Alcom-FT Technical Report Series
ALCOMFT-TR-03-110

Topology Matters: Smoothed Competitiveness of Metrical Task Systems*

Guido Schäfer *Naveen Sivadasan*

Max-Planck-Institut für Informatik, Saarbrücken, Germany
{schaefer,ns}@mpi-sb.mpg.de

Abstract

We consider metrical task systems, a general framework to model a large class of online problems. Borodin, Linial and Saks [6] gave a deterministic online algorithm, the *work function algorithm* (WFA), for metrical task systems. WFA has a competitive ratio of $2n-1$, where n is the number of nodes in the underlying graph. In fact, the competitive ratio of WFA is optimal: Every deterministic online algorithm for metrical task systems must have a competitive ratio of at least $2n-1$. Often, however, the competitive ratio is an over-pessimistic estimation of the performance of an online algorithm.

In this paper, we use the notion of *smoothed competitiveness* so as to capture the performance of WFA more adequately. We perturb, or *smoothen*, the request costs of a given task sequence by a probability distribution $\mathcal{P}$ and analyze the competitiveness of WFA on the perturbed sequence. Interstingly, our analysis reveals that the smoothed competitive ratio of WFA is much better and depends on several topological parameters of the underlying graph, such as the diameter, maximum degree, and minimum edge length.

Our analysis implies the following results for graphs with uniform edge lengths U, maximum degree D, and edge diameter $diam$.

(i) We provide the first average case analysis for WFA. We show that the expected competitive ratio of WFA is $O(\log(D))$, if the request costs are chosen randomly from an arbitrary non-increasing distribution $\mathcal{P}$ with standard deviation $\sigma = \Theta(U)$.

(ii) We prove that the smoothed competitive ratio of WFA is $O(diam(U/\sigma + \log(D)))$ and $O(\sqrt{n(U/\sigma + \log(D))})$, where σ denotes the standard deviation of the smoothing distribution. For example, already for perturbations with $\sigma = \Theta(U)$ the competitive ratio of WFA reduces to $O(\log(n))$ on a clique and to $O(\sqrt{n})$ on a line. Our analysis holds for a large class of probability distributions, including the uniform and the normal distribution.

(iii) We present lower bounds on the smoothed competitive ratio of every deterministic algorithm showing that the above bounds are tight for a large class of graphs. We also give lower bounds for arbitrary graphs.

We also extend our results to graphs with arbitrary edge lengths: For a graph with diameter $Diam$ and minimum edge length $U_{\min}$, WFA has a smoothed competitive ratio of $O(Diam/\sigma + (Diam/U_{\min})\log(n))$. This bound is tight up to a factor of at most $\log(n)$ for every deterministic algorithm.

*Partially supported by the Future and Emerging Technologies programme of the EU under contract number IST-1999-14186 (ALCOM-FT).

1 Introduction

We consider *metrical task systems*, a general framework formulated by Borodin, Linial and Saks [6] to model online problems. We are given an undirected and connected graph $G = (V, E)$, with node set V and edge set E, and a positive length function $\lambda : E \to \mathbb{R}^+$ on the edges of G. We define a distance function $\delta : V \times V \to \mathbb{R}^+$ on G such that $\delta(u, v)$ denotes the shortest path distance between any two nodes u and v in V. Let n denote the number of nodes in G. A task τ is an n-vector $(r(v_1), \ldots, r(v_n))$ of request costs, where the cost to process τ in node v_i is $r(v_i) \in \mathbb{R}^+ \cup \{\infty\}$. Given an initial state $s_0 \in V$, the online algorithm has to service a sequence $\mathcal{S} = \langle \tau_1, \ldots, \tau_r \rangle$ of tasks, one arriving at a time. If the online algorithm resides after task τ_{t-1} in node u, the cost to service task τ_t in node v is $\delta(u, v) + r_t(v)$; $\delta(u, v)$ is the transition cost and $r_t(v)$ is the processing cost. The objective is to minimize the total transition plus processing cost.

Many online problems can be formulated as metrical task systems, e.g., the paging problem, the static list accessing problem and the k-server problem.

Competitive Analysis. A measurement for the quality of an online algorithm is its *competitive ratio* [10], which is defined as the supremum over all input instances of the ratio between the cost of the online algorithm and the cost of an optimal offline algorithm. Borodin, Linial and Saks [6] gave a deterministic online algorithm, known as the *work function algorithm* (WFA), for metrical task systems. WFA has a competitive ratio of $2n - 1$, which is optimal. Borodin, Linial and Saks [6] and Manasse, McGeoch and Sleator [8] proved that *every* deterministic online algorithm for metrical task systems must have a competitive ratio of at least $2n - 1$. Put differently, if the quality of an online algorithm is measured in terms of its competitive ratio, then the work function algorithm (and any other deterministic algorithm) is not competitive for metrical task systems. On the other hand, the competitive ratio often is an over-pessimistic estimation of the performance of an online algorithm, since instances that force the online algorithm in its worst case behavior are highly artificial.

Smoothed Competitive Analysis. Spielman and Teng [11] proposed a new complexity measure, called *smoothed complexity*, which is a hybrid between average case and worst case complexity. The basic idea is to randomly perturb, or *smoothen*, the input instances and to analyze the performance of the algorithm on the perturbed instances. The smoothed complexity of an algorithm is defined as the supremum over all input instances of the expected running time on the perturbed instances. Intuitively, the smoothed complexity of an algorithm is small if the worst case instances are isolated peaks in the instance/time space.

It is natural to apply the smoothed complexity concept to other performance criteria. Becchetti et al. [3] were the first who considered *smoothed competitiveness* as an alternative to the traditional (worst case) competitiveness. The *smoothed competitive ratio* of an online algorithm ALG for a cost minimization problem is defined as follows. Let $\bar{\mathcal{S}}$ denote an input instance given by an adversary. We smoothen $\bar{\mathcal{S}}$ by some probability distribution $\mathcal{P}$. Let $\text{ALG}[\mathcal{S}]$ and $\text{OPT}[\mathcal{S}]$ denote the cost of the online and the optimal offline algorithm on a smoothed instance $\mathcal{S}$, respectively. Then, the smoothed competitive ratio c of ALG is defined as

$$c := \sup_{\bar{\mathcal{S}}} \mathbf{E}_{\mathcal{S} \xleftarrow{\mathcal{P}} \bar{\mathcal{S}}} \left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \right].$$

Our Contribution. We use the notion of smoothed competitiveness to characterize the performance of the work function algorithm. We smoothen the costs of the tasks according to an additive smoothing model: Each cost entry is smoothed by adding a random number chosen from a probability

Upper Bounds (uniform edge lengths)		
average case	$O\big(1+\frac{\sigma}{U}\cdot\log(D)\big)$	
arbitrary tasks	$O\big(diam\cdot\big(\frac{U}{\sigma}+\log(D)\big)\big)$	and $\quad O\Big(\sqrt{n\cdot\big(\frac{U}{\sigma}+\log(D)\big)}\Big)$
β**-elementary tasks**	$O\big(\beta\cdot\big(\frac{U}{\sigma}+\log(D)\big)\big)$	
Upper Bound (arbitrary edge lengths)		
arbitrary tasks	$O\big(\frac{Diam}{\sigma}+\frac{Diam}{U_{\min}}\cdot\log(n)\big)$	

Table 1: Upper bounds for WFA.

distribution $\mathcal{P}$, whose expectation coincides with the original cost entry. Spielman and Teng [11] used this smoothing model with $\mathcal{P}$ being the normal distribution. Our analysis holds for a large class of probability distributions, including, e.g., the uniform, normal and double exponential distribution. We use σ to refer to the standard deviation of $\mathcal{P}$.

Interestingly, the smoothed competitive ratio of WFA does not only depend on n and σ but also on certain *topological parameters* of the underlying graph, such as the maximum degree, the diameter, and the minimum edge length.

Graphs with Uniform Edge Lengths. We consider graphs with uniform edge lengths U, i.e., $\lambda(e)=U$ for all $e\in E$. We call these graphs *uniform graphs* subsequently. (Note that the distance function δ of G may not be uniform.) Let D be the maximum degree of a node in G and let $diam$ denote the *edge diameter* of G, i.e., the maximum length of a shortest (with respect to the number of edges) path between any two nodes.

We prove several upper bounds (see also Table 1):

(i) We show that if the request costs are chosen randomly according to some arbitrary distribution $\mathcal{P}$, which is non-increasing in $[0,\infty)$, the expected competitive ratio of WFA is $O(1+\sigma/U\cdot\log(D))$. In particular, WFA has an expected competitive ratio of $O(\log(D))$ if $\sigma=\Theta(U)$. For example, we obtain a competitive ratio of $O(\log(n))$ on a clique and of $O(1)$ on a binary tree.

(ii) We prove two upper bounds on the smoothed competitive ratio of WFA:

$$O\big(diam\cdot\big(\tfrac{U}{\sigma}+\log(D)\big)\big)\quad\text{and}\quad O\Big(\sqrt{n\cdot\big(\tfrac{U}{\sigma}+\log(D)\big)}\Big).$$

For example, on a clique WFA admits a smoothed competitive ratio of $O(U/\sigma+\log(n))$, which for $\sigma\le U/\log(n)$ is $O(U/\sigma)$ and for $\sigma\ge U/\log(n)$ becomes $O(\log(n))$. On a line, we obtain a smoothed competitive ratio of $O(\sqrt{n(U/\sigma+1)})$, which for $\sigma\le U$ becomes $O(\sqrt{n(U/\sigma)})$ and for $\sigma\ge U$ reduces to $O(\sqrt{n})$.

(iii) We obtain a better upper bound on the smoothed competitive ratio of WFA if the adversarial task sequence only consists of β*-elementary tasks*; a task is β-elementary if it has at most β non-zero entries. We prove a smoothed competitive ratio of $O(\beta\cdot(U/\sigma+\log(D)))$. For example, for elementary tasks, i.e., $\beta=1$, WFA has a smoothed competitive ratio of $O(U/\sigma+\log(n))$ on a clique and of $O(U/\sigma+1)$ on a line.

Lower Bounds (uniform edge lenghts)			
arbitrary tasks			
– existential	$\Omega\big(diam \cdot \big(\frac{U}{\sigma} + \log(D)\big)\big)$	and	$\Omega\Big(\sqrt{n \cdot \big(\frac{U}{\sigma} + \log(D)\big)}\Big)$
– universal	$\Omega\big(\frac{U}{\sigma} + \log(D)\big)$	and	$\Omega\Big(\sqrt{diam \cdot \big(\frac{U}{\sigma} + 1\big)}\Big)$
β**-elementary tasks**	$\Omega\big(\beta \cdot \big(\frac{U}{\sigma} + 1\big)\big)$ (existential)		

Lower Bound (arbitrary edge lengths)		
arbitrary tasks	$\Omega\big(\frac{Diam}{\sigma} + \frac{Diam}{U_{\min}} \cdot \epsilon \log(n)\big)$	(existential)

Table 2: Lower bounds for any deterministic algorithm.

Note that choosing σ much larger than U does not help to further reduce the expected or smoothed competitive ratio of the work function algorithm.

We also present some lower bounds (see Table 2). All lower bounds hold for *any* deterministic algorithm. We distinguish between *existential* and *universal* lower bounds. An existential lower bound, say $\Omega(f(n))$, means that there *exists* a class of graphs such that *every* deterministic algorithm has a smoothed competitive ratio of $\Omega(f(n))$ on these graphs. On the other hand, a universal lower bound $\Omega(f(n))$ states that for *any arbitrary* graph, *every* deterministic algorithm has a smoothed competitive ratio of $\Omega(f(n))$.

(iv) For a large range of values for $diam$ and D, we present existential lower bounds that match the bounds in (ii). This means (a) that the stated smoothed competitive ratio of WFA is tight, and (b) that WFA is optimal under the additive smoothing model—no other deterministic algorithm can achieve a better smoothed competitive ratio.

(v) We also prove two universal lower bounds on the smoothed competitive ratio:

$$\Omega\big(\tfrac{U}{\sigma} + \log(D)\big) \quad \text{and} \quad \Omega\Big(\sqrt{diam \cdot \big(\tfrac{U}{\sigma} + 1\big)}\Big).$$

(vi) For β-elementary tasks, we prove an existential lower bound of $\Omega(\beta \cdot (U/\sigma + 1))$. This implies that the bound in (iii) is tight up to a factor of $\log(D)$.

Graphs with Arbitrary Edge Lengths. We also present upper and lower bounds on the smoothed competitive ratio of WFA on graphs with arbitrary edge lengths.

(vii) For a graph with diameter $Diam$ and minimum edge length $U_{\min}$ the work function algorithm has a smoothed competitive ratio of $O((Diam/\sigma) + (Diam/U_{\min}) \cdot \log(n))$. For example, consider a clique with arbitrary edge lengths such that the ratio $Diam/U_{\min}$ is bounded by some constant. Then, the smoothed competitive ratio of WFA is $O(U_{\min}/\sigma + \log(n))$.

(viii) We present an existential lower bound, which shows that the stated upper bound is tight up to a factor of at most $\log(n)$.

Constrained Balls into Bins Game. Our analysis crucially relies on a lower bound for the cost of an optimal offline algorithm. We therefore have to study the growth of the work function values over time on a sequence of random requests. It turns out that the increase in the work function values can be modeled by a version of a balls into bins game with dependencies between the height of the bins, which are specified by a constraint graph. We call this game the *constrained balls into bins game*. The dependencies between the height of the bins make it difficult to analyze this stochastic process. We believe that the constrained balls into bins game is also interesting independently of the context of this paper.

Related Work. Several other attempts were made in the past to overcome the over-pessimistic estimation of the performance of an online algorithm by its competitive ratio. One idea was to enhance the capability of the online algorithm by allowing a limited lookahead [1, 2]. Another idea was to restrict the power of the adversary. For example, Borodin et al. [5] used an access graph model to restrict the input sequences in online paging problems to specific patterns. The diffuse adversary model by Koutsoupias and Papadimitriou [7] is another attempt to refine the notion of competitiveness. In this model, the actual distribution of the input is chosen by an adversary from a known class of possible distributions. We strongly believe that smoothed competitive analysis is a natural alternative to adequately characterize the performance guarantee of an online algorithm.

Organization of the Paper. In Section 2, we first review the work function algorithm and state some of its properties. In Section 3, we define the smoothing model that we use. The lower bound on the cost of an optimal offline algorithm and the related balls into bins game are presented in Section 4. Then, in Section 5, we prove the upper bounds on the expected and the smoothed competitive ratio of the work function algorithm on uniform graphs. We show the lower bounds for uniform graphs in Section 6. The improved bounds for β-elementary tasks are presented in Section 7. Finally, we extend our results to non-uniform graphs in Section 8.

Due to the lack of space, most of the proofs were put in the appendix; we therefore refer the reader to the appendix.

2 Work Function Algorithm

Let $\mathcal{S} = \langle \tau_1, \ldots, \tau_r \rangle$ be a request sequence and let s_0 denote the initial position. Let $\mathcal{S}_t$ denote the subsequence of the first t tasks of $\mathcal{S}$. For each node u, we define $w_t(u)$ as the minimum (offline) cost to process $\mathcal{S}_t$ starting in s_0 and ending in u. The function $w_t(\cdot)$ is called the *work function* of $\mathcal{S}$ at time t with respect to s_0.

Let OPT denote an optimal offline algorithm. Clearly, the optimal offline cost of the whole sequence is equal to the minimum work function value of $\mathcal{S}$ at time r, i.e., $\text{OPT}[\mathcal{S}] = \min_u\{w_r(u)\}$. The value of $w_t(u)$ can be computed via dynamic programming as follows:

$$w_0(u) = \delta(s_0, u), \quad \text{and} \quad w_t(u) = \min_{v \in V}\{w_{t-1}(v) + r_t(v) + \delta(v, u)\}. \tag{1}$$

We next describe the online work function algorithm, see also [6, 4]. Intuitively, a good strategy for an online algorithm to process task τ_t is to move to a node where OPT would reside when τ_t would be the final task. However, the competitive ratio of an algorithm that sticks to this policy can become arbitrarily bad. A slight modification gives a $2n - 1$ competitive algorithm: Instead of blindly (no matter at what cost) traveling to the node of minimum work function value, we additionally take the transition cost into account. Essentially, this is the idea underlying the work function algorithm which is defined more precisely as follows.

Work Function Algorithm (WFA): Let $s_0, \ldots, s_{t-1}$ denote the sequence of states visited by WFA to process $\mathcal{S}_{t-1}$. Then, to process task τ_t, WFA moves to a state s_t that minimizes $w_t(v) + \delta(s_{t-1}, v)$ for all $v \in V$. There is always a choice for s_t such that in addition $w_t(s_t) = w_{t-1}(s_t) + r_t(s_t)$. More formally,

$$s_t = \arg\min_{v \in V}\{w_t(v) + \delta(s_{t-1}, v)\} \quad \text{such that} \quad w_t(s_t) = w_{t-1}(s_t) + r_t(s_t). \tag{2}$$

In the sequel, we use WFA and OPT, respectively, to denote the online and the optimal offline algorithm. For a given sequence $\mathcal{S} = \langle \tau_1, \ldots, \tau_r \rangle$ of tasks, WFA$[\mathcal{S}]$ and OPT$[\mathcal{S}]$ refer to the cost incurred by WFA and OPT on $\mathcal{S}$, respectively. By $s_0, \ldots, s_r$ we denote the sequence of nodes visited by WFA. We call a task *β-elementary* if the number of non-zero entries is at most β, $\beta \geq 1$. We will also use the term *elementary task* to refer to a 1-elementary task.

We continue by observing a few properties of work functions and of the online algorithm WFA (see Appendix A for the corresponding proofs).

Fact 1. *For any node u and any time t, $w_t(u) \geq w_{t-1}(u)$.*

Fact 2. *For any node u and any time t, $w_t(u) \leq w_{t-1}(u) + r_t(u)$.*

Fact 3. *For any two nodes u and v and any time t, $|w_t(u) - w_t(v)| \leq \delta(u, v)$.*

Fact 4. *Let $s_0, \ldots, s_r$ denote the sequence of nodes visited by* WFA. *Then, at any time t, $w_t(s_t) = w_t(s_{t-1}) - \delta(s_{t-1}, s_t)$.*

3 Smoothing Models

Let the *adversarial task sequence* be given by $\bar{\mathcal{S}} = \langle \bar{\tau}_1, \ldots, \bar{\tau}_r \rangle$. We smoothen each task vector $\bar{\tau}_t = (\bar{r}_t(v_1), \ldots, \bar{r}_t(v_n))$ by perturbing each *original cost* entry $\bar{r}_t(v_j)$ according to some probability distribution $\mathcal{P}$. The obtained *smoothed* task is denoted by $\tau_t = (r_t(v_1), \ldots, r_t(v_n))$. We use μ and σ, respectively, to denote the expectation and the standard deviation of $\mathcal{P}$. We assume that $\mathcal{P}$ is symmetric around $\mu = 0$.

Next, we define two different smoothing models. Essentially, these are the models discussed by Spielman and Teng in [11]. For notational convenience, in the discussion below, we assume that the adversarial task is given by $\bar{\tau} = (\bar{r}(v_1), \ldots, \bar{r}(v_n))$ and refer to the smoothed task by $\tau = (r(v_1), \ldots, r(v_n))$.

(i) *Relative Smoothing Model.* In the relative smoothing model each cost entry is smoothed relatively to its original value.

$$r(v_i) = \max\{0, \bar{r}(v_i)\,(1 + \epsilon(v_i))\}, \text{ where } \epsilon(v_i) \leftarrow \mathcal{P}.$$

That is, the magnitude of distortion depends on the original value. Moreover, zero cost entries are retained.

(ii) *Additive Smoothing Model.* In this model each original cost entry is smoothed by

$$r(v_i) = \max\{0, \bar{r}(v_i) + \epsilon(v_i)\}, \text{ where } \epsilon(v_i) \leftarrow \mathcal{P}.$$

That is, to each original cost entry we add a random number which is chosen from $\mathcal{P}$. Observe that the magnitude of distortion does not depend on the original value but is the same for all values.

Observe that we take the maximum of zero and the smoothing outcome since we have to assure that the smoothed costs are non-negative. Thus, the probability for an original zero cost entry to remain zero is amplified to $\frac{1}{2}$.

Spielman and Teng use the additive smoothing model in their analysis in [11]. The major criticism to this model is that zero entries are destroyed. We also use the additive smoothing model. However, as we will argue in the next paragraph, concerning metrical task systems it is *crucial* that zero cost entries are destroyed.

Lower Bound for the Relative Smoothing Model. The proof of the $2n-1$ lower bound on the competitive ratio of any deterministic algorithm, see [6, 8, 4], is based only on the use of elementary tasks and the fact that the cost of the online algorithm is monotone increasing with the length of the input sequence. In the relative smoothing model zero cost entries are invariant to the smoothing, i.e., elementary tasks are smoothed to elementary tasks. In particular, this means that the above two properties still hold and therefore the lower bound proof also goes through for sequences smoothed according to the relative model.

Theorem 1. *Every deterministic online algorithm* ALG *for metrical task systems has a smoothed competitive ratio of at least* $2n-1$ *under the relative smoothing model.*

Adversary. We have two types of adversaries. The *oblivious adversary* specifies a task sequence based on the knowledge of the online algorithm and the smoothing distribution $\mathcal{P}$. That is, the adversary specifies the entire task sequence, which is then smoothed and presented to the online algorithm. We also consider a stronger adversary. The *adaptive adversary* reveals the task sequence over time, thereby taking decisions made by the online algorithm in the past into account. That is, the adaptive adversary might change its strategy depending on the behavior of the online algorithm in the past.

Our Model. We use the additive smoothing model as defined above. Our analysis holds for a large class of probability distributions, which we call *permissible*. We say $\mathcal{P}$ is permissible if (i) $\mathcal{P}$ is symmetric around $\mu = 0$, and (ii) $\mathcal{P}$ is non-increasing in $[0, \infty)$. For example, the uniform, the normal, and the double exponential distribution are permissible. The concentration of $\mathcal{P}$ around μ is given by the standard deviation σ. We already mentioned above that the smoothed competitive ratio does not improve further by choosing σ much larger than U. We therefore assume that $\sigma \leq 2U$ if we consider uniform graphs, and $\sigma \leq 2U_{\min}$ if we consider arbitrary graphs; these assumptions are not a principle limit of the analysis (any other constant instead of 2 would be fine too), but will turn out to be convenient throughout the analysis. Moreover, we use $c_{\mathcal{P}}$ to denote a constant depending on $\mathcal{P}$ such that for a random ϵ chosen from $\mathcal{P}$, $\mathbf{P}[|\epsilon| \geq \sigma / c_{\mathcal{P}}] \geq \frac{1}{2}$.

All our results hold against an adaptive adversary.

4 A Lower Bound for Offline Algorithms

In this section, we show a lower bound on the cost incurred by an optimal offline algorithm OPT when run on tasks smoothed according to the additive smoothing model.

The following lemma states that an optimal offline algorithm incurs with high probability a cost of at least $n\gamma U$ on a sequence of $\Theta\big(n\gamma\left(\frac{U}{\sigma} + \log(D)\right)\big)$ tasks.

Lemma 1. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. Let $\bar{\mathcal{S}}$ be an adversarial sequence of $\ell = \lceil c_2 n\gamma\left(\frac{U}{\sigma} + \log(D)\right)\rceil$ tasks, for a fixed constant c_2 and*

some $\gamma \geq 1$. Then, the cost incurred by OPT *on a smoothed task sequence $\mathcal{S}$ is at least $n\gamma U$ with high probability; more precisely, we have: $\mathbf{P}[\text{OPT}[\mathcal{S}] < n\gamma U] \leq 1/n^3$.*

Although the lemma is formulated for uniform graphs, we can also use it for graphs with arbitrary edge lengths: Given a graph G with arbitrary edge lengths, let $U_{\min}$ denote the minimum edge length of an edge in G. If we replace each edge length in G by $U_{\min}$, we obtain a uniform graph G' with $U = U_{\min}$. Clearly, for any task sequence the cost incurred by OPT on G' is at most the cost incurred by OPT on G.

For the purpose of proving Lemma 1, we investigate an interesting version of a balls into bins experiment, which we call the *constrained balls into bins game*.

Constrained Balls into Bins Game. We are given n bins $B_1, \ldots, B_n$. In each round, we place a ball (independently) in each bin B_i with probability p; with probability $1 - p$ no ball is placed in B_i. We define the *height* of a bin B_i in a round as the number of balls in B_i. We have dependencies between the height of different bins, which are specified by an (undirected) *constraint graph* $G_c = (V_c, E_c)$. The node set V_c of G_c contains n nodes $\{u_1, \ldots, u_n\}$, where each node u_i corresponds to a bin B_i. All edges in E_c have uniform edge lengths equal to Q. Let D be the maximum degree of a vertex in G_c. Throughout the experiment, we maintain the following invariant: The difference in height between two bins B_i and B_j is at most the shortest path distance between v_i and v_j in G_c. If the placement of a ball into a bin B_i would violate this invariant, the ball is *rejected*; otherwise we say that the ball is *accepted*. Observe that if two bins B_i and B_j do not violate the invariant in round i, then, in round $i + 1$, B_i and B_j might cause a violation only if one bin, say B_i, receives a ball, and the other, B_j, does not receive a ball; if both receive a ball, or both do not receive a ball, the invariant remains true.

Theorem 2. *Fix any bin B_z. Then, for $h \geq \log(n)$ there is a constant c_3, which only depends on p, such that after $c_3 h\,(1 + \log(D)/Q)$ rounds the height of B_z is at least h with high probability. More precisely, if R_z is a random variable denoting the number of rounds needed until the height of B_z becomes h, we have: $\mathbf{P}[R_z \geq c_3 h\,(1 + \log(D)/Q)] \leq 1/n^4$.*

We remark that there exist constraint graphs with $Q = 1$, e.g., a clique on n nodes, such that the expected number of rounds needed for the height of a bin to become h is $\Omega(h \log(n))$. Moreover, for any given maximum degree D, one can create graph instances with $Q = 1$ such that the expected number of rounds is $\Omega(h \log(D))$.

We next describe how one can model the growth of the height of B_z by an alternative, essentially equivalent, game on a *layered dependency graph*. A layered dependency graph $\mathcal{D}_h$ consists of h layers, $V_1, \ldots, V_h$, and edges are present only between adjacent layers.

We first describe the construction for $Q = 1$: Each layer of $\mathcal{D}_h$ corresponds to a subset of nodes in G_c. Layer 1 consists of z only, the node corresponding to bin B_z. Assume we have constructed layers $V_1, \ldots, V_i$, $i < h$. Then, V_{i+1} is constructed from V_i by adding all nodes, $\Gamma_{G_c}(V_i)$, that are adjacent to V_i in G_c, i.e., $V_{i+1} := V_i \cup \Gamma_{G_c}(V_i)$. For every pair $(u, v) \in V_i \times V_{i+1}$, we add an edge (u, v) to $\mathcal{D}_h$ if $(u, v) \in E_c$, or $u = v$.

Now, the following game on $\mathcal{D}_h$ is equivalent to the balls and bins game. Each node in $\mathcal{D}_h$ is in one of three states, namely UNFINISHED, READY or FINISHED. Initially, all nodes in layer h are READY and all other nodes are UNFINISHED. In each round, all READY nodes independently toss a coin; the coin turns up *head* with probability p and *tail* with probability $1-p$. A READY node changes its state to FINISHED if the outcome of its coin toss is *head*. At the end of each round, a node in layer j changes its state to READY, if all its neighbors in layer $j + 1$ are FINISHED. Note that the number

of rounds needed until the root node z in $\mathcal{D}_h$ becomes FINISHED is exactly the same as the number of rounds needed for the height of B_z to become h.

We use a similar construction if $Q > 1$. For simplicity, let h be a multiple of Q and define $h' = h/Q$. We construct a dependency graph $\mathcal{D}_{h'}$ with h' layers as described above (replace h by h' in the description above). Then, we transform $\mathcal{D}_{h'}$ into a layered graph $\mathcal{D}_h$ with h layers as follows. Let v be a node in layer j of $\mathcal{D}_{h'}$. We replace v by a path $(v_1, \ldots, v_k)$, where $k = |Q|$. v_1 is connected to all neighbors of v in layer $j-1$ and v_k is connected to all neighbors of v in layer $j+1$. This replacement makes sure that the number of rounds needed until the root node becomes FINISHED in $\mathcal{D}_h$ dominates the number of rounds needed for the height of B_z to become h.

Proof (Theorem 2). We construct a layered dependency graph $\mathcal{D}_h$ as described above. As argued above, the event $(R_z < t)$ is dominated by the event that the root node becomes FINISHED within t rounds in $\mathcal{D}_h$. Consider the event that the root node z does not become FINISHED after t rounds. Then, there must exist a *bad* path $P = (v_1, \ldots, v_h)$ from $z = v_1$ to some node v_h in the bottom layer h such that no node v_i of P was delayed by nodes other than $v_{i+1}, \ldots, v_h$. Put differently, P was delayed independently of any other path. Consider the outcome of the coin flips only for the nodes along P. If P is bad, the number of coin flips, X, that turned up *head* within t rounds for the nodes of P is at most $h-1$. Let $\alpha(t)$ denote the probability that P is bad, i.e., $\alpha(t) = \mathbf{P}[X \leq h-1]$. Clearly, $\mathbf{E}[X] = tp$.

Observe that $\mathcal{D}_h$ satisfies (i) at most h' layers have nodes with degree larger than 2, and (ii) every node has at most $D+1$ neighbors in the next larger layer. That is, the number of possible paths from z to some node v_h in layer h is bounded by $(D+1)^{h'}$.

Thus, $\mathbf{P}[R_z \geq t] \leq \alpha(t)(D+1)^{h'}$. We want to choose t such that this probability is at most $1/n^4$. If we choose $t \geq (32/p)(h + h' \log(D))$, and use a Chernoff bound (see, e.g., [9]) on X, we obtain

$$\alpha(t) = \mathbf{P}[X \leq h-1] \leq \mathbf{P}[X \leq pt/2] \leq e^{-pt/8} \leq \frac{1}{n^4(D+1)^{h'}}.$$

□

5 Upper Bounds for the Work Function Algorithm on Uniform Graphs

Average Case Analysis. We derive an upper bound on the expected competitive ratio of WFA if each request cost is chosen independently from a probability distribution $\mathcal{P}$, which is non-increasing in $[0, \infty)$.

Theorem 3. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. Then, the expected competitive ratio of* WFA *is $O(1 + (U/\sigma) \cdot \log(D))$.*

Smoothed Competitive Analysis. We derive two upper bounds on the smoothed competitive ratio of WFA for arbitrary adversarial task sequences on graphs with uniform edge lengths.

Theorem 4. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. Then, the smoothed competitive ratio of* WFA *is $O(\sqrt{n\,(U/\sigma + \log(D))})$.*

Proof (sketch). We compare the cost of WFA with two different lower bounds for OPT. The first lower bound is given below; the second lower bound follows from Lemma 1.

Fix any adversarial input sequence $\bar{S}$ of length ℓ, where $\ell = \lceil c_2 n\gamma\,(U/\sigma + \log(D))\rceil$ as defined in Lemma 1. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the smoothed sequence. For any deterministic sequence $\mathcal{K}$, let $\text{WFA}[\mathcal{K}] = \sum_{t=1}^{\ell} C(t)$, where $C(t) := r_t(s_t) + \delta(s_{t-1}, s_t)$. Let $s_0, s_1, \ldots, s_\ell$

denote the sequence of nodes visited by WFA. For sequence $\mathcal{K}$, let $T := \{t \mid w_{t-1}(s_{t-1}) - w_t(s_t) \leq \delta(s_{t-1}, s_t)/2\}$. Let $diam$ denote the edge diameter of the graph. We further split T into T^1 and T^2, where $T^1 := \left\{t \,\middle|\, w_t(s_t) - w_{t-1}(s_t) \;\leq\; 4\,U\,diam\right\}$ and $T^2 = T \setminus T^1$.

Lemma 2. *For every sufficiently long deterministic sequence $\mathcal{K}$,* $\text{WFA}[\mathcal{K}] \;\leq\; 8\sum_{t\,\in\,T} C(t)$.

Lemma 3. *For every sufficiently long deterministic sequence $\mathcal{K}$, and for a fixed constant b,* $\text{OPT}[\mathcal{K}] \geq \frac{1}{bn}\big(\frac{1}{U}\sum_{t\in T^1} C(t)^2 + n\sum_{t\in T^2} C(t)\big)$.

Let $\mathcal{E}$ denote the event that $\text{OPT}[\mathcal{S}]$ is at least $n\gamma U$. Due to Lemma 1 we have, $\mathbf{P}[\mathcal{E}] \;\geq\; 1 - 1/n^3$. Therefore,

$$\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \;=\; \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\;\middle|\;\mathcal{E}\right]\,\mathbf{P}[\mathcal{E}] + \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\;\middle|\;\neg\mathcal{E}\right]\,\mathbf{P}[\neg\mathcal{E}] \;\leq\; \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\;\middle|\;\mathcal{E}\right] + \frac{2n-1}{n^3}.$$

We can therefore concentrate on bounding $\mathbf{E}[\text{WFA}[\mathcal{S}]/\text{OPT}[\mathcal{S}] \mid \mathcal{E}]$. Let $\mathcal{S}_\mathcal{E}$ be any deterministic outcome of smoothing $\bar{\mathcal{S}}$ conditioned on $\mathcal{E}$. By the definition of $\mathcal{E}$ and by Lemma 2, we have

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\leq\; \frac{8\sum_{t\in T} C(t)}{n\gamma U}$$

By the above two lemmas, and by the definition of T, T^1 and T^2 for $\mathcal{S}_\mathcal{E}$, we also have,

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\leq\; 8b\,n\left(\frac{\sum_{t\in T^1} C(t) \;+\; \sum_{t\in T^2} C(t)}{\frac{1}{U}\sum_{t\in T^1} C(t)^2 \;+\; n\sum_{t\in T^2} C(t)}\right).$$

If $\sum_{t\in T^1} C(t) \;<\; \sum_{t\in T^2} C(t)$, the above equation reduces to $O(1)$ and we are done. Assume $\sum_{t\in T^1} C(t) \geq \sum_{t\in T^2} C(t)$. Since $|T| \geq 1$, by applying Lemma 9 (Appendix O) on the above two bounds, we obtain $\text{WFA}[\mathcal{S}_\mathcal{E}]/\text{OPT}[\mathcal{S}_\mathcal{E}] \;=\; O(\sqrt{n\,(U/\sigma + \log(D))})$. We conclude that $\mathbf{E}[\text{WFA}[\mathcal{S}]/\text{OPT}[\mathcal{S}] \mid \mathcal{E}] \;=\; O(\sqrt{n\,(U/\sigma \;+\; \log(D))})$. □

Theorem 5. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, maximum degree D, and edge diameter $diam$. Then, the smoothed competitive ratio of* WFA *is $O(diam \cdot (U/\sigma + \log(D)))$.*

6 Lower Bounds on the Smoothed Competitiveness for Uniform Graphs

Existential Lower Bounds. We show that for a large range of parameters $diam$ and D, there exists a class of graphs on which *any* deterministic algorithm has a smoothed competitive ratio matching the upper bounds stated in Theorems 4 and 5.

Theorem 6. *There exist a class of graphs such that the smoothed competitive ratio of any deterministic algorithm* ALG *is* $\Omega(\min\{n, \sqrt{n\,(U/\sigma + \log(D))}\})$.

The next bound shows that the upper bound of Theorem 5 is tight up to a factor of $\log(D)/\log(\mathcal{D}) \leq \log(n)$.

Theorem 7. *There exists a class of graphs such that the smoothed competitive ratio of any deterministic algorithm* ALG *is* $\Omega(\min\{n, diam\,(U/\sigma + \log(\mathcal{D}))\})$, *where* $\mathcal{D} = \min\{diam, D\} - 1$.

Universal Lower Bounds. We derive two universal lower bounds on the smoothed competitive ratio of any deterministic algorithm.

Theorem 8. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. Then,* any *deterministic algorithm* ALG *has a smoothed competitive ratio of* $\Omega(\min\{n, U/\sigma + \log(D)\})$.

Theorem 9. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, maximum degree D, and edge diameter $diam$. Then,* any *deterministic algorithm* ALG *has a smoothed competitive ratio of* $\Omega(\min\{diam, \sqrt{diam \cdot (U/\sigma + 1)}\})$.

These two lower bounds together imply the following. By Theorem 8, the smoothed competitive ratio of any deterministic algorithm can never be better than $\Omega(\log(D))$. Depending on the ratio U/σ, we might obtain stronger lower bounds. If $U/\sigma \leq 1$, we obtain a bound of $\Omega(\sqrt{diam})$ by Theorem 9. If $1 < U/\sigma \leq diam$, we have a bound of $\Omega(\sqrt{diam \cdot U/\sigma})$ by Theorem 9. Finally, if $diam < U/\sigma \leq n$, we obtain a lower bound of $\Omega(U/\sigma)$ by Theorem 8.

7 Better Bounds for the Work Function Algorithm on Uniform Graphs

We can strengthen the upper bound on the smoothed competitive ratio of WFA, if the adversarial task sequence only consists of β-elementary tasks.

Theorem 10. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. If the adversarial task sequence only consists of β-elementary tasks, then the smoothed competitive ratio of* WFA *is $O(\beta\,(U/\sigma + \log(D)))$.*

Proof. The theorem follows from the lower bound on OPT in Lemma 1 and from Lemma 8 below. □

The smoothed competitive ratio of WFA stated in the above theorem yields a better upper bound than those stated in Section 5 if $\beta \leq \sqrt{n/(U/\sigma + \log(D))}$.

Lower Bound. We state the following existential lower bound.

Theorem 11. *Let G be a line graph with n nodes and uniform edge lengths U. If the adversarial task sequence only consists of β-elementary tasks, where $\beta \leq \sqrt{n/(U/\sigma + 1)}$, the smoothed competitive ratio of any deterministic online algorithm* ALG *is* $\Omega(\min\{n, \beta\,(U/\sigma + 1)\})$.

8 Work Function Algorithm on Non-Uniform Graphs

We analyze the smoothed competitive ratio of WFA if the underlying graph has *arbitrary* edge lengths. We present matching upper and lower bounds.

In this section, we use $Diam$ to refer to the diameter of the underlying graph, i.e., for any two nodes in the graph there exists a path of length at most $Diam$.

Theorem 12. *Let G be an arbitrary graph with n nodes, minimum edge length U_{min}, and diameter $Diam$. Then, the smoothed competitive ratio of* WFA *is $O((Diam/\sigma) + (Diam/U_{\min}) \cdot \log(n))$.*

Theorem 13. *There exists a graph G with n nodes, minimum edge length $U_{\min}$, and diameter $Diam > 2U_{\min}$, such that the smoothed competitive ratio of* any *deterministic algorithm* ALG *is*

$$\Omega\left(\min\left\{n^{1-\epsilon}, \frac{Diam}{\sigma} + \epsilon \cdot \frac{Diam}{U_{\min}} \cdot \log(n)\right\}\right),$$

for any constant ϵ with $0 < \epsilon < 1$ and for $\sigma \leq 2U_{\min}$.

Acknowledgements

We would like to thank Alberto Marchetti-Spaccamela for comments on a first draft of this paper. Moreover, we would like to thank Kurt Mehlhorn for helpful suggestions and discussions.

References

[1] Susanne Albers, *The influence of lookahead in competitive paging algorithms (extended abstract)*, European Symposium on Algorithms, 1993, pp. 1–12.

[2] ———, *A competitive analysis of the list update problem with lookahead*, Theoretical Computer Science **197** (1998), no. 1–2, 95–109.

[3] L. Becchetti, S. Leonardi, A. Marchetti-Spaccamela, G. Schäfer, and T. Vredeveld, *Average case and smoothed competitive analysis of the multi-level feedback algorithm*, Foundations of Computer Science, 2003, to appear.

[4] A. Borodin and R. El-Yaniv, *Online computation and competitive analysis*, Cambridge University Press, 1998.

[5] A. Borodin, S. Irani, P. Raghavan, and B. Schieber, *Competitive paging with locality of reference*, Journal of Computer and System Sciences **50** (1995), no. 2, 244–258.

[6] A. Borodin, N. Linial, and M. Saks, *An optimal online algorithm for metrical task systems*, Journal of the ACM **39** (1992), 745–763.

[7] E. Koutsoupias and C. Papadimitriou, *Beyond competitive analysis*, Proceedings of the Twenty-Fifth Symposium on Foundations of Computer Science, 1994, pp. 394–400.

[8] M. S. Manasse, L. A. McGeoch, and D. D. Sleator, *Competitive algorithms for on-line problems*, ACM Symposium on Theory of Computing, 1988, pp. 322–333.

[9] R. Motwani and P. Raghavan, *Randomized algorithms*, 1 ed., Cambridge University Press, 1995.

[10] D. Sleator and R. E. Tarjan, *Amortized efficiency of list update and paging rules*, Communications of the ACM **28** (1985), 202–208.

[11] D. Spielman and S. H. Teng, *Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time*, ACM Symposium on Theory of Computing, 2001.

A Proof of Fact 3 and Fact 4

Proof (Fact 3). Assume x is the node that defines $w_t(v)$, i.e., $w_t(v) = w_{t-1}(x)+r_t(x)+\delta(x,v)$. We have $w_t(u) \leq w_{t-1}(x)+r_t(x)+\delta(x,u) \leq w_{t-1}(x)+r_t(x)+\delta(x,v)+\delta(v,u) = w_t(v)+\delta(v,u)$. □

Proof (Fact 4). By (2), we have that $w_t(s_t) + \delta(s_{t-1}, s_t) \leq w_t(v) + \delta(s_{t-1}, v)$ for all $v \in V$, in particular, for $v = s_{t-1}$ this implies $w_t(s_t) \leq w_t(s_{t-1}) - \delta(s_{t-1}, s_t)$. On the other hand, due to Fact 3, $w_t(s_t) \geq w_t(s_{t-1}) - \delta(s_{t-1}, s_t)$. □

B Proof of Lemma 1

Proof (Lemma 1). The quantity $\min_v\{w_\ell(v)\}$ is a lower bound on the cost $\text{OPT}[\mathcal{S}]$ incurred by the optimal offline algorithm on a smoothed sequence $\mathcal{S}$. That is, it suffices to prove that $\mathbf{P}[\min_v\{w_\ell(v)\} < n\gamma U] \leq 1/n^3$. Observe that we can assume that the initial work function values are all set to zero, since this can only reduce the cost of OPT.

Fix a round t and consider a smoothed task $\tau_t = (r_t(v_1), \ldots, r_t(v_n))$. Our smoothing distribution satisfies $\mathbf{P}[|\epsilon| \geq \sigma/c_{\mathcal{P}}] \geq \frac{1}{2}$ and thus $\mathbf{P}[\epsilon \geq \sigma/c_{\mathcal{P}}] \geq \frac{1}{4}$. That is, each smoothed cost $r_t(v_i)$ is at least $\sigma/c_{\mathcal{P}}$ (independently) with probability at least $\frac{1}{4}$, *irrespective* of its original cost entry. We model the growth of the work function values by the constrained balls into bins game with $p = \frac{1}{4}$. The relation between the two experiments is as follows: In round t, we throw a ball into bin B_i iff the request cost $r_t(v_i)$ of v_i is at least $\sigma/c_{\mathcal{P}}$.

Define $\Delta := \min\{U, \sigma/c_{\mathcal{P}}\}$. Let G_c be the graph obtained from G by setting all edge lengths to $Q := \lfloor U/\Delta \rfloor = \max\{1, \lfloor c_{\mathcal{P}} U/\sigma \rfloor\}$. By Lemma 4 given below, the number of rounds needed for the height of a bin B_i to become at least $h := 2n\gamma Q$ is an upper bound on the time needed for the work function value of v_i to become at least $n\gamma U$.

Now, by applying Theorem 2, we obtain that for any bin B_i, after at most $c_3 h(1 + \log(D)/Q)$ rounds, $\mathbf{P}[\text{height of } B_i < h] \leq 1/n^4$. Therefore, the probability that there exists a bin of height less than h is at most $1/n^3$. Finally, by fixing $c_2 := 4c_3\,\lceil c_{\mathcal{P}} \rceil$, we have $c_2 n\gamma(U/\sigma + \log(D)) \geq c_3 h(1 + \log(D)/Q)$, which concludes the proof. □

Lemma 4. *Let $n_t(i)$ denote the number of balls in bin B_i after t rounds. Then, for any bin B_i and for any $t \geq 0$, $w_t(v_i) \geq n_t(i)\,\Delta$.*

Proof. We proof the lemma by induction on the number of rounds t. For $t = 0$, the lemma clearly holds. (We can assume that the initial work function values are all zero.) Assume that the induction hypothesis holds after t rounds. We show that it remains true after $t + 1$ rounds. In round $t + 1$, if no ball is accepted in any bin then clearly the hypothesis remains true. Consider the case where in round $t + 1$ at least one ball is accepted for some bin B_i. By the induction hypothesis, we have $w_t(v_i) \geq n_t(i)\Delta$. Let v_k denote the node that determines the work function value $w_{t+1}(v_i)$, i.e.,

$$w_{t+1}(v_i) = w_t(v_k) + r_{t+1}(v_k) + \delta(v_i, v_k).$$

Assume that $v_k = v_i$. Then, the work function value of v_i increases by the request cost $r_{t+1}(v_i) \geq \Delta$, and since a ball was accepted in B_i, $r_{t+1}(v_i) \geq \Delta$. Thus, we have $w_{t+1}(v_i) \geq w_t(v_i) + \Delta \geq n_t(i)\Delta + \Delta$, and we are done.

Next, assume that $v_k \neq v_i$. Let r be the distance between v_i and v_k in the constraint graph. Since the ball was accepted in B_i, there is no conflict between B_i and B_k by accepting the ball in B_i, which means that

$$n_t(i) - n_t(k) \leq r - 1 + [\text{ball accepted in } B_k],$$

where "[*statement*]" is 1 if *statement* is true, and 0 otherwise. By multiplying both sides with Δ and rearranging terms, we obtain

$$(n_t(k) + r)\Delta \geq (n_t(i) + 1 - [\text{ball accepted in } B_k])\Delta.$$

Observe that $r\Delta \leq \delta(v_i, v_k)$ by the definition of r and the edge lengths Q of the constraint graph. Thus,

$$\begin{aligned} w_{t+1}(v_i) &\geq w_t(v_k) + \delta(v_i, v_k) \geq n_t(k)\Delta + r\Delta \\ &\geq (n_t(i) + 1 - [\text{ball accepted in } B_k])\Delta \geq (n_t(i) + 1)\Delta. \end{aligned}$$

□

C Proof of Theorem 3

Lemma 5. *Consider a sequence $\mathcal{S}$ of $\ell \geq nU/\sigma$ tasks, where each request is chosen randomly according to some non-increasing distribution $\mathcal{P}$ with standard deviation σ. Then, the expected cost, $\mathbf{E}[\text{WFA}[\mathcal{S}]]$, incurred by* WFA *on $\mathcal{S}$ is at most $b\ell\sigma$, for some constant b.*

Proof. The proof is similar to the proof of Lemma 8. Hence, we only discuss the differences. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the request sequence. Let the random variables $s_0, s_1, \ldots, s_\ell$ denote the sequence of nodes visited by WFA. We define a potential function Φ as $\Phi(t) = w_t(s_t) + nUt/\ell$. Clearly $\Phi(\ell) - \Phi(0) \geq 0$. Using arguments similar to those in the proof of Lemma 8, we obtain the following bound on the amortized costs

$$C_a(t) \leq 2\Delta(s_t) + 2\Delta(s_{t-1}) + nU/\ell,$$

where $\Delta(s_t) = \min_{v \in V} \{r_t(v) + \delta(v, s_t)\}$. We have that $\mathbf{E}[\Delta(s_t)] \leq \mathbf{E}[r_t(s_t)]$. By Fact 6, this implies that $\mathbf{E}[\Delta(s_t)] \leq \sqrt{12}\sigma$. Moreover, we can derive the same bound for $\mathbf{E}[\Delta(s_{t-1})]$. Hence, we obtain that $\mathbf{E}[C_a(t)] \leq c \cdot \sigma + nU/\ell$, for some constant $c \geq 1$. Finally, we conclude that $\mathbf{E}[\text{WFA}[\mathcal{S}]] \leq c\ell\sigma + nU \leq 2c\ell\sigma$, since $\ell \geq nU/\sigma$. Fixing b to $2c$ proves the lemma. □

Proof (Theorem 3). Let $\mathcal{S}$ be a random sequence generated from $\mathcal{P}$. We can use Lemma 1 to lower bound the cost of OPT: The generation of the random sequence is equivalent to smoothing (according to $\mathcal{P}$) an adversarial task sequence consisting of all-zero request vectors only. Here, we do not need that the distribution is symmetric around its mean. We only need that with probability at least $1/2$, the outcome is non-negative (which is true for $\mathcal{P}$). The analysis is in the same lines as the proof of Theorem 10. We define $\mathcal{E}$ as the event that OPT incurs a cost of at least $n\gamma U$ on $\mathcal{S}$. By Lemma 1, we have $\mathbf{P}[\mathcal{E}] \geq 1 - 1/n^3$. Thus,

$$\begin{aligned} \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] &= \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \,\middle|\, \mathcal{E}\right] \mathbf{P}[\mathcal{E}] + \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \,\middle|\, \neg\mathcal{E}\right] \mathbf{P}[\neg\mathcal{E}] \\ &\leq \frac{\mathbf{E}[\text{WFA}[\mathcal{S}] \mid \mathcal{E}]\ \mathbf{P}[\mathcal{E}]}{n\gamma U} + \frac{2n-1}{n^3} = O\Big(1 + \frac{\sigma}{U} \cdot \log(D)\Big). \end{aligned}$$

The second inequality is due to the definition of $\mathcal{E}$, the fact that the competitive ratio of WFA is $2n-1$, and Lemma 1. The last inequality follows from the definition of ℓ and by using Lemma 5 to obtain that $\mathbf{E}[\text{WFA}[\mathcal{S}] \mid \mathcal{E}]\ \mathbf{P}[\mathcal{E}] \leq \mathbf{E}[\text{WFA}[\mathcal{S}]] \leq b\ell\sigma$. □

D Proof of Theorem 4

Proof (Theorem 4). Consider any deterministic input sequence $\mathcal{K}$ of length ℓ. Let $s_0, s_1, \ldots, s_\ell$ denote the sequence of nodes visited by WFA. Let $\text{WFA}[\mathcal{K}] = \sum_{t=1}^{\ell} C(t)$, where $C(t) := r_t(s_t) + \delta(s_{t-1}, s_t)$ denotes the service cost plus the transition cost incurred by WFA in round t. With respect to $\mathcal{K}$, let T be the subset of rounds defined as

$$T := \left\{ t \;\middle|\; w_{t-1}(s_{t-1}) - w_t(s_t) \leq \frac{\delta(s_{t-1}, s_t)}{2} \right\}. \tag{3}$$

Let $\bar{T}$ denote the complement of T. Let $diam$ denote the edge diameter of the graph. We further split T into T^1 and T^2, where $T^1 := \left\{ t \;\middle|\; w_t(s_t) - w_{t-1}(s_t) \leq 4\,U\,diam \right\}$, and $T^2 = T \setminus T^1$.

We need the following to lemmas; their proofs are given below.

Lemma 2 *For every sufficiently long deterministic sequence $\mathcal{K}$, i.e., $\text{WFA}[\mathcal{K}] \geq 6nU$, we have $\text{WFA}[\mathcal{K}] \leq 8 \sum_{t \in T} C(t)$.*

Lemma 3 *For every sufficiently long deterministic sequence $\mathcal{K}$, i.e., $\text{WFA}[\mathcal{K}] \geq c_w n^2 U$ for some constant c_w, and for a fixed constant b, we have $\text{OPT}[\mathcal{K}] \geq \frac{1}{bn}\big(\frac{1}{U}\sum_{t\in T^1} C(t)^2 + n \sum_{t\in T^2} C(t)\big)$.*

Fix any adversarial input sequence $\bar{\mathcal{S}}$ of length ℓ. Let ℓ be defined as in Lemma 1, i.e., $\ell = \lceil c_2 n\gamma\,(U/\sigma + \log(D))\rceil$ and choose $\gamma \geq c_w nU$, for some constant $c_w \geq 6$ which will be fixed later. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the smoothed task sequence obtained from the adversarial sequence $\bar{\mathcal{S}}$. Let $\mathcal{E}$ denote the event that the cost of OPT on $\mathcal{S}$ is at least $n\gamma U$. Due to Lemma 1, we have $\mathbf{P}[\mathcal{E}] \geq 1 - 1/n^3$. Therefore,

$$\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] = \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \mathcal{E}\right] \mathbf{P}[\mathcal{E}] + \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \neg\mathcal{E}\right] \mathbf{P}[\neg\mathcal{E}] \leq \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \mathcal{E}\right] + \frac{2n-1}{n^3}.$$

We can therefore concentrate on bounding $\mathbf{E}[\text{WFA}[\mathcal{S}]/\text{OPT}[\mathcal{S}] \mid \mathcal{E}]$.

Let $\mathcal{S}_\mathcal{E}$ be any arbitrary deterministic outcome of smoothing $\bar{\mathcal{S}}$ such that the event $\mathcal{E}$ holds. Since we are considering sufficiently long sequences, i.e, $\gamma \geq c_w nU$, the cost of OPT and of WFA is at least $c_w n^2 U$. Hence by applying Lemma 2 and Lemma 3 and using the definition of T, T^1 and T^2 for $\mathcal{S}_\mathcal{E}$ as in (3),

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\leq\; 8b\,n \left(\frac{\sum_{t\in T^1} C(t) + \sum_{t\in T^2} C(t)}{\frac{1}{U}\sum_{t\in T^1} C(t)^2 + n\sum_{t\in T^2} C(t)} \right). \tag{4}$$

If $\sum_{t\in T^1} C(t) < \sum_{t\in T^2} C(t)$, the above equation simplifies to $\text{WFA}[\mathcal{S}_\mathcal{E}]/\text{OPT}[\mathcal{S}_\mathcal{E}] \leq 16b = O(1)$. Hence, we have $\mathbf{E}[\text{WFA}[\mathcal{S}]/\text{OPT}[\mathcal{S}] \mid \mathcal{E}] = O(1)$.

Next, assume $\sum_{t\in T^1} C(t) \geq \sum_{t\in T^2} C(t)$. By Lemma 2, we know that $|T| \geq 1$ since $\text{WFA}[\mathcal{S}_\mathcal{E}] \geq 6nU$. Hence in this case we also have that $\left|T^1\right| \geq 1$. We can rewrite (4) as

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\leq\; 16b\,n \left(\frac{\sum_{t\in T^1} C(t)}{\frac{1}{U}\sum_{t\in T^1} C(t)^2} \right). \tag{5}$$

Moreover, by the definition of $\mathcal{E}$ and Lemma 2, we also have

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \leq \frac{8\sum_{t\in T} C(t)}{n\gamma U} \leq \frac{16\sum_{t\in T^1} C(t)}{n\gamma U}. \tag{6}$$

Let $X_t := C(t)/U$. Since $\ell \le 2c_2 n\gamma\,(U/\sigma + \log(D))$ and $\ell \ge \left|T^1\right|$, we obtain two different bounds from (5) and (6), namely

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\le\; 16b\,n\left(\frac{\sum_{t\in T^1} X_t}{\sum_{t\in T^1} X_t^2}\right) \quad \text{and} \quad \frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\le\; 32c_2\left(\frac{U}{\sigma} + \log(D)\right)\left(\frac{\sum_{t\in T^1} X_t}{|T^1|}\right).$$

Since $\left|T^1\right| \ge 1$, by applying Lemma 9 on these two bounds, we obtain

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \;\le\; \sqrt{512c_2 b\,n\,(U/\sigma + \log(D))} = O\big(\sqrt{n\,(U/\sigma + \log(D))}\big).$$

We conclude that

$$\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \mathcal{E}\right] = O\big(\sqrt{n\,(U/\sigma \;+\; \log(D))}\big).$$

□

Proof (Lemma 2). Let ℓ be the length of the sequence $\mathcal{K}$ and let $s_0, s_1, \ldots, s_\ell$ denote the sequence of nodes visited by WFA starting at s_0. We have $w_\ell(s_\ell) - w_0(s_0) \ge -nU$, since $w_0(s_0) \le w_\ell(s_0)$ and due to Fact 3. Therefore,

$$(w_\ell(s_\ell) \;-\; w_{\ell-1}(s_{\ell-1})) \;+\; (w_{\ell-1}(s_{\ell-1}) \;-\; w_{\ell-2}(s_{\ell-2})) \;+\; \cdots \;+\; (w_1(s_1) \;-\; w_0(s_0)) \;\ge\; -nU.$$

Let R^- denote the set of rounds t, where $w_t(s_t) - w_{t-1}(s_{t-1}) < 0$, and let R^+ denote the rounds not in R^-. Then, the last inequality can be rewritten as

$$\sum_{t\,\in\, R^-} (w_{t-1}(s_{t-1}) - w_t(s_t)) \le nU + \sum_{t\,\in\, R^+} (w_t(s_t) \;-\; w_{t-1}(s_{t-1})).$$

Since $\bar{T} \;\subseteq R^-$, we have

$$\sum_{t\,\in\, \bar{T}} (w_{t-1}(s_{t-1}) \;-\; w_t(s_t)) \le nU \;+\; \sum_{t\,\in\, R^+} (w_t(s_t) \;-\; w_{t-1}(s_{t-1})). \tag{7}$$

For any $t \;\in\; \bar{T}$, we have $C(t) \;\le\; 3\,(w_{t-1}(s_{t-1}) - w_t(s_t))$, which can be seen as follows. Let $h_t := w_{t-1}(s_{t-1}) - w_t(s_t)$. We have $w_{t-1}(s_t) \ge w_{t-1}(s_{t-1}) - \delta(s_{t-1}, s_t)$ (by Fact 3) and $r_t(s_t) = w_t(s_t) - w_{t-1}(s_t)$ (by (2)). Therefore, $r_t(s_t) \le \delta(s_{t-1}, s_t) - h_t$. Since $t \in \bar{T}$ and by the definition (3) of T, for $t \in \bar{T},\ \delta(s_{t-1}, s_t) < 2h_t$. Thus, we have $C(t) = r_t(s_t) + \delta(s_{t-1}, s_t) \le 3h_t$. Hence,

$$\frac{1}{3}\sum_{t\in\bar{T}} C(t) \le \sum_{t\,\in\,\bar{T}} (w_{t-1}(s_{t-1}) \;-\; w_t(s_t)). \tag{8}$$

Moreover, for any $t \;\in\; R^+$, we have $w_t(s_t) \;-\; w_{t-1}(s_{t-1}) \;\le\; C(t)$. This follows from the facts that $w_t(s_t) \;-\; w_{t-1}(s_{t-1}) = r_t(s_t) + w_{t-1}(s_t) - w_{t-1}(s_{t-1})$ (by (2)) and $w_{t-1}(s_t) - w_{t-1}(s_{t-1}) \le \delta(s_{t-1}, s_t)$. Since $R^+ \subseteq T$, we can write $\sum_{t\,\in\, R^+} w_t(s_t) - w_{t-1}(s_{t-1}) \;\le\; \sum_{t\,\in\, T} C(t)$. This together with (7) and (8) implies that

$$\sum_{t\in\bar{T}} C(t) \;\le\; 3\sum_{t\in T} C(t) \;+\; 3nU.$$

Since $\text{WFA}[\mathcal{K}]$ is at least $6nU$, we have that $\sum_{t\in T} C(t) \;\ge\; \text{WFA}[\mathcal{S}]/8$. □

Proof (Lemma 3). Let $s_0, s_1, \ldots, s_\ell$ denote the sequence of nodes visited by WFA for the input sequence $\mathcal{K}$ of length ℓ. We recall the definition of T, T^1 and T^2 for the sequence $\mathcal{K}$. For any round t, let $W_t := \sum_{i=1}^{n} w_t(v_i)$ and $\Delta W_t := W_t - W_{t-1}$.

Claim 1. *Fix a round t and consider any node u. Let $w_t(u) - w_{t-1}(u) \geq H$. If $H \leq 4U\,diam$, then $\Delta W_t \geq H^2/(10U)$; otherwise, $\Delta W_t \geq nH/2$.*

Proof. For the case $H \leq 4U\,diam$, let $d := \lfloor H/(8U) \rfloor$. For $d = 0$, we can already verify the claim. Let the node u be denoted as u_0. Consider a shortest path $P = (u_0, u_1, \ldots, u_d)$. Since $d \leq \lfloor diam/2 \rfloor$, there always exists a shortest path of length d. We have that for all $i \in \{0, \ldots, d\}$, $w_{t-1}(u_i) \leq w_{t-1}(u_0) + iU$ and $w_t(u_i) \geq w_t(u_0) - iU$. Since $d \leq H/(8U) \leq d+1$,

$$\sum_{v \in P} (w_t(v) - w_{t-1}(v)) \;\geq\; (d+1)H - 2U \sum_{i=1}^{d} i \;\geq\; (d+1)(H - dU) \;\geq\; \frac{H^2}{10U}.$$

If $H > 4U\,diam$, then, since for any node v of G, $w_{t-1}(v) \leq w_{t-1}(u) + U\,diam$ and $w_t(v) \geq w_t(u) - U\,diam$, we have that $\sum_{v \in G} (w_t(u) - w_{t-1}(u)) \geq nH - 2nU\,diam \geq nH/2$. □

Claim 2. *For any $t \in T^1$, $\Delta W_t \geq C(t)^2/(160U)$.*

Proof. Consider any round $t \in T^1$. Let $C(t) = dU + r$, where dU is the travel cost and r is the service cost. We already have that $r = w_t(s_t) - w_{t-1}(s_t)$ by (2). We will show that

$$\Delta W_t \;\;\geq\;\; \left((dU)^2 + r^2\right)/(80U). \tag{9}$$

Since we know that $C(t)^2 = (dU+r)^2 \leq 2\left((dU)^2 + r^2\right)$, we conclude that $\Delta W_t \geq C(t)^2/(160U)$. Now, all that remains to be shown is (9). We distinguish two cases:

Case $dU \geq r$: We have $w_t(s_{t-1}) - w_{t-1}(s_{t-1}) \geq dU/2$. This is because of the following two facts, namely $w_t(s_{t-1}) = w_t(s_t) + dU$ (Fact 4) and $w_{t-1}(s_{t-1}) \leq w_t(s_t) + dU/2$ (definition of T). Hence, by applying Claim 1 to node s_{t-1}, $\Delta W_t \geq (dU)^2/(40U) \geq \left((dU)^2 + r^2\right)/(80U)$.

Case $r > dU$: Since by (2), $w_t(s_t) - w_{t-1}(s_t) = r$, and since $r \leq 4\,U\,diam$ by the definition of t, applying Claim 1 to s_t, we obtain $\Delta W_t \geq r^2/10U \geq \left((dU)^2 + r^2\right)/(20U)$. □

Claim 3. *For any $t \in T^2$, $\Delta W_t \geq nC(t)/4$.*

Proof. Consider any round $t \in T^2$. Let $C(t) = dU + r$, where dU is the travel cost and r is the service cost. We recall that $r = w_t(s_t) - w_{t-1}(s_t)$ by (2). We have that $w_t(s_t) - w_{t-1}(s_t) \geq C(t)/2$. This is because, $w_t(s_t) - w_{t-1}(s_t) = r = C(t) - dU$, and $C(t) - dU \geq C(t)/2$, since $C(t) > 4diamU$ (definition of T^2) and $d \leq diam$. Thus, we have $w_t(s_t) - w_{t-1}(s_t) = r \geq 4diamU$ and $r \geq C(t)/2$. By applying Claim 1 to node s_t, the proof follows. □

Claims 2 and 3 together imply that

$$\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i) = \sum_{t=1}^{\ell} \Delta W_t \geq \sum_{t \in T} \Delta W_t \geq \frac{1}{160U} \sum_{t \in T^1} C(t)^2 + \frac{n}{4} \sum_{t \in T^2} C(t). \tag{10}$$

Below, we will show that

$$\textsc{opt}[\mathcal{K}] \geq \left(\frac{1}{2n}\right) \left(\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i)\right). \tag{11}$$

By combining (10) and (11), we obtain the lemma for an appropriate choice of b.

Now, only (11) remains to be shown. Since for any v_i, $w_\ell(v_i) \leq \min\{w_\ell(\cdot)\} + nU$ (Fact 3) and since $\text{OPT}[\mathcal{K}] \geq \min\{w_\ell(\cdot)\}$, we obtain

$$\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i) \;\leq\; n \cdot \text{OPT}[\mathcal{K}] + n^2U - \sum_{i=1}^{n} w_0(v_i) \;\leq\; n \cdot \text{OPT}[\mathcal{K}] + n^2U.$$

Thus,

$$\text{OPT}[\mathcal{K}] \;\geq\; (1/n)\left(\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i) \;-\; n^2U\right). \tag{12}$$

Moreover, for those rounds in T^1, where $C(t) \;<\; U$, we have $\Delta W_t \geq C(t)$. This follows from the fact that if $C(t) < U$, then there is no travel cost and $C(t) = w_t(s_t) - w_{t-1}(s_t) \;\leq\; \Delta W_t$. Otherwise, for those rounds in T^1, where $C(t) \geq U$, we have due to Claim 2 that $\Delta W_t \;\geq\; C^2(t)/(160U) \;\geq\; C(t)/160$. The above observations can be summarized as

$$\sum_{t \in T^1,\; C(t) < U} \Delta W_t \geq C(t) \quad \text{and} \quad \sum_{t \in T^1,\; C(t) \geq U} \Delta W_t \geq \frac{C(t)}{160}.$$

Therefore,

$$\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i) \geq \sum_{t \in T^1} \Delta W_t + \sum_{t \in T^2} \Delta W_t \geq \frac{1}{160} \sum_{t \in T^1} C(t) + \frac{n}{4} \sum_{t \in T^2} C(t) \geq \frac{1}{160} \sum_{t \in T} C(t).$$

We fix the constant c_w appropriately such that the fact that $\text{WFA}[\mathcal{K}] \;\geq\; c_w n^2 U$ and the fact that $\sum_{t \in T} C(t) \geq \text{WFA}[\mathcal{K}]/8$ (by Lemma 2) together with the above inequality imply that $\sum_{i=1}^{n} w_\ell(v_i) - \sum_{i=1}^{n} w_0(v_i) \geq 2n^2U$. This together with (12) implies (11), which concludes the proof. □

E Proof of Theorem 5

Proof (Theorem 5). Fix any adversarial sequence $\bar{\mathcal{S}}$ of length ℓ, where ℓ is as defined in Lemma 1 for some $\gamma \;\geq\; 1$. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote a smoothed task sequence obtained from $\bar{\mathcal{S}}$. Let $\mathcal{E}$ be the event which denotes that the cost incurred by OPT on $\mathcal{S}$ is at least $n\gamma U$. From Lemma 1, we have that $\mathbf{P}[\mathcal{E}] \;\geq\; 1 - 1/n^3$. Using Proposition 2 (Appendix M), we have that for any task sequence $\mathcal{S}$ of length ℓ,

$$\text{WFA}[\mathcal{K}] \;\leq\; \text{OPT}[\mathcal{K}] + \mathit{diam} U \ell.$$

Let $\mathcal{S}_\mathcal{E}$ be a fixed outcome of smoothing $\bar{\mathcal{S}}$ conditioned on the event $\mathcal{E}$. From the definition of $\mathcal{E}$, we have that $\text{OPT}[\mathcal{S}_\mathcal{E}] \geq n\gamma U$ and $\text{WFA}[\mathcal{S}_\mathcal{E}] \leq \text{OPT}[\mathcal{S}_\mathcal{E}] + \mathit{diam} U \ell$. Using the relation between γ and ℓ as stated in Lemma 1, we obtain $\text{WFA}[\mathcal{S}_\mathcal{E}]/\text{OPT}[\mathcal{S}_\mathcal{E}] \leq b \cdot \mathit{diam} \cdot (U/\sigma + \log(D))$ for an appropriate constant b. Hence,

$$\begin{aligned}
\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \;&=\; \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \mathcal{E}\right] \mathbf{P}[\mathcal{E}] \;+\; \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \;\middle|\; \neg\mathcal{E}\right] \mathbf{P}[\neg\mathcal{E}] \\
&\leq\; b \cdot \mathit{diam} \cdot \left(\frac{U}{\sigma} + \log(D)\right) \;+\; \frac{2n-1}{n^3} \;=\; O\left(\mathit{diam} \cdot \left(\frac{U}{\sigma} \;+\; \log(D)\right)\right).
\end{aligned}$$

□

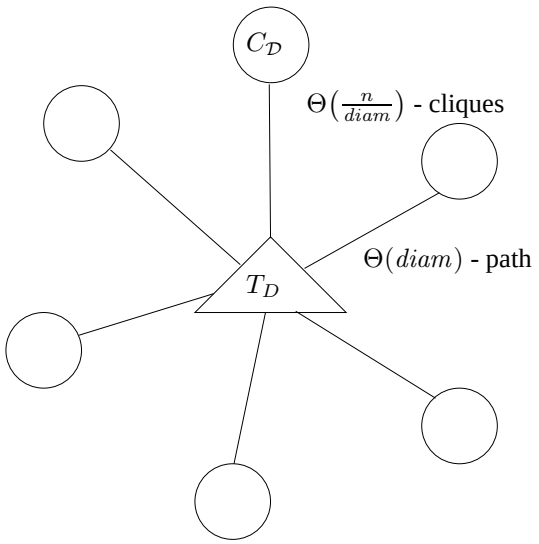

Figure 1:

F Proof of Theorem 6 and Theorem 7

We use the following lemma to prove the two lower bounds. For all given values of n, $diam$ and D that satisfy the conditions in the proof of Lemma 6, we can construct graphs which give the desired lower bounds as stated in either Theorem 6 or Theorem 7, depending on whichever is tighter. But we note that these conditions only exclude marginal cases and are not restrictive. Moreover, for the cases $D > n/4$ and $diam > n/4$, desired bounds are already obtained by the universal lower bounds of Theorem 8 and Theorem 9, respectively.

Lemma 6. *Given diameter $diam$ and maximum degree D, such that $b < diam \leq n/4$ and $b < D \leq n/4$ for some constant b, there exists a graph with n nodes and uniform edge lengths U such that the smoothed competitive ratio of any deterministic algorithm* ALG *is*

$$\Omega(\min\{n/diam, diam \cdot (U/\sigma + \log(\mathcal{D}))\},$$

where $\mathcal{D} = \min\{diam, D\} - 1$

Proof. The graph looks like the one depicted in Figure 1. The graph consists of $m := \Theta(\frac{n}{diam})$ lollipops, i.e, each one consists of a line having $r := \Theta(diam)$ nodes which has a clique $C_{\mathcal{D}}$ of size $\mathcal{D}$ attached to one end of it. The free ends of all these subgraphs are connected to each other by a $(D-1)$-ary tree T_D. This is to maintain the maximum degree constraint. The subgraphs are joined to the leaves of this tree T_D. For simplicity, assume that the distance between the corresponding nodes in any pair of subgraphs (lollipops) is exactly the same. For any appropriately fixed constants m and r, the asymptotic bounds remain the same. If, after creating and connecting these subgraphs, there are some nodes remaining, we let them become part of the connecting tree T_D and always impose ∞ costs on them. Note that the height of the connecting tree T_D is $O(diam)$, since $diam \geq \log_D(n)$. The constant b is fixed as specified in Lemma 7 and we make sure that $b \geq 3$. Moreover m and r are chosen appropriately to maintain the diameter constraint. The constraints on D and $diam$ in the statement of the lemma are imposed to avoid the extreme cases of having only one clique, etc; they do not have any other impact.

First, we consider the case $\frac{U}{\sigma} \geq \log(\mathcal{D})$. In this case it is enough to show a lower bound of $O(\min\{n/diam,\ diam \cdot U/\sigma\})$. We can achieve this, by imposing in each round an ∞ cost on all nodes in the graph except on those, which join a line segment with its clique. A lower bound on this 'virtual' clique of size m and unifom edge length $\Theta(diam \cdot U)$ is our desired lower bound. For that we apply Theorem 8 on this clique.

For the case $\log(\mathcal{D}) > \frac{U}{\sigma}$, we use the averaging technique. We will compare the cost of ALG with a collection $\mathbf{B}$ of $m-1$ offline algorithms. Let us view the smoothed request sequence $\mathcal{S}$ of length ℓ being generated in two stages. We note that the two stage view of request generation does not make any difference to ALG as well as any offline algorithm.

Stage 1: In all the ℓ request vectors, all the non-clique nodes (nodes which do not belong to any $C_{\mathcal{D}}$) are given a cost of ∞ and clique nodes have cost zero. We smoothen this sequence and obtain a sequence $\mathcal{S}'$.

Stage 2: From $\mathcal{S}'$, we get the final smoothed sequence $\mathcal{S}$ as follows. In round t, all nodes in the clique where the ALG ends up after round $t-1$ gets ∞ cost. This forces ALG to move out of the current clique to another one in that round.

Each algorithm $\mathbf{B}_i$ in $\mathbf{B}$ simulates the behavior of the offline algorithm as mentioned in Lemma 7 with the following modifications. Each clique is occupied by atmost one $\mathbf{B}_i$. If an ∞ cost is imposed on a clique, the offline algoritm which is present there moves to the free clique. For a given sequence $\mathcal{S}$, there is a fixed offline strategy for each clique as described by Lemma 7. If say B_i travels from clique C_j to C_k because of ∞ costs, in C_k, B_i proceed with the offline strategy for C_k as long as it is in C_k. Without loss of generality, assume that each $\mathbf{B}_i$ starts in a different clique (see Appendix N). Let $\mathbf{B}_i[\mathcal{S}]$ denote the total cost for $\mathbf{B}_i$ on $\mathcal{S}$ and let $\mathbf{B}_{av} := 1/(m-1)\sum_{i=1}^{m-1}\mathbf{B}_i[\mathcal{S}]$. It is easy to see that

$$\mathbf{B}[\mathcal{S}] \;=\; \sum_{i=1}^{m-1}\mathbf{B}_i[\mathcal{S}] \;\leq\; \ell U\, diam \;+\; \sum_{i=1}^{m-1}\mathbf{B}_i[\mathcal{S}'],$$

where the term $\ell \cdot U \cdot diam$ accounts for the additional travelling cost incurred for travelling away from the clique with all ∞ costs.
Since $\mathcal{D} \geq a$, where a is a constant as required in Lemma 7, by applying Lemma 7, we get

$$\mathbf{E}\left[\sum_{i=1}^{m-1}\mathbf{B}_i[\mathcal{S}']\right] \;\leq\; \frac{6m\ell U}{\log(\mathcal{D})}.$$

Let $\mathcal{E}$ denote the event that $\sum_{i=1}^{m-1}\mathbf{B}_i[\mathcal{S}'] \leq 12m\ell U/\log(\mathcal{D})$. By Markov's inequality, $\mathbf{P}[\mathcal{E}] \geq 1/2$. Hence under the event $\mathcal{E}$, $\mathbf{B}[\mathcal{S}] \leq \ell U\, diam + 12m\ell U/\log(\mathcal{D})$ and $\text{ALG}[\mathcal{S}] \geq \ell U\, diam$. Hence

$$\mathbf{B}_{av} = \mathbf{B}[\mathcal{S}]/(m-1) \leq \text{ALG}[\mathcal{S}]/(m-1) + 24\ell U/\log(\mathcal{D})$$

Putting everything together, we obtain

$$\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \;\geq\; \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}}\;\middle|\;\mathcal{E}\right]\mathbf{P}[\mathcal{E}] \geq \frac{1}{2}\left(\frac{\text{ALG}[\mathcal{S}]}{\frac{\text{ALG}[\mathcal{S}]}{m-1} + \frac{24\ell U}{\log(\mathcal{D})}}\right)$$

Depending on the dominating term in the denominator sum, the above expression gives either an $\Omega(diam \cdot \log(\mathcal{D}))$ or an $\Omega(n/diam)$ lower bound. □

Proof (Theorem 6). Depending on whether $U/\sigma \geq \log(D)$ or not, the graph we construct will be different. First consider the case $U/\sigma \geq \log(D)$. We will fix the diameter of the graph such that for

a suitable constant c, $diam \cdot U/\sigma \geq c\sqrt{nU/\sigma}$. This implies that $diam \geq c\sqrt{n\sigma/U}$. The graph we consider in this case will be as given by Lemma 6 with diameter $diam = \left\lceil c\sqrt{\frac{n\sigma}{U}} \right\rceil + b$ and n nodes and maximum degree D. The constant b is as required by Lemma 6. Under this setting, we have that $n/diam = \frac{1}{c}\sqrt{nU/\sigma} = \frac{1}{c^2} diam \cdot (U/\sigma)$. Hence by Lemma 6 there exists a graph where $\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] = \Omega\left(\sqrt{nU/\sigma}\right)$. For the case $U/\sigma < \log(D)$, by similar arguments we will fix the diameter to be $c\sqrt{n/\log(D)}$. Under this setting we also have that $n/diam = \frac{1}{c}\sqrt{n\log(D)} = \frac{1}{c^2} diam \cdot \log(D)$. Hence by Lemma 6, there exists a graph where $\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] = \Omega\left(\sqrt{n\log(D)}\right)$. □

Proof (Theorem 7). The class of graphs which we consider are such that, $diam\left(\frac{U}{\sigma} + \log(D)\right) \leq c\sqrt{n\left(\frac{U}{\sigma} + \log(D)\right)}$, for a suitable constant c. In fact this condition is necessary because of the upper bounds given by Theorem 5 and Theorem 4. In other words we have, $\frac{n}{diam} \geq \frac{1}{c^2} \cdot diam \cdot \left(\frac{U}{\sigma} + \log(D)\right)$. Under this condition, by Lemma 6 we have that for $diam \geq b$, where b is some constant, there are graphs with $\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] = \Omega\big(diam\left(\frac{U}{\sigma} + \log(\mathcal{D})\right)\big)$ where $\mathcal{D} = \min\{diam, D\} - 1$. For constant values of $diam$, Theorem 8 gives the desired lower bound. Since $diam \leq \sqrt{n}$, the requirement that $diam \leq n/4$ by Lemma 6 is anyway satisfied here for large n. □

G Proof of Theorem 8

Proof (Theorem 8). Consider any deterministic algorithm ALG. We distinguish two cases.

Case 1: $U/\sigma \geq \log(D)$. For a sufficiently large ℓ, which we will fix later, the adversary constructs a request sequence $\bar{\mathcal{S}}$ of length ℓ consisting only of elementary tasks as follows. Let v denote the node in which ALG resides after round $i-1$. Then, in round i, the adversary puts a request cost of U on v. (Recall that the adversary is adaptive and thus knows where ALG resides after round $i-1$.) Let $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the random variable of the ℓ tasks after smoothing.

We use an averaging technique similar to the one presented in [6]. Let $\mathbf{B} = \{\mathbf{B}_1, \ldots, \mathbf{B}_{2n-1}\}$ be a collection of $2n-1$ (not necessarily optimal) offline algorithms. We compare the average cost of the algorithms in $\mathbf{B}$ with the cost of ALG. The following invariant is maintained true by the algorithms. In round i, let v be the node where ALG resides. Then all nodes of G other than v are occupied by exactly two algorithms and v itself is occupied by one algorithm. The invariant is easily maintained when ALG changes its state from v to, say w: one algorithm from w travels to v. Without loss of generality, we ignore the additional costs incurred by the algorithms in $\mathbf{B}$ to reach their initial position (see Appendix N). Let $\mathbf{B}[\mathcal{S}]$ denote the sum of the costs incurred by all the algorithms in $\mathbf{B}$ for processing $\mathcal{S}$. Clearly, $\mathbf{B}[\mathcal{S}]/(2n-1)$ is an upper bound on $\text{OPT}[\mathcal{S}]$. Let $\mathbf{B}_{av}$ denote $\mathbf{B}[\mathcal{S}]/(2n-1)$. It suffices to show that $\text{ALG}[\mathcal{S}]/\mathbf{B}_{av} = \Omega(\min\{n, U/\sigma\})$ with constant probability.

For the analysis, we view the smoothing as being done in two stages. *Stage 1:* Initially there are ℓ zero vectors, i.e., all cost components are zero. We smoothen these vectors according to $\mathcal{P}$ and obtain a new sequence $\mathcal{S}' = \langle \tau'_1, \ldots, \tau'_\ell \rangle$ of smoothed cost vectors. *Stage 2:* For each i, the request vector τ_i presented to ALG is generated from τ'_i as follows. If ALG resides in node v at the end of round $i-1$, then the cost component for v is replaced by a value obtained by smoothing an initial value of U. From the perspective of ALG as well as any algorithm in $\mathbf{B}$, this view does not make any difference. Let $\mathbf{C}'(j) = \sum_{t=1}^{\ell} r'_t(j)$ denote the total request cost accumulated in j due to the sequence $\mathcal{S}'$. From Fact 5, we have that for any j, $\mathbf{E}[\mathbf{C}'(j)] \leq \ell\sigma$.

Let $\mathcal{E}$ denote the event that $\sum_{j=1}^{n} \mathbf{C}'(j) \leq 2n\ell\sigma$. By Markov's inequality, we have $\mathbf{P}[\mathcal{E}] \geq 1/2$. Let the random variables $U_1, \ldots, U_\ell$ denote the ℓ cost components, one for each round, which were introduced during step 2 to obtain the final sequence $\mathcal{S}$. Using Fact 5 and the fact that $\sigma \leq U$,

since $U/\sigma \geq \log(D)$, we have $\mathbf{E}[U_i] \leq 2U$, for any i. Hence by Markov inequality, we have that, $\mathbf{P}\left[\sum_{i=1}^{\ell} U_i \geq 4\ell U\right] \leq 1/2$. Note that the U_i's are independent. For each U_i, let z_i denote a 0/1 random variable which takes value 1 iff $U_i \geq U$. Let $Z := z_1 + \cdots + z_\ell$. Since the smoothing distribution $\mathcal{P}$ is a symmetric, for any i: $\mathbf{P}[U_i \geq U] \geq 1/2$, and therefore $\mathbf{E}[z_i] \geq 1/2$. Applying a Chernoff bound, see, e.g., [9], we obtain

$$\mathbf{P}[Z \leq \ell/4] \leq e^{-\ell/16}. \tag{13}$$

Let $\mathcal{A}$ denote the event $(Z \geq \ell/4) \wedge (\sum_{i=1}^{r} U_i \leq 4\ell U)$. The event $(Z \geq \ell/4)$ means that at least in $\ell/4$ of the smoothed requests, ALG incurs a cost of at least U if it does not change its position in these rounds. The fact that $\mathbf{P}\left[\sum_{i=1}^{\ell} U_i \geq 4\ell U\right] \leq 1/2$ and (13) together imply that $\mathbf{P}[\mathcal{A}] = 1 - \mathbf{P}[\neg\mathcal{A}] \geq 1 - \left(1/2 + e^{-\ell/16}\right)$, which for sufficiently large ℓ ($\ell \geq 32$ would suffice) is at least $1/4$.

We define a new event $(\mathcal{E} \wedge \mathcal{A})$, which can be interpreted as that after stage 1 of the smoothing event $\mathcal{E}$ holds and that after stage 2 the event $\mathcal{A}$ holds. From the description of stage 1 and stage 2, we see that the events $\mathcal{E}$ and $\mathcal{A}$ are independent. Hence,

$$\mathbf{P}[\mathcal{E} \wedge \mathcal{A}] \geq 1/8 \tag{14}$$

Also note that the occurrence of the events $\mathcal{E}$ and $\mathcal{A}$ are independent of the moves made by ALG. We analyze the ratio of $\text{ALG}[\mathcal{S}]$ and $\mathbf{B}_{av}$ conditioned on the event $(\mathcal{E} \wedge \mathcal{A})$. Let $\mathcal{S}_g$ be any fixed sequence which is the outcome of smoothing $\bar{\mathcal{S}}$ conditioned on $(\mathcal{E} \wedge \mathcal{A})$. For $\mathcal{S}_g$, let ALG change its position in k of the ℓ rounds. The set of ℓ rounds can be partitioned into T_k and $T_{\ell-k}$, where T_k are those rounds where ALG changes its position. From the description of $\mathbf{B}$ and the event $\mathcal{E} \wedge \mathcal{A}$, we have that for $\mathcal{S}_g$

$$\mathbf{B}[\mathcal{S}_g] \leq \text{ALG}[\mathcal{S}_g] + 2\sum_{t \in T_k} U_t^g + 4n\ell\sigma,$$

where for $t \in T_k$ U_t^g denotes the request cost imposed on the node where ALG was at the beginning of round t. Since in this round ALG moved to another node, two offline algorithms incur this cost. Because of the event $(\mathcal{E} \wedge \mathcal{A})$, since in at least $\ell/4$ of the rounds, ALG has a cost of at least U, we also have $\text{ALG}[\mathcal{S}_g] \geq \ell U/4$. This together with the definition of the event $\mathcal{A}$ implies that

$$\text{ALG}[\mathcal{S}_g] + 2\sum_{t \in T_k} U_t^g \leq \text{ALG}[\mathcal{S}_g] + 8\ell U \leq 33\,\text{ALG}[\mathcal{S}_g].$$

That is, for $\mathcal{S}_g$: $\mathbf{B}_{av} = \mathbf{B}[\mathcal{S}_g]/n \leq 33\text{ALG}[\mathcal{S}_g]/n + 4\ell\sigma$. Moreover, since $\text{ALG}[\mathcal{S}_g] \geq \ell U/4$ we conclude that

$$\frac{\text{ALG}[\mathcal{S}_g]}{\mathbf{B}_{av}} \geq \frac{\text{ALG}[\mathcal{S}_g]}{33\text{ALG}[\mathcal{S}_g]/n + 4\ell\sigma} \geq b \cdot \min\{n, U/\sigma\},$$

where b is a fixed constant chosen appropriately. Finally, from (14) we obtain

$$\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}} \,\middle|\, \mathcal{E} \wedge \mathcal{A}\right] \mathbf{P}[\mathcal{E} \wedge \mathcal{A}] \geq \frac{b}{8} \cdot \min\left\{n, \frac{U}{\sigma}\right\} = \Omega\left(\min\left\{n, \frac{U}{\sigma}\right\}\right).$$

Case 2: $U/\sigma < \log(D)$. We will prove this case for $D = \omega(1)$; for $D = O(1)$, the term $\log(D)$ is irrelevant in the lower bound expression. Let s be a node of G with degree D. Let G_s be the induced subgraph of G with vertex set V_s containing only the vertex s itself and its neighbors in G; the edge set of G_s contains all edges of G incident between vertices of V_s. For a sufficiently large ℓ, which will be fixed later, the adversarial request sequence $\bar{\mathcal{S}}$ consists of ℓ tasks as follows. In any round i, the

cost component of each vertex not in G_s is ∞. In addition, there is an ∞ cost component on the node in which ALG resides after round $i-1$. (Again, recall that the adversary is adaptive.) All remaining cost components are zero. We assume that ALG always stays within G_s; any reasonable algorithm avoids ∞ costs. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the ℓ tasks after smoothing.

We compare the cost of ALG with the cost of an offline algorithm $\mathbf{B}$ as specified in Lemma 7. Since G_s is a star with $D+1$ nodes and the cost between any two nodes in G_s is at most $2U$, by Lemma 7 we have $\mathbf{E}[\mathbf{B}[\mathcal{S}]] \leq \frac{12\ell U}{\log(D)}$. We can apply Lemma 7 here since, with respect to G_s, the request sequence is elementary. Now by Markov's inequality, we obtain $\mathbf{P}[\mathbf{B}[\mathcal{S}] \geq 24\ell U/\log(D)] \leq 1/2$. Since ALG has to move in each round to avoid ∞ cost, the cost of ALG for any smoothed sequence is at least ℓU. Putting all these things together, we obtain

$$\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}[\mathcal{S}]}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}[\mathcal{S}]}\middle| \mathbf{B}[\mathcal{S}] \leq \frac{24\ell U}{\log(D)}\right]\left(\frac{1}{2}\right) = \Omega(\log(D))\,.$$

□

Lemma 7. *Let G_c be a clique of $m+1$ nodes such that all edge lengths are at most U. Consider an adversarial sequence $\bar{\mathcal{S}}$ of ℓ elementary tasks, for some sufficiently large ℓ. Let the random variable $\mathcal{S}$ denote the smoothed sequence. Then, there exists an offline algorithm $\mathbf{B}$ such that for $m \geq \alpha$ and α a fixed constant, $\mathbf{E}[\mathbf{B}[\mathcal{S}]] \leq \frac{6\ell U}{\log(m)}$*

Proof. For convenience assume that $\ell = \gamma k$, where $k := \lceil \log(m)/3 \rceil$ and $\gamma \geq 1$. We split the sequence $\mathcal{S}$ into γ phases $\mathcal{S}_1, \ldots, \mathcal{S}_\gamma$, each consisting of k contiguous tasks. We will specify the strategy of $\mathbf{B}$ shortly. Let us consider any phase $\mathcal{S}_j$ of k tasks. For the analysis, we view the smoothing as being done in two stages. We describe it here for a phase, say j: *Stage 1:* Initially there are k zero vectors (all components are zero). Smoothen all these vectors with the given distribution to get a new sequence which we denote by $\mathcal{S}'_j = \langle \tau'_1, \ldots, \tau'_k \rangle$. *Stage 2:* For each round i, the request vector τ_i that is presented to $\mathbf{B}$ is generated from τ'_i as follows. The component of τ'_i which corresponds to the non-zero component in the original task vector is replaced by the smoothed outcome of the corresponding non-zero cost. The original task is elementary, hence this is done on at most one component. For the offline algorithm this view does not make any difference. For any node v_i of G_c, let $\mathbf{C}'(i)$ denote the total cost accumulated in v_i because of the sequence $\mathcal{S}'_j$. For any node v_i of G_c, let X_i denote a $0/1$ random variable which takes value 1 iff $\mathbf{C}'(i) = 0$. Since the smoothing distribution is symmetric, we have $\mathbf{P}[X_i = 1] \geq (1/2)^k \geq 1/\sqrt{m}$ (where the latter inequality holds for $m \geq 64$). Let $\mathbf{X} := X_1 + \cdots + X_{m+1}$. Note that the X_i's are independent. We have that $\mathbf{E}[\mathbf{X}] \geq \sqrt{m}$. Let $\mathcal{E}_j$ denote the event that for a sequence $\mathcal{S}'_j$, $\mathbf{X} \geq \sqrt{m}/2$. By Chernoff bound, see, e.g., [9], we have

$$\mathbf{P}[\neg\mathcal{E}_j] \leq \mathbf{P}\big[\mathbf{X} \leq \sqrt{m}/2\big] \leq e^{-\sqrt{m}/16}. \tag{15}$$

The offline algorithm $\mathbf{B}$ has two different strategies based on the event $\mathcal{E}_j$. If the event $\mathcal{E}_j$ is true for phase j, then $\mathbf{B}$ follows *Strategy 1:* At the beginning of phase j, $\mathbf{B}$ travels to a node whose total accumulated cost is zero at the end of phase j ($\mathbf{B}$ is offline), and stays in that node for the rest of the phase. This is possible since under the event $\mathcal{E}_j$ there are at least $\sqrt{m}/2 - k$ such nodes. This is immediately seen from the description of stages 1 and 2 and the event $\mathcal{E}_j$. For $m \geq 128$, $\sqrt{m}/2 - k$ is at least one. If the event $\mathcal{E}_j$ is not true for phase j, then $\mathbf{B}$ follows *Strategy 2:* In every round t of phase j, move to the node with the lowest (smoothed) request cost. Let $\mathbf{B}^1(j)$ denote the cost spent by $\mathbf{B}$ in phase j when following strategy 1; analogously, let $\mathbf{B}^2(j)$ denote the cost spent by $\mathbf{B}$ in phase j following strategy 2. Moreover, let $\mathbf{B}(j)$ denote the actual cost spent by $\mathbf{B}$ in phase j. We have

$$\mathbf{E}[\mathbf{B}(j)] = \mathbf{E}\big[\mathbf{B}^1(j) \mid \mathcal{E}_j\big] \cdot \mathbf{P}[\mathcal{E}_j] + \mathbf{E}\big[\mathbf{B}^2(j) \mid \neg\mathcal{E}_j\big] \cdot \mathbf{P}[\neg\mathcal{E}_j] \tag{16}$$

We have that $\mathbf{E}\big[\mathbf{B}^1(j) \mid \mathcal{E}_j\big] \leq U$. This is because under the condition $\mathcal{E}_j$, $\mathbf{B}$ incurs only the initial travel cost to the good node, which is at most U. Now we will show that $\mathbf{E}\big[\mathbf{B}^2(j) \mid \neg\mathcal{E}_j\big] \leq 3kU$. To show this, for round t of phase j, let $\min_2(t)$ denote the second smallest request cost of all nodes in G_c. We concentrate on the second smallest request cost, because the cost incurred by B^2 during round t is at most $\min_2(t) + U$. The smallest request cost could possibly get corrupted by the non-zero cost component introduced in stage 2. Note that $\mathbf{E}[\min_2(t)] \leq 2\sigma m/2^m$. This is due to Fact 5 and since zeros are retained with probability at least $1/2$ in our smoothing model. Hence, the probability of having at most one zero request component $m+1$ nodes in G_c is at most $m/2^m$ (for $m \geq 2$). Let $\mathrm{MIN}_2 := \sum_{t=1}^{k} \min_2(t)$. Since $\mathbf{B}^2(j) \leq kU + \mathrm{MIN}_2$, we can write $\mathbf{E}\big[\mathbf{B}^2(j)\big] \leq kU + \mathbf{E}[\mathrm{MIN}_2]$, and

$$\mathbf{E}\big[\mathbf{B}^2(j) \mid \neg\mathcal{E}_j\big] \leq kU + \mathbf{E}[\mathrm{MIN}_2 \mid \neg\mathcal{E}_j]\,. \tag{17}$$

Now we will show that $\mathbf{E}\big[\mathbf{B}^2(j) \mid \neg\mathcal{E}_j\big] \leq 4\sigma + kU + \mathbf{E}[\mathrm{MIN}_2]$. To show this, we apply the following trick. We will look at a modified event $\neg\mathcal{E}_j^t$ for every round t. For $t = k$, the event $\neg\mathcal{E}_j^k$ is same as $\neg\mathcal{E}_j$, which means that there are at most $\sqrt{m}/2$ nodes with a total accumulated cost equal to zero after stage 1. For $t \neq k$, the event $\neg\mathcal{E}_j^t$ denotes that after t rounds there are at most $\sqrt{m}/2$ nodes in G_c with total accumulated cost of zero after stage 1. The modified event only means that the bad event ($\neg\mathcal{E}_j^t$) occurs somewhat earlier in the sequence. First we observe that $\mathbf{E}\Big[\mathrm{MIN}_2 \mid \neg\mathcal{E}_j^1\Big] \leq 4\sigma + \mathbf{E}[\mathrm{MIN}_2]$. This is because $\mathbf{E}\Big[\min_2(1) \mid \neg\mathcal{E}_j^1\Big] \leq 4\sigma$, which is a weak but sufficient bound for our purpose. (This is obtained by the fact that the expected second smallest cost is at most the expected maximum cost of two arbitrarily fixed nodes). Moreover, we also have $\mathbf{E}\Big[\min_2(t) \mid \neg\mathcal{E}_j^1\Big] \leq \mathbf{E}[\min_2(t) \mid \text{all request costs are non-zero in round } t]$. For the rest of the sequence, since the cost generation is Markovian and does not change because of $\neg\mathcal{E}_j^1$, for all $t > 1$, $\mathbf{E}\Big[\min_2(t) \,|\, \neg\mathcal{E}_j^1\Big] = \mathbf{E}[\min_2(t)]$.

By a simple induction on t, we will show that $\mathbf{E}\Big[\mathrm{MIN}_2 \;\Big|\; \neg\mathcal{E}_j^{(t+1)}\Big] \leq 4\sigma + \mathbf{E}[\mathrm{MIN}_2]$. Assume that the hypothesis is true for t. Base case ($t = 1$) is already verified. For $t > 1$, we can write

$$\mathbf{E}\Big[\mathrm{MIN}_2 \;\Big|\; \neg\mathcal{E}_j^{(t+1)}\Big] \;=\; \mathbf{E}\big[\mathrm{MIN}_2 \;\big|\; \neg\mathcal{E}_j^{t}\big]\,\mathbf{P}\Big[\neg\mathcal{E}_j^t \,\Big|\, \neg\mathcal{E}_j^{(t+1)}\Big] + \mathbf{E}\Big[\mathrm{MIN}_2 \;\Big|\; \mathcal{E}_j^t \wedge \neg\mathcal{E}_j^{(t+1)}\Big]\,\mathbf{P}\Big[\mathcal{E}_j^t \,\Big|\, \neg\mathcal{E}_j^{(t+1)}\Big]\,.$$

Since for any round $\min_2(t)$ is Markovian, we have

$$\begin{aligned}
\mathbf{E}\Big[\mathrm{MIN}_2 \;\Big|\; \mathcal{E}_j^t \wedge \neg\mathcal{E}_j^{(t+1)}\Big] \;&\leq\; \mathbf{E}\left[\sum_{r=1}^{t} \min_2(r)\right] \;+\; \mathbf{E}\Big[\min_2(t+1) \;\Big|\; \neg\mathcal{E}_j^{(t+1)}\Big] \;+\; \mathbf{E}\left[\sum_{r=t+2}^{k} \min_2(r)\right] \\
&\leq\; 4\sigma + \mathbf{E}[\mathrm{MIN}_2]
\end{aligned}$$

Having shown that $\mathbf{E}[\mathrm{MIN}_2 \mid \neg\mathcal{E}_j] \leq 4\sigma + \mathbf{E}[\mathrm{MIN}_2]$, from (17) we obtain that $\mathbf{E}\big[\mathbf{B}^2(j) \mid \neg\mathcal{E}_j\big] \leq kU + 4\sigma + \mathbf{E}[\mathrm{MIN}_2] \leq kU + 4\sigma + 2k\sigma m/2^m \leq 2kU$ for $\sigma \leq 2U$ and for $m \geq \alpha$, α a fixed constant. By (16) and (15) we also conclude that for any phase j, $\mathbf{E}[\mathbf{B}(j)] \leq 2U$ for an appropriate choice of α. Finally, we obtain

$$\mathbf{E}[\mathbf{B}[\mathcal{S}]] = \mathbf{E}\left[\sum_{t=1}^{\gamma} \mathbf{B}_t[\mathcal{S}]\right] \leq 2\gamma U \;\leq\; \frac{6\ell U}{\log(m)}.$$

from the relation between ℓ and γ. □

H Proof of Theorem 9

Corollary 1. *Let G be a line graph of n nodes with uniform edge lengths U. For any deterministic algorithm* ALG *and for general adversarial tasks, the smoothed competitive ratio is* $\Omega(\min\{n, \sqrt{n\,(U/\sigma+1)}\})$

Proof. The bound given in Theorem 11 is maximized for $\beta = \lceil\sqrt{n/(U/\sigma+1)}\rceil$. □

Proof (Theorem 9). We extend Theorem 11 to arbitrary graphs in the obvious way. Consider a shortest path in G of edge diameter $diam$; we can view the shortest path as a line in G. We restrict the metrical task system to this line by enforcing ∞ costs on all nodes that are not part of the line. This defines a new metrical task system on a line with uniform edge lengths U and having $diam+1$ nodes. Applying Corollary 1 to this line, we obtain a lower bound of $\Omega(\min\{diam, \sqrt{diam\,(U/\sigma+1)}\})$ on the smoothed competitive ratio for any deterministic algorithm. □

I Proof of Theorem 10

Proof (Theorem 10). Consider any adversarial input sequence $\bar{\mathcal{S}}$ of length ℓ as defined in Lemma 1 for some appropriate γ. Let the random variable $\mathcal{S} = \langle\tau_1, \ldots, \tau_\ell\rangle$ denote the smoothed sequence obtained from $\bar{\mathcal{S}}$. We define $\mathcal{E}$ as the event that OPT incurs a cost of at least $n\gamma U$ on $\mathcal{S}$. By Lemma 8, which is given below, we have

$$\mathbf{E}[\text{WFA}[\mathcal{S}] \mid \mathcal{E}]\ \mathbf{P}[\mathcal{E}] \ \le\ \mathbf{E}[\text{WFA}[\mathcal{S}]] \ \le\ c_1\ell\beta U. \tag{18}$$

We estimate

$$\begin{aligned}\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] &= \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\,\middle|\,\mathcal{E}\right]\mathbf{P}[\mathcal{E}] + \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\,\middle|\,\neg\mathcal{E}\right]\mathbf{P}[\neg\mathcal{E}]\\ &\le \frac{\mathbf{E}[\text{WFA}[\mathcal{S}] \mid \mathcal{E}]\ \mathbf{P}[\mathcal{E}]}{n\gamma U} + \frac{2n-1}{n^3} = O\left(\beta\left(\frac{U}{\sigma} + \log(D)\right)\right).\end{aligned}$$

The second inequality follows from the definition of $\mathcal{E}$, the fact that the competitive ratio of WFA is $2n-1$ and Lemma 1. The last inequality is due to (18) and the definition of ℓ. □

Lemma 8. *Let G be an arbitrary graph with n nodes, uniform edge lengths U, and maximum degree D. Let $\bar{\mathcal{S}}$ be an adversarial sequence of length ℓ only consisting of β-elementary tasks. Then, for $\ell \ge n-1$, the expected cost incurred by* WFA *on a smoothed task sequence $\mathcal{S}$ is at most $c_1\ell\beta U$ for some constant c_1, i.e., $\mathbf{E}[\text{WFA}[\mathcal{S}]] \le c_1\ell\beta U$.*

Proof. We use a potential function argument. Note that although each of the adversarial request vectors has at most β non-zero entries, it is possible that in some rounds, WFA has to travel large distances irrespective of β. This is because the behavior of WFA is tied to the way the work function values change. This is true even if the input request is not smoothed.

Let $\bar{\mathcal{S}}$ denote the adversarial request sequence of length ℓ. Let the random variable $\mathcal{S} = \langle\tau_1, \ldots, \tau_\ell\rangle$ denote a smoothed sequence obtained from $\bar{\mathcal{S}}$. Here $\mathcal{S}$ is a random variable. Let $w_0(\cdot)$ denote an arbitrary configuration of the work function values in the beginning. For $t \in \{0, \ldots, \ell\}$, let s_t denote the node in which WFA resides after round t. We use s_0 to denote the state where WFA is present at the beginning of this phase. We define a potential function Φ for any round t as

$$\Phi(t) = w_t(s_t) + tU.$$

Observe that $\Phi(\ell) - \Phi(0) = w_\ell(s_\ell) - w_0(s_0) + \ell U \geq w_\ell(s_\ell) - w_\ell(s_0) + \ell U$. Due to Fact 3, the latter term is non-negative for $\ell \geq n-1$. We define the amortized cost $C_a(t)$ incurred by WFA for task τ_t as

$$\begin{aligned} C_a(t) &= r_t(s_t) + \delta(s_{t-1}, s_t) + \Phi(t) - \Phi(t-1) \\ &= r_t(s_t) + \delta(s_{t-1}, s_t) + w_t(s_t) - w_{t-1}(s_{t-1}) + U \\ &= w_t(s_t) - w_{t-1}(s_t) + w_t(s_{t-1}) - w_{t-1}(s_{t-1}) + U \end{aligned} \tag{19}$$

where the last equality follows from (2) and 4. Note that $C_a(t)$ is a random variable. Our goal is to bound $\mathbf{E}[C_a(t)]$. For round t, let us consider two random variables, namely, $\Delta(s_t) = \min_{v \in V} \{r_t(v) + \delta(v, s_t)\}$ and $\Delta(s_{t-1}) = \min_{v \in V} \{r_t(v) + \delta(v, s_{t-1})\}$. We observe that for node s_t, $w_t(s_t) - w_{t-1}(s_t) \leq 2\Delta(s_t)$. This is obtained from the following two inequalities. Let m be a node which satisfies the $\Delta(s_t)$ expression. Using Fact 3 with respect to s_t and m, we have that $w_{t-1}(s_t) \geq w_{t-1}(m) - \delta(m, s_t)$. Further more, using the Fact 2 for m and Fact 3 with respect to s_t and m, we also have that $w_t(s_t) \leq w_{t-1}(m) + r_t(m) + \delta(m, s_t)$. Combining these two, we obtain the desired inequality. Using analogous arguments for node s_{t-1}, we also obtain that $w_t(s_{t-1}) - w_{t-1}(s_{t-1}) \leq 2\Delta(s_{t-1})$. Hence we can rewrite (19) as

$$C_a(t) \leq 2\Delta(s_t) + 2\Delta(s_{t-1}) + U.$$

Next, we show that $\mathbf{E}[\Delta(s_t)]$ and $\mathbf{E}[\Delta(s_{t-1})]$ are at most $(\beta+2)U$. More precisely, in Proposition 1 below, we show that $\mathbf{E}[\Delta(s_t)] \leq (\beta+1)U + (\beta U + 2\sigma)/2^{n-\beta}$. This is at most $(\beta+2)U$ for $\beta \leq \sqrt{n}$ and $\sigma \leq U 2^{n-\sqrt{n}-2}$ and for n greater than a fixed constant. Analogously, we obtain the same bound for $\mathbf{E}[\Delta(s_{t-1})]$. Since $\text{WFA}[\mathcal{S}] = \sum_{t=1}^{\ell} C_a - \Phi(\ell) + \Phi(0)$ and since $\Phi(\ell) - \Phi(0)$ is always non-negative,

$$\mathbf{E}[\text{WFA}[\mathcal{S}]] \leq \mathbf{E}\left[\sum_{t=1}^{\ell} C_a(t)\right] \leq 2\mathbf{E}\left[\sum_{t=1}^{\ell} (\Delta(s_t) + \Delta(s_{t-1}))\right] + \ell U \leq 9\ell\beta U.$$

□

Proposition 1. *Consider a graph G with n nodes and uniform edge lengths U. Assume that the adversarial task sequence is β-elementary and fix a round t. Let s be an arbitrary node of G and define a random variable $\Delta(s) = \min_{v \in V} \{r_t(v) + \delta(v, s)\}$ Then, $\mathbf{E}[\Delta(s)] \leq (\beta+1)U + \frac{\beta U + 2\sigma}{2^{n-\beta}}$.*

Proof. Consider an β-elementary task. Order the nodes with zero cost entries by increasing distances from s. For the ith node in the ordering, the distance from s is at most $(\beta + i - 1)U$. Each zero cost component of the input vector is retained with probability at least 1/2. If we sum the expectation over the mutually exclusive events, namely that the ith node is the first in the ordering that retains zero costs and the last case that all nodes have cost more than zero, we get that $\mathbf{E}[\Delta(s)] \leq (\beta+1)U + \frac{\beta U + 2\sigma}{2^{n-\beta}}$. The second term follows because, for the last case, for any arbitrary node in the ordering, the conditional expected cost is at most 2σ using Fact 5. □

J Proof of Theorem 11

Proof (Theorem 11). We can assume that $U/\sigma \leq n$. Otherwise we have the tight lower bound of $\Omega(n)$ on any graph by Theorem 8. In addition, since $\beta \leq \sqrt{n/(U/\sigma + 1)}$, we have that $\beta(U/\sigma + 1) \leq n$. Hence for our lower bound expression to be proved, the minimum is always due to the second item. Hence we only need to show a lower bound of $\Omega(\beta(U/\sigma + 1))$. Fix any

deterministic algorithm ALG. For the case $U/\sigma \geq 1$, we will a lower bound of $\Omega(\beta U/\sigma)$ and for the complement case we will show a bound of $\Omega(\beta)$. Putting them together we have the desired result.
We will use the standard averaging technique to show the lower bound. We consider only a small collection of offline algorithms that is appropriate for deriving the bound. Divide the line into $h := n/2\beta$ contiguous segments of 2β nodes. For simplicity assume that h is an integer. This does not affect the asymptotic lower bound. Let $S_1, S_2, \ldots, S_h$ denote these segments in order. For any round i, let v be the node where ALG ends up after round $i-1$. Then in round i, the β-elementary task will have ∞ cost component on v as well as all the nodes that are within a distance of $\lceil \beta/2 \rceil - 1$ from it. We recall that the adversary is adaptive. All other cost component in it are zeroes. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote the ℓ tasks after smoothing.

Let $\mathbf{B}$ be a set of h offline algorithms (not necessarily optimal), one for each of the h segments. Let $\mathbf{B}_i$ denote the offline algorithm which is always present in segment S_i. Let $\mathbf{B}[\mathcal{S}] = \sum_{i=1}^{h} \mathbf{B}_i[\mathcal{S}]$, where $\mathbf{B}_i[\mathcal{S}]$ is a random variable which denotes the total cost incurred by $\mathbf{B}_i$. Clearly, the value $\mathbf{B}[\mathcal{S}]/h$ is an upper bound for OPT$[\mathcal{S}]$. Let $\mathbf{B}[\mathcal{S}]/h$ be denoted as $\mathbf{B}_{av}$. In every round, each $\mathbf{B}_i$ travel to to the nearest node in the same segment where the request cost is at most U. If no such node is present, it travels to a node in that segment with minimum request cost.

Consider any round t. All the ∞ cost components of the vector are present in at most two consecutive line segments. Moreover, in each of such segments, at most β of its 2β nodes can get ∞ cost. Let $b_j(t)$ denote the cost incurred by algorithm $\mathbf{B}_j$ in round t. Let S_j be a segment that does not get any ∞ cost in round t. If $\sigma \leq U$ then $\mathbf{E}[b_j(t)] \leq \sigma$ by Fact 5 and by the fact that $b_j(t)$ is at most the cost on the position where $\mathbf{B}_j$ is present in the beginning of this round. If $\sigma > U$, by using Proposition 1 on this segment, since $\sigma \leq 2U$ we get that $\mathbf{E}[b_j(t)] \leq 3U$. If S_j happens to be a segment which contains ∞ cost, then $\mathbf{E}[b_j(t)] \leq \beta U + \sigma$ if $\sigma \leq U$ and $\mathbf{E}[b_j(t)] \leq \beta U + 3U$ otherwise by Proposition 1. Since there are at most 2 'bad' segments which can get ∞ cost in any round, we can conclude that

$$\mathbf{E}[\mathbf{B}_{av}] = \frac{1}{h}\mathbf{E}[\mathbf{B}[\mathcal{S}]] = \frac{1}{h}\mathbf{E}\left[\sum_{i=1}^{h} \mathbf{B}_i[\mathcal{S}]\right] = \frac{1}{h}\mathbf{E}\left[\sum_{t=1}^{\ell}\sum_{i=1}^{h} b_i(t)\right] \leq \ell\left(\frac{2\beta U}{h} + \sigma\right) \quad (\text{if } \sigma \leq U)$$

and if $\sigma > U$, then

$$\mathbf{E}[\mathbf{B}_{av}] \leq \ell\left(2\beta U/h + 3U\right).$$

We derive separate lower bound depending on these two cases. By Markov inequality, we have $\mathbf{P}[\mathbf{B}_{av} \geq 2\mathbf{E}[\mathbf{B}_{av}]] \leq 1/2$. In any case, ALG$[\mathcal{S}]$ is at least $\frac{\ell\beta U}{2}$ because of the ∞ cost in its neighborhood.
For the case $\sigma \leq U$, we can write that

$$\begin{aligned}\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] &\geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}} \,\middle|\, \mathbf{B}_{av} \leq 2\ell\left(\frac{2\beta U}{h} + \sigma\right)\right]\left(\frac{1}{2}\right) \\ &\geq \frac{1}{4}\left(\frac{n\beta U}{n\sigma + 4U\beta^2}\right) \geq b\left(\frac{\beta U}{\sigma}\right)\end{aligned}$$

for an appropriate constant b. This follows from the facts (a) $h = n/2\beta$; (b) ALG$[\mathcal{S}]$ is always at least $\frac{\ell\beta U}{2}$; (c) $\beta \leq \sqrt{\frac{n}{U/\sigma+1}}$; and (d) $\sigma \leq U$. Based on these facts, the expression follows irrespective of which one is the dominating term in the denominator sum. Using the same facts and analogous arguments, for the case $\sigma > U$, we get,

$$\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}} \,\middle|\, \mathbf{B}_{av} \leq 2\ell\left(\frac{2\beta U}{h} + 3U\right)\right]\left(\frac{1}{2}\right) \geq c \cdot \beta$$

for an appropriate constant c. Hence the lower bound is shown. □

K Proof of Theorem 12

Proof (Theorem 12). Let $\bar{\mathcal{S}}$ be any adversarial task of length ℓ. Let the random variable $\mathcal{S} = \langle \tau_1, \ldots, \tau_\ell \rangle$ denote a smoothed task sequence obtained from $\bar{\mathcal{S}}$. For convenience assume that ℓ is as defined in Lemma 1 for some $\gamma \geq 1$. For getting a lower bound for the cost of OPT, we assume that all the edge lengths in the graph are U_{min}. Clearly, a lower bound on such a graph is a lower bound on the original graph. On such graph, we can apply Lemma 1. Let $\mathcal{E}$ be an event that the cost incurred by OPT on $\mathcal{S}$ is at least $n\gamma U_{min}$. From Lemma 1, we have that $\mathbf{P}[\mathcal{E}] \geq 1 - 1/n^3$. Let $\mathcal{S}_\mathcal{E}$ be a fixed outcome of smoothing $\bar{\mathcal{S}}$ conditioned on the event $\mathcal{E}$. From the definition of $\mathcal{E}$, we have that $\text{OPT}[\mathcal{S}_\mathcal{E}] \geq n\gamma U_{min}$. Moreover, from Proposition 2, we also have that $\text{WFA}[\mathcal{S}_\mathcal{E}] \leq \text{OPT}[\mathcal{S}_\mathcal{E}] + Diam \cdot \ell$. By using the relation between γ and ℓ as stated in Lemma 1, we have that

$$\frac{\text{WFA}[\mathcal{S}_\mathcal{E}]}{\text{OPT}[\mathcal{S}_\mathcal{E}]} \leq b \cdot \frac{Diam}{U_{min}} \left(U/\sigma + \log(D)\right) \leq b \cdot \frac{Diam}{U_{min}} \left(U/\sigma + \log(n)\right)$$

for an appropriately fixed constant positive constant b. Hence,

$$\begin{aligned}
\mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] &= \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \,\middle|\, \mathcal{E}\right] \mathbf{P}[\mathcal{E}] + \mathbf{E}\left[\frac{\text{WFA}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]} \,\middle|\, \neg\mathcal{E}\right] \mathbf{P}[\neg\mathcal{E}] \\
&\leq b \cdot \frac{Diam}{U_{min}} \cdot \left(U_{min}/\sigma + \log(n)\right) + (2n - 1)\left(1/n^3\right) \\
&= O\left(\frac{Diam}{U_{min}} \cdot \left(U_{min}/\sigma + \log(n)\right)\right)
\end{aligned}$$

□

L Proof of Theorem 13

Proof (Theorem 13). The proof is exactly like in Lemma 6 and the graph also is similar. Each of the cliques have size $\mathcal{D} := \Theta(n^\epsilon)$ which is at least the fixed constant b as required in the lemma for any large n. Within the clique, all edge lengths are U_{min}. There are $m := \Theta\left(n^{1-\epsilon}\right)$ such cliques. Each clique is connected to a central node say v_c by a single edge of length $Diam/2 - U_{min}$. We do not need the central connection tree here.

The cost generation is also identical here in the sense that always there is an ∞ cost for the central node as well as all the nodes of the clique where ALG resides after the previous round. We compare the cost of ALG with a collection of m offline algorithms. To avoid repetition, we only sketch the major steps of the proof which follows by the same arguments. Let the final sequence $\mathcal{S}$ of length ℓ be viewed as generated in two stages (as in the previous proof). Let $\mathcal{S}'$ be the random variable which denote the sequence after stage 1. We also define here the similar event $\mathcal{E}$ which is $\mathbf{B}_m[\mathcal{S}'] + \sum_{i=1}^{m-1} \mathbf{B}_i[\mathcal{S}'] \leq \frac{12m\ell U_{min}}{\log(\mathcal{D})}$. By using Lemma 7, we have that $\mathbf{P}[\mathcal{E}] \geq 1/2$. Using similar arguments we get that

$$\mathbf{B}[\mathcal{S}] = \sum_{i=1}^{m-1} \mathbf{B}_i[\mathcal{S}] \leq \ell \cdot Diam + \mathbf{B}_m[\mathcal{S}'] + \sum_{i=1}^{m-1} \mathbf{B}_i[\mathcal{S}'].$$

Hence under the event $\mathcal{E}$, we have that $\mathbf{B}[\mathcal{S}] \leq \ell \cdot Diam + \frac{12m\ell U_{min}}{\log(\mathcal{D})}$ and $\text{ALG}[\mathcal{S}] \geq \ell \cdot Diam$ (because of ∞ costs). Hence we write that conditioned on the event $\mathcal{E}$,

$$\mathbf{B}_{av} = \frac{\mathbf{B}[\mathcal{S}]}{m-1} \leq \frac{\text{ALG}[\mathcal{S}]}{m-1} + \frac{24\ell U_{min}}{\log(\mathcal{D})}$$

Finally we end up with the claimed bound using the relation, $\mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\text{OPT}[\mathcal{S}]}\right] \geq \mathbf{E}\left[\frac{\text{ALG}[\mathcal{S}]}{\mathbf{B}_{av}} \middle| \mathcal{E}\right] \mathbf{P}[\mathcal{E}]$. Recall that $\mathcal{D} = n^{\epsilon}$. □

M

Proposition 2. *On a graph $G = (V, E)$ with diameter $Diam$, for any ℓ and for any request sequence $\mathcal{K}$ of length ℓ,* $\text{WFA}[\mathcal{K}] \leq \text{OPT}[\mathcal{K}] + Diam \cdot \ell$. *In particular, for graphs with uniform edge length U and edge diameter $diam$, we have* $\text{WFA}[\mathcal{K}] \leq \text{OPT}[\mathcal{K}] + diam \cdot U\ell$.

Proof. Let the starting round be denoted as 1 and let $s_0, s_1, \ldots, s_\ell$ denote the sequence of nodes visited by WFA (WFA starts at s_0). For any round t, let $w_t^{min} := \min_{\forall v \in G} w_t(v)$. Let $C(t)$ denote the cost incurred by WFA in round t. From the description of the WFA algorithm and (2), we know that for any round t, $C(t) \leq Diam + w_t^{min}$. Since $\text{OPT}[\mathcal{K}] \geq w_\ell^{min}$, and since $w(\cdot)$ values are non-negative, we conclude that $\text{WFA}[\mathcal{K}] = \sum_{t=1}^{C}(t) \leq Diam \cdot \ell + \text{OPT}[\mathcal{K}]$. □

N Constant additive cost for the offline optimum

We would like to point out that in all our lower bound proofs, we can assume without loss of generality that OPT has an additive cost of say α which is independent of the length of the input sequence. This does not change the bound in the asymptotic sense. This is because of the following. In all the lower bound proofs, if we want to show a lower bound of say $\Omega(y/x)$, we show that for any (smoothed) input sequence of length ℓ, with at least a constant probability, the cost incurred by OPT is at most $x\ell$; call this event $\mathcal{E}$. Moreover, under this event we also show that the expected cost incurred by any algorithm ALG is at least $y\ell$. Now to take care of the constant additive cost of α for OPT, we only need to make sure that the length of the input sequence is long enough. That is, ℓ is such that $x\ell \geq \alpha$. Hence we can conclude that $\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] \geq \mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}} \middle| \mathcal{E}\right] \mathbf{P}[\mathcal{E}] \geq \frac{2y\ell}{x\ell+\alpha}$. Since $x\ell \geq \alpha$, we have that $\mathbf{E}\left[\frac{\text{ALG}}{\text{OPT}}\right] \geq 4y/x$ which is still $\Omega(y/x)$.

O Estimations

Fact 5. *Let $\mathcal{P}$ be a continuous distribution with mean $\mu = 0$ and symmetric around the mean. Let σ be its standard deviation. Let x be a random variable with distribution $\mathcal{P}$. Let y be another random variable defined as $y := \max\{0, x\}$. Then,* $\mathbf{E}[y] \leq \sigma$.

Proof. Since $\mu = 0$, $\sigma^2 = \mathbf{E}\left[x^2\right]$. Let $\sigma(y)$ denote the standard deviation for the distribution of y. From the definition of $\mathbf{E}\left[x^2\right]$, we see that $\mathbf{E}\left[y^2\right] = \mathbf{E}\left[x^2\right]/2$. Since $\sigma^2(y) = \mathbf{E}\left[y^2\right] - \mathbf{E}[y]^2$, and since $\sigma^2(y) \geq 0$, we have that $\mathbf{E}[y]^2 \leq \mathbf{E}\left[y^2\right]$. This in turn means that $\mathbf{E}[y] \leq \sigma/\sqrt{2}$. □

Fact 6. *Let $\mathcal{P}$ be a continuous non-negative distribution which is non increasing in $[0, \infty)$. Let μ and σ be its mean and standard deviation. Then $\mu \leq \sqrt{12}\,\sigma$.*

Proof. Let X be a random variable which follows the distribution $\mathcal{P}$. Fix $\alpha = \mu/2\sigma$. Let $\mathcal{E}$ denote the event that $|X - \mu| \geq \mu/2$. Using Chebyshev Inequality, we have that

$$\mathbf{P}[\mathcal{E}] = \mathbf{P}[|X - \mu| \geq \mu/2] \leq 4\sigma^2/\mu^2 \tag{20}$$

Since $\mathcal{P}$ is continuous and non-increasing in $[0, \infty)$,

$$\mathbf{P}[\mathcal{E}] = \mathbf{P}[|X - \mu| \geq \mu/2] \geq \mathbf{P}[X \leq \mu/2] \geq 1/2\mathbf{P}[\mu/2 < X \leq 3/2\mu] \geq 1/2\mathbf{P}[\neg\mathcal{E}]\,.$$

That is, $\mathbf{P}[\mathcal{E}] \geq 1/2\mathbf{P}[\neg\mathcal{E}]$. This implies that $\mathbf{P}[\mathcal{E}] \geq 1/3$. Hence (20) gives that $\mu^2 \leq 12\sigma^2$. Since both μ and σ are non-negative, we have the result. □

Lemma 9. *Let*

$$\mathcal{X} := \min\left\{\frac{A\sum_{i=1}^{m} X_i}{\sum_{i=1}^{m} X_i^2},\ \frac{B\sum_{i=1}^{m} X_i}{m}\right\},$$

where A,B,X_i's and m are all positive quantities. Then,

$$\mathcal{X} \leq \sqrt{A\,B}$$

Proof. First we will show that

$$m(X_1^2 + X_2^2 + \cdots + X_m^2) \geq (X_1 + X_2 + \cdots + X_m)^2. \tag{21}$$

This is because this can be re-written as

$$(1/2)\sum_{\forall i,j}\left(X_i^2 + X_j^2\right) \geq \sum_{i=1}^{m} X_i^2 + \sum_{\forall i,j; i\neq j} X_i X_j.$$

which is always true. Define $Y := \frac{\sum_{i=1}^{m} X_i}{m}$. Note that Y is positive. Because of (21), we can write that,

$$\mathcal{X} \leq \min\left\{\frac{A}{Y},\ B\,Y\right\}.$$

The above expression is maximized when $\frac{A}{Y} = BY$. This means that at maximum, $Y = \sqrt{\frac{A}{B}}$. In other words, at maximum, we have that $\mathcal{X} \leq \sqrt{A\,B}$. Hence the lemma is shown. □