an additional factor of $\log n$ in the update time.[3] First, we assume without loss of generality that all shortest paths of the graph have the same number of edges. This can be ensured by adding a sufficiently small penalty to every edge weight. Now assume an arbitrary but fixed order on the nodes and identify a path with its sequence of nodes. We say that a path π_1 is lexicographically smaller than a path π_2 if either π_1 is a prefix of π_2 or $\pi_1 = \pi \circ v_1 \circ \pi_1''$ and $\pi_2 = \pi \circ v_2 \circ \pi_2''$ where $v_1 < v_2$. The latter condition means that at the first position where π_1 and π_2 differ, the node of π_1 is a smaller than the one of π_2. The lexicographic shortest path from s to t is the path that among all shortest paths from s to t is smallest according to the lexicographic order. To compare paths lexicographically we use the *minimum-index tree structure* (short: MITS) of Cabello, Chambers, and Erickson [CCE13], which can be implemented using Euler-tour trees [RHK99, Tar97] or self-adjusting top trees [TW05]. This minimum index tree structure allows adding and removing edges to and from a tree with given root s. Given two nodes u and v as its input it can answer the following query: is the path from s to u in the tree lexicographically smaller than the path from s to v? All operations of this data structure take time $O(\log n)$. Cabello, Chambers, and Erickson [CCE13] explain how use the MITS to adapt Dijkstra's algorithm for computing lexicographic shortest paths at the cost of an additional factor of $O(\log n)$ in the running time. In a similar way the Bellman-Ford algorithm can be modified to compute the lexicographic $\leq h$ hop shortest paths.

We now explain how to modify our algorithm to compute, for every pair of nodes s and t, the first edge (s, x) on the lexicographic shortest path from s to t in $G \setminus D$. In the preprocessing algorithm the only modification is to use the modified Bellman-Ford algorithm to compute the lexicographic $\leq h_i$ hop shortest paths. In the update procedure we first run the algorithm for computing the distance matrix completely so that we know the value of $d_{G\setminus D}(s,t)$ for all pairs of nodes s and t. We then initialize an MITS for every node s and add to it the following paths computed by our algorithm.

1. For every $1 \leq i \leq \lfloor \log h \rfloor$ and every $v \in R_i$ we add the shortest $\leq h_i$ hop path $\pi_i^v(s,v)$ from the preprocessing to the MITS if the weight of $\pi_i^v(s,v)$ is equal to $d_{G\setminus D}(s,v)$.

2. For every $1 \leq i \leq \lfloor \log h \rfloor$ and every $v \in R_i$ such that $s \in U_i^v$, we add the lexicographic shortest path π from s to v in H_i^v to the MITS if π has at most h_i edges and the weight of π is equal to $d_{G\setminus D}(s,v)$.

3. For every $v \in R_{\lceil \log h \rceil}$ we add the lexicographic shortest path from s to v in $G \setminus D$ to the

[3] An intuitive alternative for enforcing that shortest paths are unique is to add random perturbation to the edge weights. However we do not know how to use this scheme in our dynamic setting. While one can guarantee that a random perturbation makes shortest paths unique in the initial graph, it is not clear how to obtain this property in all versions of the graph.

MITS.

Using similar arguments as in Section 3.2 it follows that (1) we only add lexicographic shortest paths to the MITS leading to a tree structure and (2) for every lexicographic shortest path π from s to some node t there is some node v on π such that the subpath from s to v is contained in the MITS of s. The running time for these insertions of paths to the MITS's can be bounded as follows: As, for each $1 \leq i \leq \lfloor \log h \rfloor$, $R_i = O((n \log n)/h_i)$ and each inserted path has at most h_i edges, the total time for the insertions in step 1 is $O(n^2 \log^3 n)$. The total time for step 2 can be bounded by recalling that the total number of affected nodes is $\sum_{v \in R_i} |U_i^v| = O(\Delta h_i n \log n)$, yielding a total time of $O(\Delta h_i n^2 \log^2 n)$. Since $|R_{\lceil \log h \rceil}| = O((n \log n)/h)$ the total time for step 3 is $O((n^3 \log^2 n)/h)$.

To obtain the lexicographic shortest paths of $G \setminus D$ for every pair of nodes s and t (or rather the first edge of this path) we do the following: We let $X(s,t)$ be the set of nodes v satisfying one of the following three conditions:

1. $v \in R_i$ for some $1 \leq i \leq \lfloor \log h \rfloor$ and $\pi_i^v(s,v) + \pi_i^v(v,t) = d_G(s,t)$,
2. $v \in R_i$ such that $s \in U_i^v$ or $t \in U_i^v$ for some $1 \leq i \leq \lfloor \log h \rfloor$ and $d_{H_i^v}(s,v) + d_{H_i^v}(v,t) = d_G(s,t)$, or
3. $v \in R_{\lceil \log h \rceil}$ and $d_{G \setminus D}(s,v) + d_{G \setminus D}(v,t) = d_G(s,t)$.

Among the nodes in $X(s,t)$ we now want to find the node v whose lexicographic shortest path from x to v is smallest in $G \setminus D$ and output the first edge on this path. This can be done in time $O(|X(s,t)| \log n)$ using the MITS of s that we have prepared above. By our arguments from above, the total size of all these sets is bounded by $O(|D| h n^2 \log n + (n^3 \log n)/h)$. Using $h = \sqrt{n/|D|}$ we thus obtain a running time of $O(n^{2.5} \sqrt{|D|} \log n)$.

5 Conclusions

In this paper we considered the fully dynamic APSP problem with a worst-case update time guarantee of $O(n^{2+2/3} \log^{4/3} n)$. Our algorithm is simple and independent of any other sophisticated algorithms. Our current knowledge of lower bounds for this problem seems quite rudimentary. A natural barrier for the current approaches seems to be $\Omega(n^{2+1/2})$. One reason for this barrier is that the only way we know to deal with insertions is to use the naive approach, in which for every insertion (since the last time the data structure was reconstructed) in every update we compute a SSSP tree and recompute all pair-wise distances in these trees. This naive approach sets a barrier of $\Omega(n^{2+1/2})$. A natural question is on the existence of fully dynamic APSP algorithm that meet this barrier or prove impossibility results.

For unweighted graphs, our upper bound indeed meets this barrier. Weighted graphs seem to be inherently harder: for example, extending the algebraic techniques of Sankowski [San05] to weighted graphs is an open question. Our techniques for weighted graphs incur a cost related to computing single source h-hop shortest paths: the best known is time $\tilde{O}(n^2 h)$ for weighted graphs, and time $\tilde{O}(n^2)$ for unweighted graphs. If h-hop shortest paths could be solved in time

$\tilde{O}(n^2)$ for weighted graphs, then our techniques would immediately provide an improved results that would meet the natural $\Omega(n^{2+1/2})$ barrier.

We believe that it would be interesting to also explore the potential opposite connection: could hardness of h-hop shortest paths in weighted graphs imply lower bounds for dynamic shortest paths in weighted graphs? Another interesting direction is to derandomize our algorithm or prove an existential gap between randomized and deterministic algorithms.

References

[AC13] Ittai Abraham and Shiri Chechik. "Dynamic Decremental Approximate Distance Oracles with $(1+\epsilon, 2)$ stretch". In: *CoRR* abs/1307.1516 (2013) (cit. on p. 4).

[ACG12] Ittai Abraham, Shiri Chechik, and Cyril Gavoille. "Fully Dynamic Approximate Distance Oracles for Planar Graphs via Forbidden-Set Distance Labels". In: *Symposium on Theory of Computing (STOC)*. 2012, pp. 1199–1218 (cit. on p. 4).

[ACT14] Ittai Abraham, Shiri Chechik, and Kunal Talwar. "Fully Dynamic All-Pairs Shortest Paths: Breaking the $O(n)$ Barrier". In: *International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX)*. 2014, pp. 1–16 (cit. on p. 4).

[AIMS+91] Giorgio Ausiello, Giuseppe F. Italiano, Alberto Marchetti-Spaccamela, and Umberto Nanni. "Incremental Algorithms for Minimal Length Paths". In: *Journal of Algorithms* 12.4 (1991). Announced at SODA'90, pp. 615–638 (cit. on p. 4).

[BBK+94] Shai Ben-David, Allan Borodin, Richard M. Karp, Gábor Tardos, and Avi Wigderson. "On the Power of Randomization in On-Line Algorithms". In: *Algorithmica* 11.1 (1994). Announced at STOC'90, pp. 2–14 (cit. on p. 2).

[BC16] Aaron Bernstein and Shiri Chechik. "Deterministic decremental single source shortest paths: beyond the $O(mn)$ bound". In: *Symposium on Theory of Computing (STOC)*. 2016, pp. 389–397 (cit. on p. 4).

[Ber09] Aaron Bernstein. "Fully Dynamic $(2+\epsilon)$ Approximate All-Pairs Shortest Paths with Fast Query and Close to Linear Update Time". In: *Symposium on Foundations of Computer Science (FOCS)*. 2009, pp. 693–702 (cit. on p. 4).

[Ber16] Aaron Bernstein. "Maintaining Shortest Paths Under Deletions in Weighted Directed Graphs". In: *SIAM Journal on Computing* 45.2 (2016). Announced at STOC'13, pp. 548–574 (cit. on p. 4).

[BHS03] Surender Baswana, Ramesh Hariharan, and Sandeep Sen. "Maintaining All-Pairs Approximate Shortest Paths Under Deletion of Edges". In: *Symposium on Discrete Algorithms (SODA)*. 2003, pp. 394–403 (cit. on p. 4).

[BHS07] Surender Baswana, Ramesh Hariharan, and Sandeep Sen. "Improved decremental algorithms for maintaining transitive closure and all-pairs shortest paths". In: *Journal of Algorithms* 62.2 (2007). Announced at STOC'02, pp. 74–92 (cit. on p. 4).

[BR11] Aaron Bernstein and Liam Roditty. "Improved Dynamic Algorithms for Maintaining Approximate Shortest Paths Under Deletions". In: *Symposium on Discrete Algorithms (SODA)*. 2011, pp. 1355–1365 (cit. on p. 4).

[CCE13] Sergio Cabello, Erin W. Chambers, and Jeff Erickson. "Multiple-Source Shortest Paths in Embedded Graphs". In: *SIAM Journal on Computing* 42.4 (2013), pp. 1542–1571 (cit. on p. 16).

[CW16] Timothy M. Chan and Ryan Williams. "Deterministic APSP, Orthogonal Vectors, and More: Quickly Derandomizing Razborov-Smolensky". In: *Symposium on Discrete Algorithms (SODA)*. 2016 (cit. on p. 3).

[DI04] Camil Demetrescu and Giuseppe F. Italiano. "A New Approach to Dynamic All Pairs Shortest Paths". In: *Journal of the ACM* 51.6 (2004). Announced at STOC'03, pp. 968–992 (cit. on pp. 2, 4).

[DI06] Camil Demetrescu and Giuseppe F. Italiano. "Fully dynamic all pairs shortest paths with real edge weights". In: *Journal of Computer and System Sciences* 72.5 (2006). Announced at FOCS'01, pp. 813–837 (cit. on p. 4).

[ES81] Shimon Even and Yossi Shiloach. "An On-Line Edge-Deletion Problem". In: *Journal of the ACM* 28.1 (1981), pp. 1–4 (cit. on p. 4).

[FR06] Jittat Fakcharoenphol and Satish Rao. "Planar graphs, negative weight edges, shortest paths, and near linear time". In: *Journal of Computer and System Sciences* 72.5 (2006). Announced at FOCS'01, pp. 868–889 (cit. on p. 4).

[Gal12] François Le Gall. "Faster Algorithms for Rectangular Matrix Multiplication". In: *Symposium on Foundations of Computer Science (FOCS)*. 2012, pp. 514–523 (cit. on p. 3).

[HK01] Monika Rauch Henzinger and Valerie King. "Maintaining Minimum Spanning Forests in Dynamic Graphs". In: *SIAM Journal on Computing* 31.2 (2001). Announced at ICALP'97, pp. 364–374 (cit. on pp. 5, 6).

[HKN13] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. "Sublinear-Time Maintenance of Breadth-First Spanning Tree in Partially Dynamic Networks". In: *International Colloquium on Automata, Languages and Programming (ICALP)*. 2013, pp. 607–619 (cit. on p. 4).

[HKN14a] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. "A Subquadratic-Time Algorithm for Dynamic Single-Source Shortest Paths". In: *Symposium on Discrete Algorithms (SODA)*. 2014, pp. 1053–1072 (cit. on p. 4).

[HKN14b] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. "Sublinear-Time Decremental Algorithms for Single-Source Reachability and Shortest Paths on Directed Graphs". In: *Symposium on Theory of Computing (STOC)*. 2014, pp. 674–683 (cit. on p. 4).

[HKN16] Monika Henzinger, Sebastian Krinninger, and Danupon Nanongkai. "Dynamic Approximate All-Pairs Shortest Paths: Breaking the $O(mn)$ Barrier and Derandomization". In: *SIAM Journal on Computing* 45.3 (2016). Announced at FOCS'13, pp. 947–1006 (cit. on p. 4).

[HKR+97] Monika Rauch Henzinger, Philip N. Klein, Satish Rao, and Sairam Subramanian. "Faster Shortest-Path Algorithms for Planar Graphs". In: *Journal of Computer and System Sciences* 55.1 (1997). Announced at STOC'94, pp. 3–23 (cit. on p. 4).

[Joh77] Donald B. Johnson. "Efficient Algorithms for Shortest Paths in Sparse Networks". In: *Journal of the ACM* 24.1 (1977), pp. 1–13 (cit. on p. 13).

[Kin99] Valerie King. "Fully Dynamic Algorithms for Maintaining All-Pairs Shortest Paths and Transitive Closure in Digraphs". In: *Symposium on Foundations of Computer Science (FOCS)*. 1999, pp. 81–91 (cit. on p. 3).

[KS98] Philip N. Klein and Sairam Subramanian. "A Fully Dynamic Approximation Scheme for Shortest Paths in Planar Graphs". In: *Algorithmica* 22.3 (1998). Announced at WADS'93, pp. 235–249 (cit. on p. 4).

[LO88] Christos Levcopoulos and Mark H. Overmars. "A Balanced Search Tree with $O(1)$ Worst-case Update Time". In: *Acta Informatica* 26.3 (1988), pp. 269–277 (cit. on p. 5).

[RHK99] Monika Rauch Henzinger and Valerie King. "Randomized Fully Dynamic Graph Algorithms with Polylogarithmic Time per Operation". In: *Journal of the ACM* 46.4 (1999). Announced at STOC'95, pp. 502–516 (cit. on p. 16).

[RZ12] Liam Roditty and Uri Zwick. "Dynamic Approximate All-Pairs Shortest Paths in Undirected Graphs". In: *SIAM Journal on Computing* 41.3 (2012). Announced at FOCS'04, pp. 670–683 (cit. on p. 4).

[San04] Piotr Sankowski. "Dynamic Transitive Closure via Dynamic Matrix Inverse". In: *Symposium on Foundations of Computer Science (FOCS)*. 2004, pp. 509–517 (cit. on p. 3).

[San05] Piotr Sankowski. "Subquadratic Algorithm for Dynamic Shortest Distances". In: *International Computing and Combinatorics Conference (COCOON)*. 2005, pp. 461–470 (cit. on pp. 3, 4, 17).

[Tar97] Robert E. Tarjan. "Dynamic trees as search trees via Euler tours, applied to the network simplex algorithm". In: *Mathematical Programming* 77 (1997), pp. 169–177 (cit. on p. 16).

[Tho04] Mikkel Thorup. "Fully-Dynamic All-Pairs Shortest Paths: Faster and Allowing Negative Cycles". In: *Scandinavian Workshop on Algorithm Theory (SWAT)*. 2004, pp. 384–396 (cit. on p. 4).

[Tho05] Mikkel Thorup. "Worst-Case Update Times for Fully-Dynamic All-Pairs Shortest Paths". In: *Symposium on Theory of Computing (STOC)*. 2005, pp. 112–119 (cit. on pp. 1, 2, 4–6).

[TW05] Robert E. Tarjan and Renato F. Werneck. "Self-adjusting top trees". In: *Symposium on Discrete Algorithms (SODA)*. 2005, pp. 813–822 (cit. on p. 16).

[UY91] Jeffrey D. Ullman and Mihalis Yannakakis. "High-Probability Parallel Transitive-Closure Algorithms". In: *SIAM Journal on Computing* 20.1 (1991). Announced at SPAA'90, pp. 100–125 (cit. on p. 5).

[Wil14] Ryan Williams. "Faster all-pairs shortest paths via circuit complexity". In: *Symposium on Theory of Computing (STOC)*. 2014, pp. 664–673 (cit. on p. 3).

[Zwi02] Uri Zwick. "All Pairs Shortest Paths using Bridging Sets and Rectangular Matrix Multiplication". In: *Journal of the ACM* 49.3 (2002). Announced at FOCS'98, pp. 289–317 (cit. on pp. 3, 15).