APPENDIX A
PROOF OF LEMMA 1

Proof. Consider the estimate $\hat{O}_w(t)$ that the algorithm uses at node v for neighbor w at time t. By definition of $T_{\max}$, the measurement is based on clock values $L_v(t_v)$ and $L_w(t_w)$ for some $t_v, t_w \in [t - T_{\max}, t)$. Without loss of generality, we assume that to measure whether $L_w - L_v \geq T \in \mathbf{R}$, the signals are sent at logical times satisfying $L_w(t_w) - T = L_v(t_v)$.[4] Denote by $t'_v \in (t_v, t)$ and $t'_w \in (t_w, t)$ the times when the respective signals arrive at the data or clock input, respectively, of the register[5] indicating whether $\hat{O}_w \geq T$ for a given threshold T. By definition of δ_0, we have that

$$|t'_v - t_v - (t'_w - t_w)| \leq \delta_0.$$

Note that the register indicates $\hat{O}_w(t) \geq T$, i.e., latches 1, if and only if $t'_w < t'_v$.[6] Thus, we need to show

$$L_w(t) - L_v(t) \geq T + \delta \implies t'_w < t'_v$$
$$L_w(t) - L_v(t) \leq T - \delta \implies t'_w > t'_v.$$

Assume first that $L_w(t) - L_v(t) \geq T + \delta$. Then, using **I4** and that $L_w(t_w) - T = L_v(t_v)$, we can bound

$$\begin{aligned}
T + \delta &\leq L_w(t) - L_v(t) \\
&\leq L_w(t_v) - L_v(t_v) + ((1+\mu)(1+\rho) - 1)(t - t_v) \\
&= L_w(t_v) - L_w(t_w) + T + (\mu + \rho + \rho\mu)(t - t_v) \\
&\leq t_v - t_w + T + (\mu + \rho + \rho\mu)(t - \min\{t_v, t_w\}) \\
&< t_v - t_w + T + (\mu + \rho + \rho\mu)T_{\max}.
\end{aligned}$$

Hence,

$$t'_w - t'_v \geq t_w - t_v - \delta_0 > \delta - \delta_0 - (\mu + \rho + \rho\mu)T_{\max} = 0.$$

For the second implication, observe that it is equivalent to

$$L_v(t) - L_w(t) \geq -T + \delta \implies t'_v > t'_w.$$

As we have shown the first implication for any $T \in \mathbf{R}$, the second follows analogously by exchanging the roles of v and w. $\square$

APPENDIX B
PROOF OF THEOREM 1

In this appendix, we prove Theorem 1. We assume that at (Newtonian) time $t = 0$, the system satisfies *some* bound on local skew. The analysis we provide shows that the GCS algorithm maintains a (slightly larger) bound on local skew for all $t \geq 0$. An upper bound on the local skew also bounds the number of values of s for which **FC** or **SC** (Definition 1) can hold, as a large s implies a large local skew. (For example, if a node v satisfies **FC1** for some s, then v has a neighbor x satisfying $L_x(t) - L_v(t) \geq (2s+1)\kappa$, implying that $\mathcal{L}(t) \geq (2s+1)\kappa$.) Accordingly, an implementation need only test for values of s satisfying $|s| < \frac{1}{2\kappa}\mathcal{L}_{\max}$, where $\mathcal{L}_{\max}$ is a theoretical upper bound on the local skew. Our analysis also shows that given an arbitrary initial global skew $\mathcal{G}(0)$, the system will converge to the skew bounds claimed in Theorem 1 within time $O(\mathcal{G}(0)/\mu)$. We note that the skew upper bounds of Theorem 1 match the theoretical lower bounds of [11] up to a factor of approximately 2, and the theoretical lower bounds can be acheived even by systems with initially perfect synchronization (i.e., systems with $\mathcal{L}(0) = \mathcal{G}(0) = 0$).

Our analysis also assumes that logical clocks are differentiable functions. This assumption is without loss of generality: By the Stone-Weierstrass Theorem (cf. Theorem 7.26 in [20]) every continuous function on a compact interval can be approximated arbitrarily closely by a differentiable function.

We will rely on the following technical result. We provide a proof in Section B-E.

Lemma 3. *For $k \in \mathbf{Z}$ and $t_0, t_1 \in \mathbf{R}_{\geq 0}$ with $t_0 < t_1$, let $\mathcal{F} = \{f_i \mid i \in [k]\}$, where each $f_i \colon [t_0, t_1] \to \mathbf{R}$ is a differentiable function. Define $F \colon [t_0, t_1] \to \mathbf{R}$ by $F(t) = \max_{i \in [k]}\{f_i(t)\}$. Suppose $\mathcal{F}$ has the property that for every i and t, if $f_i(t) = F(t)$, then $\frac{d}{dt} f_i(t) \leq r$. Then for all $t \in [t_0, t_1]$, we have $F(t) \leq F(t_0) + r(t - t_0)$.*

Throughout this section, we assume that each node runs an algorithm satisfying the invariants stated in Definition 2. By Lemma 2, Algorithm 1 meets this requirement if $\kappa > 2\delta + 2(\rho + \mu + \rho\mu)T_{\max}$.

A. Leading Nodes

We start by showing that skew cannot build up too quickly. This is captured by analyzing the following functions.

Definition 4 (Ψ and Leading Nodes)**.** *For each $v \in V$, $s \in \mathbf{N}$, and $t \in \mathbf{R}_{\geq 0}$, we define*

$$\Psi^s_v(t) = \max_{w \in V}\{L_w(t) - L_v(t) - 2s\kappa d(v, w)\},$$

where $d(v, w)$ denotes the distance between v and w in G. Moreover, set

$$\Psi^s(t) = \max_{v \in V}\{\Psi^s_v(t)\}.$$

Finally, we say that $w \in V$ is a leading node *if there is some $v \in V$ satisfying*

$$\Psi^s_v(t) = L_w(t) - L_v(t) - 2s\kappa d(v, w) > 0.$$

Observe that any bound on Ψ^s implies a corresponding bound on $\mathcal{L}$: If $\Psi^s(t) \leq \kappa$, then for any adjacent nodes v, w we have $L_w(t) - L_v(t) - 2s\kappa \leq \Psi^s(t) \leq \kappa$. Therefore, $\Psi^s(t) \leq \kappa \implies \mathcal{L} \leq (2s+1)\kappa$. Our analysis will show that in general, $\Psi^s(t) \leq \mathcal{G}_{\max}/\sigma^s$ for every $s \in \mathbf{N}$ and all times t. In particular, considering $s = \lceil \log_{\mu/\rho} \mathcal{G}_{\max}/\kappa \rceil$ gives a bound on $\mathcal{L}$ in terms of $\mathcal{G}_{\max}$. Because $\mathcal{G}(t) = \Psi^0(t)$, the skew bounds will then follow if we can suitably bound Ψ^0 at all times.

Note that the definition of Ψ^s_v is closely related to the definition of the slow condition. In fact, the following lemma

[4] One can account for asymmetric propagation times by shifting $L_w(t_w)$ and $L_v(t_v)$ accordingly, so long as this is accounted for in $T_{\max}$ and carry out the proof analogously.

[5] We assume a register here, but the same argument applies to any state-holding component serving this purpose in the measurement circuit.

[6] For simplicity of the presentation we neglect the setup/hold time ε (accounted for in δ_0) and metastability; see Section III for a discussion.

shows that if w is a leading node, then w satisfies the slow condition. Thus, Ψ^s cannot increase quickly: **I4** (Def. 2) then stipulates that leading nodes increase their logical clocks at rate at most $1+\rho$. This behavior allows nodes in fast mode to catch up to leading nodes.

Lemma 4 (Leading Lemma)**.** *Suppose $w \in V$ is a leading node at time t. Then $\frac{\mathrm{d}}{\mathrm{d}t}L_w(t) = \frac{\mathrm{d}}{\mathrm{d}t}H_w(t) \in [1, 1+\rho]$.*

Proof. By **I4**, the claim follows if w satisfies the slow condition at time t. As w is a leading node at time t, there are $s \in \mathbf{N}$ and $v \in V$ satisfying

$$\Psi_v^s(t) = L_w(t) - L_v(t) - 2s\kappa d(v,w) > 0.$$

In particular, $L_w(t) > L_v(t)$, so $w \neq v$. For any $y \in V$, we have

$$\begin{aligned} L_w(t) - L_v(t) - 2s\kappa d(v,w) &= \Psi_v^s(t) \\ &\geq L_y(t) - L_v(t) - 2s\kappa d(y,w). \end{aligned}$$

Rearranging this expression yields

$$L_w(t) - L_y(t) \geq 2s\kappa(d(v,w) - d(y,w)).$$

In particular, for any $y \in N_v$, $d(v,w) \geq d(y,w) - 1$ and hence

$$L_y(t) - L_w(t) \leq 2s\kappa,$$

i.e., SC2 holds for s at w.

Now consider $x \in N_v$ so that $d(x,w) = d(v,w) - 1$. Such a node exists because $v \neq w$. We obtain

$$L_w(t) - L_y(t) \geq 2s\kappa.$$

Thus SC1 is satisfied for s, i.e., indeed the slow condition holds at w at time t. □

Lemma 4 can readily be translated into a bound on the growth of Ψ_w^s whenever $\Psi_w^s > 0$.

Lemma 5 (Wait-up Lemma)**.** *Suppose $w \in V$ satisfies $\Psi_w^s(t) > 0$ for all $t \in (t_0, t_1]$. Then*

$$\Psi_w^s(t_1) \leq \Psi_w^s(t_0) - (L_w(t_1) - L_w(t_0)) + (1+\rho)(t_1 - t_0).$$

Proof. Fix $w \in V$, $s \in \mathbf{N}$ and $(t_0, t_1]$ as in the hypothesis of the lemma. For $v \in V$ and $t \in (t_0, t_1]$, define the function $f_v(t) = L_v(t) - 2s\kappa d(v,w)$. Observe that

$$\max_{v \in V}\{f_v(t)\} - L_w(t) = \Psi_w^s(t)\,.$$

Moreover, for any v satisfying $f_v(t) = L_w(t) + \Psi_w^s(t)$, we have $L_v(t) - L_w(t) - 2s\kappa d(v,w) = \Psi_w^s(t) > 0$. Thus, Lemma 4 shows that v is in slow mode at time t. As (we assume that) logical clocks are differentiable, so is f_v, and it follows that $\frac{\mathrm{d}}{\mathrm{d}t} f_v(t) \leq 1+\rho$ for any $v \in V$ and time $t \in (t_0, t_1]$ satisfying $f_v(t) = \max_{x \in V}\{f_x(t)\}$. By Lemma 3, it follows that $\max_{v \in V}\{f_v(t)\}$ grows at most at rate $1+\rho$:

$$\max_{v \in V}\{f_v(t_1)\} \leq \max_{v \in V}\{f_v(t_0)\} + (1+\rho)(t_1 - t_0)\,.$$

We conclude that

$$\begin{aligned} \Psi_w^s(t_1) - \Psi_w^s(t_0) &= \max_{v \in V}\{f_v(t_1)\} - L_w(t_1) \\ &\quad - (\max_{v \in V}\{f_v(t_0)\} - L_w(t_0)) \\ &\leq (1+\rho)(t_1 - t_0) - (L_w(t_1) - L_w(t_0)), \end{aligned}$$

which can be rearranged into the desired result. □

Corollary 2. *For all $s \in \mathbf{N}$ and times $t_1 \geq t_0$, $\Psi^s(t_1) \leq \Psi^s(t_0) + \rho(t_1 - t_0)$.*

Proof. Choose $w \in V$ such that $\Psi^s(t_1) = \Psi_w^s(t_1)$. As $\Psi_w^s(t) \geq 0$ for all times t, nothing is to show if $\Psi^s(t_1) = 0$. Let $t \in [t_0, t_1)$ be the supremum of times from $t' \in [t_0, t_1)$ with the property that $\Psi_w^s(t') = 0$. Because Ψ_w^s is continuous, $t \neq t_0$ implies that $\Psi_w^s(t) = 0$. Hence, $\Psi_w^s(t) \leq \Psi_w^s(t_0)$. By **I2** and Lemma 5, we get that

$$\begin{aligned} \Psi^s(t_1) &= \Psi_w^s(t_1) \\ &\leq \Psi_w^s(t) - (L_w(t_1) - L_w(t)) + (1+\rho)(t_1 - t) \\ &\leq \Psi_w^s(t) + \rho(t_1 - t) \\ &\leq \Psi_w^s(t_0) + \rho(t_1 - t_0) \\ &\leq \Psi^s(t_0) + \rho(t_1 - t_0). \end{aligned}$$

□

Trailing Nodes

As $L_w(t_1) - L_w(t_0) \geq t_1 - t_0$ at all times by **I2**, Lemma 7 implies that Ψ^s cannot grow faster than at rate ρ when $\Psi^s(t) > 0$. This means that nodes whose clocks are far behind leading nodes can catch up, so long as the lagging nodes satisfy the fast condition and thus run at rate at least $1+\mu$ by **I3**. Our next task is to show that "trailing nodes" always satisfy the fast condition so that they are never too far behind leading nodes. The approach to showing this is similar to the one for Lemma 5, where now we need to exploit the fast condition.

Definition 5 (Ξ and Trailing Nodes)**.** *For each $v \in V$, $s \in \mathbf{N}$, and $t \in \mathbf{R}_{\geq 0}$, we define*

$$\Xi_v^s(t) = \max_{w \in V}\{L_v(t) - L_w(t) - (2s+1)\kappa d(v,w)\},$$

where $d(v,w)$ denotes the distance between v and w in G. Moreover, set

$$\Xi^s(t) = \max_{v \in V}\{\Xi_v^s(t)\}.$$

Finally, we say that $w \in V$ is a trailing node *at time t, if there is some $v \in V$ satisfying*

$$\Xi_v^s(t) = L_v(t) - L_w(t) - (2s+1)\kappa d(v,w) > 0.$$

Lemma 6 (Trailing Lemma)**.** *If $w \in V$ is a trailing node at time t, then $\frac{\mathrm{d}}{\mathrm{d}t}L_w(t) = (1+\mu)\frac{\mathrm{d}}{\mathrm{d}t}H_w(t) \in [1+\mu, (1+\rho)(1+\mu)]$.*

Proof. By **I3**, it suffices to show that w satisfies the fast condition at time t. Let s and v satisfy

$$\begin{aligned} L_v(t) &- L_w(t) - (2s+1)\kappa d(v,w) \\ &= \max_{x \in V}\{L_v(t) - L_x(t) - (2s+1)\kappa d(v,x)\} > 0. \end{aligned}$$

In particular, $L_v(t) > L_w(t)$, implying that $v \neq w$. For $y \in V$, we have

$$L_v(t) - L_w(t) - (2s+1)\kappa d(v,w) \geq L_v(t) - L_y(t) - (2s+1)\kappa d(v,y).$$

Thus for all neighbors $y \in N_w$,

$$L_y(t) - L_w(t) + (2s+1)\kappa(d(v,y) - d(v,w)) \geq 0.$$

It follows that

$$\forall y \in N_v\colon L_w(t) - L_y(t) \leq (2s+1)\kappa,$$

i.e., FC2 holds for s. As $v \neq w$, there is some node $x \in N_v$ with $d(v,x) = d(v,w) - 1$. Thus we obtain

$$\exists x \in N_v\colon L_y(t) - L_w(t) \geq (2s+1)\kappa,$$

showing FC1 for s, i.e., indeed the fast condition holds at w at time t. □

Using Lemma 6, we can show that if $\Psi^s_w(t_0) > 0$, w will eventually catch up. How long this takes can be expressed in terms of $\Psi^{s-1}(t_0)$, or, if $s = 0$, $\mathcal{G}$.

Lemma 7 (Catch-up Lemma)**.** *Let $s \in \mathbf{N}$ and $v, w \in V$. Let t_0 and t_1 be times satisfying that*

$$t_1 \geq t_0 + \frac{\Xi^s_v(t_0)}{\mu}.$$

Then

$$L_w(t_1) \geq t_1 - t_0 + L_v(t_0) - (2s+1)\kappa d(v,w).$$

Proof. W.l.o.g., we may assume that $t_1 = t_0 + \Xi^s_v(t_0)/\mu$, as **I2** ensures that $\frac{\mathrm{d}}{\mathrm{dt}} L_w(t) \geq 1$ at all times, i.e., the general statement readily follows. For any $x \in V$, define

$$f_x(t) = t - t_0 + L_v(t_0) - L_x(t) - (2s+1)\kappa d(v,x).$$

Again by **I2**, it thus suffices to show that $f_w(t) \leq 0$ for some $t \in [t_0, t_1]$.

Observe that $\Xi^s_v(t_0) = \max_{x\in V}\{f_x(t_0)\}$. Thus, it suffices to show that $\max_{x\in V}\{f_x(t)\}$ decreases at rate μ so long as it is positive, as then $f_w(t_1) \leq \max_{x\in V}\{f_x(t_1)\} \leq 0$. To this end, consider any time $t \in [t_0, t_1]$ satisfying $\max_{x\in V}\{f_x(t)\} > 0$ and let $y \in V$ be any node such that $\max_{x\in V}\{f_x(t)\} = f_y(t)$. Then y is trailing, as

$$\begin{aligned}
\Xi^s_v(t) &= \max_{x\in V}\{L_v(t) - L_x(t) - (2s+1)\kappa d(v,x)\}\\
&= L_v(t) - L_v(t_0) - (t - t_0) + \max_{x\in V}\{f_x(t)\}\\
&= L_v(t) - L_v(t_0) - (t - t_0) + f_y(t)\\
&= L_v(t) - L_y(t) - (2s+1)\kappa d(v,y)
\end{aligned}$$

and

$$\begin{aligned}
\Xi^s_v(t) &= L_v(t) - L_v(t_0) - (t - t_0) + \max_{x\in V}\{f_x(t)\}\\
&> L_v(t) - L_v(t_0) - (t - t_0) \geq 0.
\end{aligned}$$

Thus, by Lemma 6 we have that $\frac{\mathrm{d}}{\mathrm{dt}} L_y(t) \geq 1 + \mu$, implying $\frac{\mathrm{d}}{\mathrm{dt}} f_y(t) = 1 - \frac{\mathrm{d}}{\mathrm{dt}} L_y(t) \leq -\mu$.

To complete the proof, assume towards a contradiction that $\max_{x\in V}\{f_x(t)\} > 0$ for all $t \in [t_0, t_1]$. Then, applying Lemma 3 again, we conclude that

$$\begin{aligned}
\Xi^s_v(t_0) &= \max_{x\in V}\{f_x(t_0)\}\\
&> -(\max_{x\in V}\{f_x(t_1)\} - \max_{x\in V}\{f_x(t_0)\})\\
&\geq \mu(t_1 - t_0) = \Xi^s_v(t_0),
\end{aligned}$$

i.e., it must hold that $f_w(t) \leq \max_{x\in V}\{f_x(t)\} \leq 0$ for some $t \in [t_0, t_1]$. □

B. Base Case and Global Skew

We now prove that if $\Psi^s(0)$ is bounded for some $s \in \mathbf{N}$, it cannot grow significantly and thus remains bounded. This will both serve as an induction anchor for establishing our bound on the local skew and for bounding the global skew, as $\Psi^0(t) = \mathcal{G}(t)$. In addition, we will deduce that even if the initial global skew $\mathcal{G}(0)$ is large, at times $t \geq \mathcal{G}(0)/\mu$, $\mathcal{G}(t)$ is bounded by $\mathcal{G}_{\max} = (1 - 2\rho/\mu)\kappa D$.

To this end, we will apply Lemma 7 in the following form.

Corollary 3. *Let $s \in \mathbf{N}$ and t_0, t_1 be times satisfying*

$$t_1 \geq t_0 + \frac{\Psi^s(t_0)}{\mu}.$$

Then, for any $w \in V$ we have

$$L_w(t_1) - L_w(t_0) \geq t_1 - t_0 + \Psi^s_w(t_0) - \kappa \cdot D.$$

Proof. If $\Psi^s_w(t_0) - \kappa \cdot D \leq 0$, the claim is trivially satisfied due to **I2** guaranteeing that $\frac{\mathrm{d}}{\mathrm{dt}} L_w(t) \geq 1$ at all times t. Hence, assume that $\Psi^s_w(t_0) - \kappa \cdot D > 0$ and choose any v so that

$$\Psi^s_w(t_0) = L_v(t) - L_w(t) - 2s\kappa d(v,w).$$

We have that

$$\begin{aligned}
\Xi^s_v(t_0) &\geq L_v(t) - L_w(t) - (2s+1)\kappa d(v,w)\\
&\geq L_v(t) - L_w(t) - 2s\kappa d(v,w) - \kappa \cdot D\\
&= \Psi^s_w(t_0) - \kappa \cdot D.
\end{aligned}$$

As trivially $\Psi^s(t_0) \geq \Xi^s(t_0) \geq \Xi^s_v(t_0)$, we have that $t_1 \geq t_0 + \Xi^s_v(t_0)/\mu$ and the claim follows by applying Lemma 7. □

Combining this corollary with Lemma 5, we can bound Ψ^s at all times.

Lemma 8. *Fix $s \in \mathbf{N}$. If $\Psi^s(0) \leq \kappa \cdot D/(1 - \rho^2/\mu^2)$, then*

$$\Psi^s(t) \leq \frac{\mu}{\mu - \rho} \cdot \kappa \cdot D.$$

at all times t. Otherwise,

$$\Psi^s(t) \leq \begin{cases} \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) & \text{if } t \leq \frac{\Psi^s(0)}{\mu}\\ \kappa \cdot D + \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) & \text{else.} \end{cases}$$

Proof. For $t \leq \Psi^s(0)/\mu$, the claim follows immediately from Corollary 2 (and possibly using that $\Psi^s(0) \leq \kappa \cdot D/(1 - \rho^2/\mu^2)$). Concerning larger times, denote by B the bound that needs to be shown and suppose that $\Psi^s(t_1) = B + \varepsilon$ for some

$\varepsilon > 0$ and minimal $t_1 > \Psi^s(0)/\mu$. Choose $w \in V$ so that $\Psi^s_w(t_1) = \Psi^s(t_1)$ and t_0 such that $t_1 = t_0 + \Psi^s(t_0)/\mu$. Such a time must exist, because the function $f(t) = t_1 - t - \Psi^s(t)/\mu$ is continuous and satisfies

$$f(t_1) = -\frac{\Psi^s(t_1)}{\mu} < 0 < t_1 - \frac{\Psi^s(0)}{\mu} = f(t_0).$$

We apply Lemma 5 and Corollary 3, showing that

$$\begin{aligned}\Psi^s_w(t_1) &\le \Psi^s_w(t_0) - (L_w(t_1) - L_w(t_0)) + (1+\rho)(t_1 - t_0)\\ &\le \kappa \cdot D + \rho(t_1 - t_0)\\ &= \kappa \cdot D + \frac{\rho}{\mu}\Psi^s(t_0).\end{aligned}$$

We distinguish two cases. If $\Psi^s(0) \le \kappa \cdot D/(1 - \rho^2/\mu^2)$, we have that

$$\Psi^s(t_0) < \frac{\mu}{\mu - \rho} \cdot \kappa \cdot D + \varepsilon,$$

because $t_0 < t_1$, leading to the contradiction

$$\frac{\mu}{\mu - \rho} \cdot \kappa \cdot D + \varepsilon = \Psi^s(t_1) < \left(1 + \frac{\rho}{\mu - \rho}\right) \cdot \kappa \cdot D + \varepsilon.$$

On the other hand, if $\Psi^s(0) > \kappa \cdot D/(1 - \rho^2/\mu^2)$, this is equivalent to

$$\kappa \cdot D + \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) > \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0).$$

Hence,

$$\Psi^s(t_0) < \kappa \cdot D + \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) + \varepsilon$$

and we get that

$$\begin{aligned}&\kappa \cdot D + \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) + \varepsilon\\ &\quad = \Psi^s(t_1)\\ &\quad < \kappa \cdot D + \frac{\rho}{\mu} \cdot \left(\kappa \cdot D + \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \cdot \Psi^s(0) + \varepsilon\right).\end{aligned}$$

This implies the contradiction

$$\Psi^s(0) < \frac{\kappa \cdot D}{1 + \rho/\mu} + \frac{\rho}{\mu} \cdot \Psi^s(0)$$

to $\Psi^s(0) > \kappa \cdot D/(1 - \rho^2/\mu^2)$. □

Corollary 4. *Abbreviate* $q = \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right)$ *and assume that* $q \le \frac{3}{4}$. *For* $i, s \in \mathbf{N}$ *and times* $t \ge 4(\Psi^s(0) + i \cdot \kappa \cdot D)/\mu$, *it holds that*

$$\Psi^s(t) \le \frac{\kappa D}{1 - q} + q^i \left(1 + \frac{\rho}{\mu}\right) \Psi^s(0).$$

Proof. Consider the series given by $x_0 = (1+\rho/\mu)\Psi^s_0$, $x_{i+1} = \kappa \cdot D + q x_i$, $t_0 = 0$, and $t_{i+1} = t_i + \frac{x_i}{\mu}$. By applying Lemma 8 with time 0 replaced by time t_i (i.e., shifting time) and $\Psi^s(0)$ by x_i, we can conclude that x_i upper bounds $\Psi^s(t)$ at times $t \ge t_i$. Simple calculations show that $x_i \le \frac{\kappa D}{1-q} + q^i \Psi^s(0)$ and $t_i \le 4(\Psi^s(0) + i \cdot \kappa \cdot D)/\mu$, so the claim follows. □

In particular, Ψ^s becomes bounded by $(1 + O(\rho/\mu))\kappa D$ within $O(\Psi^s(0)/\mu)$ time. Plugging in $s = 0$, we obtain a bound on the global skew.

Corollary 5. *If* $\frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \le \frac{3}{4}$, *it holds that*

$$\mathcal{G}(t) \le \frac{\kappa D}{1 - q} + q^i \left(1 + \frac{\rho}{\mu}\right) \mathcal{G}(0)$$

at all times $t \ge 4(\mathcal{G}(0) + i \cdot \kappa \cdot D)/\mu$.

Proof. By applying Corollary 4 for $s = 0$, noting that $\mathcal{G}(t) = \Psi^0(t)$. □

C. Bounding the Local Skew

In order to bound the local skew, we analyze the *average* skew over paths in G of various lengths. For long paths of $\Omega(D)$ hops, we will simply exploit that we already bounded the global skew, i.e., the skew between *any* pair of nodes. For successively shorter paths, we inductively show that the average skew between endpoints cannot increase too quickly: reducing the length of a path by factor σ can only increase the skew between endpoints by an additive constant term. Thus, paths of constant length (in particular edges) can only have a(n average) skew that is logarithmic in the network diameter.

In order to bound Ψ^s in terms of Ψ^{s-1}, we need to apply the catch-up lemma in a different form.

Corollary 6. *Let* $s \in \mathbf{Z}$ *and* t_0, t_1 *be times satisfying*

$$t_1 \ge t_0 + \frac{\Psi^{s-1}(t_0)}{\mu}.$$

Then, for any $w \in V$ *we have*

$$L_w(t_1) - L_w(t_0) \ge t_1 - t_0 + \Psi^s_w(t_0).$$

Proof. We have that $\Psi^{s-1}(t_0) \ge \Xi^{s-1}(t_0)$ and there is some $v \in V$ satisfying

$$\Psi^s_w(t_0) = L_v(t_0) - L_w(t_0) - 2s\kappa d(v, w).$$

We apply Lemma 7 to t_0, t_1, v, w and level $s - 1$, yielding that

$$\begin{aligned}&L_w(t_1) - L_w(t_0)\\ &\quad \ge t_1 - t_0 + L_v(t_0) - L_w(t_0) - (2s - 1)\kappa d(v, w)\\ &\quad \ge t_1 - t_0 + L_v(t_0) - L_w(t_0) - 2s\kappa d(v, w)\\ &\quad = t_1 - t_0 + \Psi^s_w(t_0).\end{aligned}$$

□

Combining this corollary with Lemma 5, we can bound Ψ^s at all times.

Lemma 9. *Fix* $s \in \mathbf{Z}$ *and suppose that* $\Psi^{s-1}(t) \le \psi^{s-1}$ *for all times* t. *Then*

$$\Psi^s(t) \le \begin{cases} \Psi^s(0) + \frac{\rho}{\mu} \cdot \psi^{s-1} & \text{if } t \le \frac{\psi^{s-1}}{\mu} \\ \frac{\rho}{\mu} \cdot \psi^{s-1} & \text{else.} \end{cases}$$

Proof. For $t \le \psi^{s-1}/\mu$, the claim follows immediately from Corollary 2. To show the claim for $t > \psi^{s-1}/\mu$, assume for contradiction that it does not hold true and let t_1 be minimal such that there $\Psi^s(t_1) > \rho\psi^{s-1}/\mu + \varepsilon$ for some $\varepsilon > 0$. Thus, there is some $w \in V$ so that

$$\Psi^s_w(t_1) = \Psi^s(t_1) = \frac{\rho}{\mu} \cdot \psi^{s-1} + \varepsilon.$$

Applying Corollary 6 with $t_0 = t_1 - \psi^{s-1}/\mu$ together with Lemma 5 yields the contradiction

$$\begin{aligned}\Psi^s_w(t_1) &\leq \Psi^s_w(t_0) - (L_w(t_1) - L_w(t_0)) + (1+\rho)(t_1 - t_0)\\ &\leq \rho(t_1 - t_0)\\ &= \frac{\rho}{\mu} \cdot \psi^{s-1}.\end{aligned}$$

□

Corollary 7. *Fix $s \in \mathbf{N}$. Suppose that $\Psi^s(t) \leq \psi^s$ for all times t and that $\mathcal{L}(0) \leq 2(s+1)\kappa$. Then*

$$\Psi^{s'}(t) \leq \left(\frac{\rho}{\mu}\right)^{s'-s} \psi^s$$

for all $s' \geq s$ and times t.

Proof. Observe that $\mathcal{L}(0) \leq 2(s+1)\kappa$ implies that $\Psi^{s'}(0) = 0$ for all $s' > s$. Thus, the statement follows from Lemma 9 by induction on s', where $\psi^{s'} = \rho \cdot \psi^{s'-1}/\mu$ and the base case is $s' = s$. □

Corollary 8. *Fix $s \in \mathbf{N}$. Suppose that $\Psi^s(t) \leq \psi^s$ for all times t. Then*

$$\Psi^{s'}(t) \leq \left(\frac{\rho}{\mu}\right)^{s'-s} \psi^s$$

for all $s' \geq s$ and times $t \geq \psi^s/(\mu - \rho)$.

Proof. Consider the times

$$t_{s'} = \sum_{i=1}^{s'-s} \left(\frac{\rho}{\mu}\right)^i \cdot \frac{\psi^s}{\mu} \leq \frac{\psi^s}{\mu} \cdot \frac{1}{1-\rho/\mu} = \frac{\psi^s}{\mu - \rho}.$$

We apply Lemma 9 inductively, where in step $s' > s$ we shift times by $-t_{s'}$. Thus, all considered times fall under the second case of Lemma 9, i.e., the initial values $\Psi^{s'}(0)$ (or rather $\Psi^{s'}(t_{s'})$) do not matter. □

D. Putting Things Together

It remains to combine the results on global and local skew to derive bounds that depend on the system parameters and initialization conditions only. First, we state the bounds on global and local skew that hold at all times. We emphasize that this bound on the local skew also bounds up to which level $s \in \mathbf{N}$ the algorithm needs to check **FT**1 and **FT**2, as larger local skews are impossible.

Theorem 2. *Suppose that $\mathcal{L}(0) \leq (2s+1)\kappa$ for some $s \in \mathbf{N}$. Then*

$$\mathcal{G}(t) \leq \left(2s + \frac{\mu}{\mu - \rho}\right)\kappa D$$
$$\text{and} \qquad \mathcal{L}(t) \leq \left(2s + \left\lceil \log_{\mu/\rho} \frac{\mu D}{\mu - \rho}\right\rceil + 1\right)\kappa$$

for all $t \in \mathbf{R}_{\geq 0}$.

Proof. As $\mathcal{L}(0) \leq (2s+1)\kappa$, we have that

$$\Psi^s(0) \leq \max_{v,w \in V}\{d(v,w)\} \cdot \kappa = \kappa \cdot D.$$

By Lemma 8, hence $\Psi^s(t) \leq \frac{\mu}{\mu-\rho} \cdot \kappa \cdot D$ at all times t. Thus,

$$\begin{aligned}L_v(t) - L_w(t) - 2s\kappa D &\leq L_v(t) - L_w(t) - 2s\kappa d(v,w)\\ &\leq \Psi^s(t)\\ &\leq \frac{\mu}{\mu - \rho} \cdot \kappa \cdot D\end{aligned}$$

for all $v, w \in V$ and times t, implying the stated bound on the global skew.

Concerning the local skew, apply Corollary 7 with $\psi^s = \frac{\mu}{\mu-\rho} \cdot \kappa \cdot D$ and $s' = s + \left\lceil \log_{\mu/\rho} \frac{\mu D}{\mu-\rho}\right\rceil$, yielding that

$$\Psi^{s'}(t) \leq \left(\frac{\rho}{\mu}\right)^{\lceil \log_{\mu/\rho}(\psi^s/\kappa)\rceil} \psi^s \leq \kappa.$$

Hence, for all neighbors $v, w \in V$ and all times t,

$$\begin{aligned}L_v(t) - L_w(t) - 2s'\kappa &= L_v(t) - L_w(t) - 2s'\kappa d(v,w)\\ &\leq \Psi^{s'}(t) \leq \kappa,\end{aligned}$$

implying the claimed bound on the local skew. □

In particular, if the system can be initialized with local skew at most κ, the system maintains the strongest bounds the algorithm guarantees at all times.

Corollary 9. *Suppose that $\mathcal{L}(0) \leq \kappa$. Then*

$$\mathcal{G}(t) \leq \frac{\mu}{\mu - \rho} \cdot \kappa D$$
$$\text{and} \qquad \mathcal{L}(t) \leq \left(\left\lceil \log_{\mu/\rho} \frac{\mu D}{\mu - \rho}\right\rceil + 1\right)\kappa$$

for all $t \in \mathbf{R}_{\geq 0}$.

If such highly accurate intialization is not possible, the algorithm will converge to the bounds from Corollary 9.

Theorem 3. *Suppose that $\mu > 2\rho$. Then there is some $T \in O\left(\frac{\mathcal{G}(0)+\kappa D}{\mu - 2\rho}\right)$ such that*

$$\mathcal{G}(t) \leq \frac{\mu}{\mu - 2\rho} \cdot \kappa D$$
$$\text{and} \qquad \mathcal{L}(t) \leq \left(\left\lceil \log_{\mu/\rho} \frac{\mu D}{\mu - 2\rho}\right\rceil + 1\right)\kappa$$

for all times $t \geq T$.

Proof. By assumption,

$$q = \frac{\rho}{\mu} \cdot \left(1 + \frac{\rho}{\mu}\right) \leq \frac{1}{2} \cdot \frac{3}{2} = \frac{3}{4}.$$

Fix some sufficiently small constant $\varepsilon > 0$ such that

$$\frac{\kappa D}{1-q} + \varepsilon\kappa D \leq \frac{\kappa D}{1 - 2\rho/\mu};$$

since $q \le \frac{3}{2} \cdot \frac{\rho}{\mu}$, such a constant exists. Choose $i \in \mathbf{N}$ minimal with the property that $q^i \left(1 + \frac{\rho}{\mu}\right) \mathcal{G}(0) \le \varepsilon\kappa D$. Therefore, by Corollary 5,

$$\mathcal{G}(t) \le \frac{\mu\kappa D}{\mu - 2\rho}$$

at all times $t \ge 4(\mathcal{G}(0) + i\kappa D)/\mu$. Noting that $\Psi^0(t) = \mathcal{G}(t)$, analogously to Theorem 2 we can now apply Corollary 8 to infer the desired bound on the local skew for times

$$t \ge \frac{4(\mathcal{G}(0) + i\kappa D)}{\mu} + \frac{\mu\kappa D}{(\mu - \rho)(\mu - 2\rho)}.$$

Consequently, it remains to show that the right hand side of this inequality is indeed in $O\left(\frac{\mathcal{G}(0)+\kappa D}{\mu - 2\rho}\right)$. As $\mu - \rho \ge \mu/2$, this is immediate for the second term. Concerning the first term, our choice of i and $q \le 3/4$ yield that $i \in O\left(\log \frac{\mathcal{G}(0)}{\kappa D}\right)$. Because for $x \ge y > 0$ it holds that $x \ge \log(x/y) \cdot y$, we can bound

$$\frac{4(\mathcal{G}(0) + i\kappa D)}{\mu} \in O\left(\frac{\mathcal{G}(0) + \kappa D}{\mu - 2\rho}\right). \qquad \square$$

E. Proof of Lemma 3

Proof. We prove the stronger claim that for all a, b satisfying $t_0 \le a < b \le t_1$, we have

$$\frac{F(b) - F(a)}{b - a} \le r. \tag{9}$$

To this end, suppose to the contrary that there exist $a_0 < b_0$ satisfying $(F(b_0) - F(a_0))/(b_0 - a_0) \ge r + \varepsilon$ for some $\varepsilon > 0$. We define a sequence of nested intervals $[a_0, b_0] \supset [a_1, b_1] \supset \cdots$ as follows. Given $[a_j, b_j]$, let $c_j = (b_j + a_j)/2$ be the midpoint of a_j and b_j. Observe that

$$\begin{aligned}\frac{F(b_j) - F(a_j)}{b_j - a_j} &= \frac{1}{2}\frac{F(b_j) - F(c_j)}{b_j - c_j} + \frac{1}{2}\frac{F(c_j) - F(a_j)}{c_j - a_j} \\ &\ge r + \varepsilon,\end{aligned}$$

so that

$$\frac{F(b_j) - F(c_j)}{b_j - c_j} \ge r + \varepsilon \quad \text{or} \quad \frac{F(c_j) - F(a_j)}{c_j - a_j} \ge r + \varepsilon.$$

If the first inequality holds, define $a_{j+1} = c_j$, $b_{j+1} = b_j$, and otherwise define $a_{j+1} = a_j$, $b_j = c_j$. From the construction of the sequence, it is clear that for all j we have

$$\frac{F(b_j) - F(a_j)}{b_j - a_j} \ge r + \varepsilon. \tag{10}$$

Observe that the sequences $\{a_j\}_{j=0}^{\infty}$ and $\{b_j\}_{j=0}^{\infty}$ are both bounded and monotonic, hence convergent. Further, since $b_j - a_j = \frac{1}{2^j}(b_0 - a_0)$, the two sequences share the same limit.

Define

$$c = \lim_{j\to\infty} a_j = \lim_{j\to\infty} b_j,$$

and let $f \in \mathcal{F}$ be a function satisfying $f(c) = F(c)$. By the hypothesis of the lemma, we have $f'(c) \le r$, so that

$$\lim_{h\to 0} \frac{f(c+h) - f(h)}{h} \le r.$$

Therefore, there exists some $h > 0$ such that for all $t \in [c - h, c + h]$, $t \ne c$, we have

$$\frac{f(t) - f(c)}{t - c} \le r + \frac{1}{2}\varepsilon.$$

Further, from the definition of c, there exists $N \in \mathbf{N}$ such that for all $j \ge N$, we have $a_j, b_j \in [c - h, c + h]$. In particular this implies that for all sufficiently large j, we have

$$\frac{f(c) - f(a_j)}{c - a_j} \le r + \frac{1}{2}\varepsilon, \tag{11}$$

$$\frac{f(b_j) - f(c)}{b_j - c} \le r + \frac{1}{2}\varepsilon. \tag{12}$$

Since $f(a_j) \le F(a_j)$ and $f(c) = F(c)$, (11) implies that for all $j \ge N$,

$$\frac{F(c) - F(a_j)}{c - a_j} \le r + \frac{1}{2}\varepsilon.$$

However, this expression combined with (10) implies that for all $j \ge N$

$$\frac{F(b_j) - F(c)}{b_j - c} \ge r + \varepsilon. \tag{13}$$

Since $F(c) = f(c)$, the previous expression together with (12) implies that for all $j \ge N$ we have $f(b_j) < F(b_j)$.

For each $j \ge N$, let $g_j \in \mathcal{F}$ be a function such that $g_j(b_j) = F(b_j)$. Since $\mathcal{F}$ is finite, there exists some $g \in \mathcal{F}$ such that $g = g_j$ for infinitely many values j. Let $j_0 < j_1 < \cdots$ be the subsequence such that $g = g_{j_k}$ for all $k \in \mathbf{N}$. Then for all j_k, we have $F(b_{j_k}) = g(b_{j_k})$. Further, since F and g are continuous, we have

$$g(c) = \lim_{k\to\infty} g(b_{j_k}) = \lim_{k\to\infty} F(b_{j_k}) = F(c) = f(c).$$

By (13), we therefore have that for all k

$$\frac{g(b_{j_k}) - g(c)}{b_{j_k} - c} = \frac{F(b_j) - F(c)}{b_j - c} \ge r + \varepsilon.$$

However, this final expression contradicts the assumption that $g'(c) \le r$. Therefore, (9) holds, as desired. $\square$