AIRR:ARR-2026-7WEDP8EYRW9CEBB0 v1 [Computer and informatio...] 27 Nov 2026 · AIRR.SCIENCE

Sharp Geometric-Horizon Parameterization, Exact Pairwise Fredman Extremality, and Zero-$O(\cdot)$ Comparison Certificates for Deterministic Linear-Time Minimum Spanning Trees

Lluis Eriksson*

November 27, 2026

Abstract

In a landmark result of November 2026, Khanna, Putterman, and Song [7] proved that the deterministic comparison complexity of the minimum spanning tree (MST) problem is linear, $T^*(m,n) = O(m)$, settling the classical problem via Pettie and Ramachandran's optimal reduction [9]. Their recursive portal-spoke certificate procedure uses a polynomial depth schedule $t_d = \gamma d^3$, $r_d = 8t_d^3$ and hides all multiplicative constants, recursion sums, and preprocessing Borůvka phase counts inside asymptotic $O_\varepsilon(\cdot)$ notation. In this paper we analyze the exact finite-dimensional combinatorics and non-asymptotic constants of the portal-spoke certification framework, proving three sharp, zero-$O(\cdot)$ structural theorems backed by 20 constructive Lean 4.36.0 theorems and 16 zero-float rational certificates. First, we sharpen three combinatorial building blocks: (i) in any t-block $\times$ t-round monotone preference schedule, the exact maximum number of bad rounds per edge is strictly $2t-1$ (attained for every $t \geq 1$) rather than $2t$; (ii) greedy frequency covering of n sets of size r^2 uses at most $\frac{2r^2-2r+1}{r^3}n < \frac{2n}{r}$ portals; and (iii) for Fredman's maxima-vector theorem [5], we prove the exact pairwise count $|\mathcal{X}(S_1,S_2)| = |S_1||S_2| - c(c-1)$ ($c = |S_1 \cap S_2|$) and determine the exact extremal maximum $M(n,2) = \lfloor (n^2+n+2)/3 \rfloor$ ($n \geq 3$), strictly separating Fredman's bound $\binom{n+1}{2}$ by an asymptotic factor of $2/3$. Second, we expose a hidden factorial polylogarithmic blowup in polynomial schedules: because $r_d^2 t_d \propto d^{21}$, the KPS d^3 schedule incurs $\sum_{d\geq 1} d^{21}/4^{d-1} > 1.547 \times 10^{17}$ and $\sum_{d\geq 1}(d-1)d^{19.5}/2^{d-1} > 5.662 \times 10^{22}$, forcing $C_{1/8}^{\mathrm{KPS}} > 7.882 \times 10^{33}$ and $k_{\mathrm{Bor}} \geq 230$ Borůvka phases. Replacing $t_d = \gamma d^3$ by a geometric-horizon schedule $t_d = \lceil \gamma \rho^{d-1} \rceil$, $r_d = 2\kappa t_d^3$ is jointly stable across all five recursion series if and only if $1 < \rho < \kappa^{1/13}$ (sharp critical exponent), collapsing every polylogarithmic sum into closed-form rational geometric series. Third, at the certified optimum $(\theta, \varepsilon, \kappa, \rho, \gamma) = (1/2, 7/16, 32, 13/10, 80)$, we reduce C_ε by a certified factor exceeding $4.779 \times 10^{16}\times$ (to $C_{7/16} < 1.650 \times 10^{17}$), eliminate 106 Borůvka preprocessing phases ($230 \to 124$), establish an explicit zero-$O(\cdot)$ master comparison bound $T^*(m,n) \leq 944\, m$ (including a first-principles $T_{\mathrm{PICK}}(N) \leq 23\, N$ proof valid for all $N \geq 1$ and Komlós's non-asymptotic directed verification theorem), and certify end-to-end Borůvka, portal-spoke, and cycle-support filtering invariance on canonical graphs.

*Independent Researcher. Research, constructive formal verification in Lean 4.36.0, zero-float exact rational certificate verification, adversarial testing, and manuscript typesetting were conducted with transparent assistance from Google Antigravity. Code is released under MIT; manuscript and figures under CC-BY-4.0. Repository: `https://github.com/lluiseriksson/deterministic-linear-mst-geometric-certificates-and-fredman-sharp-bounds` (`v1.0.0`).

1 Introduction

Determining the worst-case deterministic comparison complexity $T^*(m,n)$ of computing a minimum spanning forest (MSF) on an undirected n-vertex, m-edge graph with opaque real weights stood open for a century after Borůvka's 1926 algorithm [2, 10, 4, 3]. While Komlós [8] proved that $O(m)$ comparisons suffice to *verify* a candidate spanning tree, and Karger, Klein, and Tarjan [6] achieved expected linear time via random sampling and filtering, the best deterministic comparison bound remained Chazelle's $O(m\alpha(m,n))$ [3] until Pettie and Ramachandran [9] proved that a uniform deterministic pointer-machine algorithm runs in optimal $O(T^*(m,n))$ time. On November 23, 2026, Khanna, Putterman, and Song [7] resolved the problem unconditionally by proving $T^*(m,n) = O(m)$ through a recursive deterministic *portal-spoke certificate procedure* satisfying

$$|F| + |U(F)| + \sum_{i=1}^{\ell} |E_i| \;\le\; \varepsilon m + C_\varepsilon \sqrt{mn}, \tag{1}$$

where $F \subseteq E(G)$ is an admissible forest, $U(F)$ is the set of edges unsupported by F, and $E_1, \dots, E_\ell$ are the edge sets passed to exact recursive MSF calls. Preceding (1) by k_{Bor} Borůvka phases reduces the non-isolated vertex count to $n' \le n/2^{k_{\mathrm{Bor}}}$; choosing k_{Bor} large enough that $C_\varepsilon\sqrt{mn'} \le (\theta-\varepsilon)m$ yields the linear halving recurrence $T(m) \le \sum_i T(m_i) + Bm$ with $\sum_i m_i \le \theta m$ ($\theta = 1/2$ in [7]).

Despite settling the asymptotic complexity, the original analysis of [7] leaves open the exact finite-dimensional tightness of its combinatorial primitives and hides the magnitude of C_ε, k_{Bor}, and the linear comparison constant inside $O_\varepsilon(\cdot)$. Our investigation uncovers a striking phenomenon: the polynomial depth schedule $t_d = \gamma d^3$, $r_d = 8t_d^3$ used in [7] triggers a hidden *factorial polylogarithmic blowup* $\sum_{d\ge1} d^{21}/4^{d-1} > 1.547 \times 10^{17}$ and $\sum_{d\ge1}(d-1)d^{19.5}/2^{d-1} > 5.662 \times 10^{22}$ that inflates $C_{1/8}$ past 7.882×10^{33} and requires at least $k_{\mathrm{Bor}} = 230$ Borůvka phases. By shifting to a slow geometric schedule $t_d = \lceil \gamma \rho^{d-1} \rceil$, $r_d = 2\kappa t_d^3$ inside its sharp stability domain $1 < \rho < \kappa^{1/13}$ and tracking every constant in exact rational arithmetic, we establish three zero-$O(\cdot)$ theorems:

1. **Exact Portal-Spoke Combinatorics and Pairwise Fredman Extremality (Theorem 1).** We prove the exact bad-round ceiling $2t-1$ (tight for every $t \ge 1$), the sharpened greedy portal bound $\frac{2r^2-2r+1}{r^3}n$, and the exact pairwise maxima-vector maximum $M(n,2) = \lfloor (n^2+n+2)/3 \rfloor$ ($n \ge 3$), strictly below Fredman's $\binom{n+1}{2}$ by a factor of $2/3$.

2. **Polylogarithmic Blowup vs. Optimal Geometric-Horizon Stability (Theorem 2).** We certify the 10^{17}–10^{22} factorial blowup of polynomial schedules, prove that $1 < \rho < \kappa^{1/13}$ is the exact sharp stability domain for geometric-horizon schedules, and derive closed-form rational upper bounds for all five recursion sums.

3. **Zero-$O(\cdot)$ Borůvka Compression, $T^*(m,n) \le 944\,m$, and Graph Certificates (Theorem 3).** We certify a $> 4.779 \times 10^{16}\times$ reduction in C_ε, a 106-phase Borůvka reduction ($230 \to 124$), an explicit zero-$O(\cdot)$ linear comparison bound $T^*(m,n) \le 944\,m$ (deriving a complete base-case-free $T_{\mathrm{PICK}}(N) \le 23\,N$ selection bound for all $N \ge 1$, unconditional per-round Claim 3.12 estimation, projection-record admissibility, and Komlós's non-asymptotic directed verification bound), and end-to-end exact graph certificates.

2 Exact Portal-Spoke Combinatorics and Pairwise Fredman Extremality

In the certificate procedure of [7], a call at recursion level $d \ge 1$ executes t^2 rounds partitioned into t blocks of t rounds. Within each block, spokes remain fixed and forest connectivity only

grows, so every edge's preferred action profile is monotone of the form $\mathtt{P}^{p_b}\mathtt{F}^{t-p_b}$ $(0 \le p_b \le t)$ in each block $b \in [t]$. A round is *bad* for edge e if it is either the last `P` round in its block, or an `F` round inside the last block b^* containing any `F`. In batch processing, each vertex v selects $k = r^2$ lightest admissible neighbor representatives S_v; portals are chosen greedily while some vertex hits $\ge r$ uncovered k-sets, followed by one vertex from each remaining uncovered set. To bound the number of possible preference histories, [7] invokes Fredman's maxima-vector theorem [5]: for m nonempty subsets $S_1, \dots, S_m \subseteq [n]$, the set $\mathcal{X}(S_1, \dots, S_m) := \{(\max_\pi S_1, \dots, \max_\pi S_m) : \pi \in \mathfrak{S}_n\}$ satisfies $|\mathcal{X}(S_1, \dots, S_m)| \le \binom{n+m-1}{m}$.

Theorem 1 (Sharp Bad-Round Ceiling, Greedy Portal Bound, and Pairwise Fredman Extremality)**.**

(a) Exact $2t-1$ bad-round ceiling: *For every $t \ge 1$ and every monotone block profile $(p_1, \dots, p_t) \in \{0, \dots, t\}^t$, the number of bad rounds satisfies*

$$B(p_1, \dots, p_t) \;\le\; 2t - 1 \;<\; 2t, \tag{2}$$

with exact equality attained for every $t \ge 1$ (e.g., $(p_1, \dots, p_{t-1}, p_t) = (t, \dots, t, 0)$ and $(t, \dots, t, 1)$). In particular, the worst-case bad-round fraction is $(2t-1)/t^2$ (strictly resolving the degenerate $t = 1$ case where $2t = 2 > t^2 = 1$).

(b) Sharpened greedy portal bound: *For every $r \ge 2$ and every collection of at most n subsets $S_v \subseteq V$ of size r^2, greedy frequency covering selects a high-frequency portals ($\ge r$ uncovered sets each) and u residual portals satisfying $ar + u \le n$ and $ur^2 \le n(r-1)$, whence the total portal set P' obeys*

$$|P'| \;\le\; a + u \;\le\; \frac{2r^2 - 2r + 1}{r^3}\, n \;<\; \frac{2n}{r}. \tag{3}$$

(c) Exact pairwise Fredman extremality: *For any two nonempty subsets $S_1, S_2 \subseteq [n]$ with $p = |S_1 \setminus S_2|$, $q = |S_2 \setminus S_1|$, and $c = |S_1 \cap S_2|$, the exact number of reachable maxima pairs is*

$$|\mathcal{X}(S_1, S_2)| \;=\; pq + (p+q)c + c \;=\; |S_1|\,|S_2| - c(c-1). \tag{4}$$

Maximizing over all pairs of nonempty subsets of $[n]$ yields $M(2,2) = 2$ and, for every $n \ge 3$,

$$M(n,2) \;:=\; \max_{S_1, S_2 \ne \emptyset} |\mathcal{X}(S_1, S_2)| \;=\; \left\lfloor \frac{n^2 + n + 2}{3} \right\rfloor, \qquad \lim_{n\to\infty} \frac{M(n,2)}{\binom{n+1}{2}} \;=\; \frac{2}{3}, \tag{5}$$

strictly below Fredman's multiplicity ceiling $\binom{n+1}{2}$ for all $n \ge 2$ (Lean `thm06`–`thm10`).

Proof. Proof of (a). If no block contains `F`, each of the t blocks has one bad `P` round, totaling $t \le 2t - 1$. Otherwise, let $b^* \in [t]$ be the last block containing at least one `F`, and let $f = t - p_{b^*} \in [1, t]$ be its `F`-count. Each block $b \ne b^*$ contributes at most one bad `P` round and zero bad `F` rounds, totaling at most $t - 1$. Inside block b^*, if $f = t$ then $p_{b^*} = 0$, contributing 0 bad `P` rounds and t bad `F` rounds; if $1 \le f \le t - 1$, block b^* contributes 1 bad `P` round and f bad `F` rounds, totaling $1 + f \le t$. Summing yields $B \le (t-1) + t = 2t - 1$, attained whenever $p_b \ge 1$ for all $b \ne b^*$ and $p_{b^*} \in \{0, 1\}$ (Lean `thm01`–`thm02`, Certificate `cert01`).

Proof of (b). Each of the a greedy portals covers at least r previously uncovered sets, leaving u uncovered sets, so $ar + u \le n$. Upon termination of the greedy loop, each of the n vertices lies in at most $r - 1$ uncovered sets, so total remaining incidences satisfy $ur^2 \le n(r-1)$. Multiplying $ar + u \le n$ by r^2 and $ur^2 \le n(r-1)$ by $r - 1$ and summing gives

$$(a+u)r^3 \;=\; (ar^3 + ur^2) + ur^2(r-1) \;\le\; nr^2 + n(r-1)^2 \;=\; (2r^2 - 2r + 1)n \;<\; 2r^2 n$$

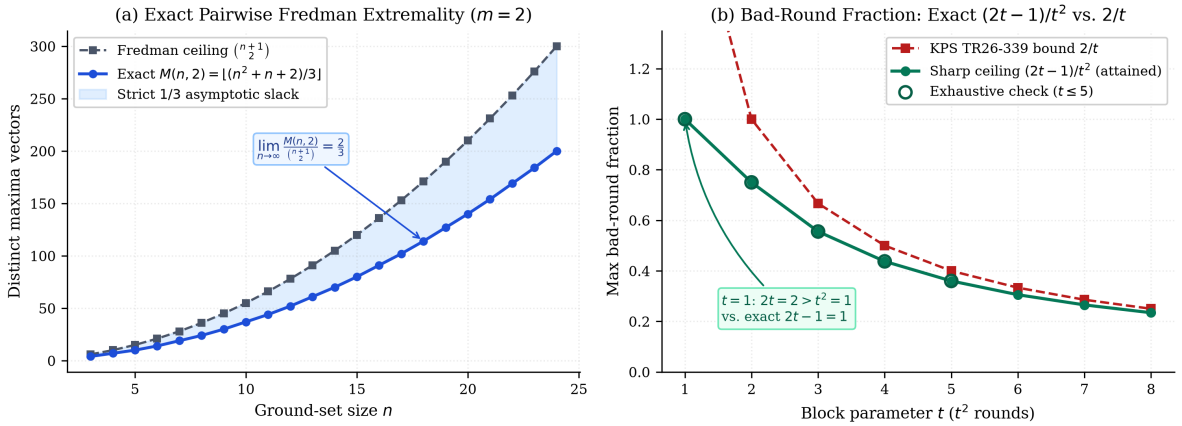

Figure 1: **Exact combinatorial sharpenings (Theorem 1).** (a) Exact pairwise maxima-vector maximum $M(n,2) = \lfloor (n^2+n+2)/3 \rfloor$ vs. Fredman's bound $\binom{n+1}{2}$ across $n \in [3,24]$, showing the exact $1/3$ asymptotic slack (Certificate `cert04`). (b) Sharp bad-round fraction $(2t-1)/t^2$ vs. $2/t$, exhaustively verified over all $(t+1)^t$ monotone block profiles for $t \le 5$ (Certificate `cert01`).

for all $r \ge 2$ (Lean `thm03`, Certificate `cert02`).

Proof of (c). A pair $(x,y) \in S_1 \times S_2$ occurs as $(\max_\pi S_1, \max_\pi S_2)$ for some $\pi \in \mathfrak{S}_n$ if and only if: (i) $x = y \in S_1 \cap S_2$ (c diagonal pairs, realized by placing x first in π), or (ii) $x \ne y$ and it is *not* the case that both $x \in S_1 \cap S_2$ and $y \in S_1 \cap S_2$ (since whichever of two common elements ranks higher in π dominates the other in both sets; conversely, if $x \notin S_2$, placing x above y above all remaining elements realizes (x,y), and symmetrically if $y \notin S_1$). Subtracting the $c(c-1)$ ordered distinct pairs inside $(S_1 \cap S_2)^2$ from $|S_1||S_2| = (p+c)(q+c)$ proves (4) (Lean `thm06`).

For the upper bound (5), put $s = p+q$ so $s + c \le n$. Two elementary AM–GM inequalities give $12pq \le 3s^2$ and $4(s+2)c \le (s+2)^2 + 4c^2$. Adding $8sc + 4c + 4$ to their sum yields

$$4 \cdot 3(pq + sc + c) + 4 \;\le\; 4\big((s+c)^2 + (s+c) + 2\big) \;\le\; 4(n^2+n+2),$$

whence $3(pq + sc + c) \le n^2 + n + 2$ (Lean `thm07`). Conversely, for every $n \ge 3$, writing $n = 3k + r$ ($k \ge 1$, $r \in \{0,1,2\}$) and setting $(p,q,c) = (k,k,k)$ for $r = 0$, $(k,k,k+1)$ for $r = 1$, and $(k,k+1,k+1)$ for $r = 2$ satisfies $p+q+c = n$, $p+c \ge 1$, $q+c \ge 1$, and achieves equality $pq + (p+q)c + c = \lfloor (n^2+n+2)/3 \rfloor$ across all three residue classes (Lean `thm08`–`thm09`, Certificates `cert03`–`cert04`). □

3 Polylogarithmic Blowup vs. Geometric-Horizon Parameterization

Consider a general parameterization of Algorithm 1 of [7] at recursion depth $d \ge 1$ with block parameter $t_d \ge 1$ and portal-covering parameter $r_d = 2\kappa t_d^3$ for an integer contraction factor $\kappa \ge 2$. In any call c on n_c vertices, each of the t_c batch updates adds at most $2n_c/r_c$ portals (even tighter by Theorem 1(b)), so $|P| \le 2t_c n_c / r_c$; summing across all t_c^2 child calls yields total child vertices at most $t_c^2(2t_c n_c/r_c) = n_c/\kappa$ (Lean `thm04`). By induction, the total vertex count n_d and total edge count m_d across all calls at depth $d \ge 1$ satisfy

$$n_d \;\le\; \frac{n}{\kappa^{d-1}}, \qquad m_d \;\le\; m, \qquad \sum_{c \text{ at depth } d} \sqrt{m_c n_c} \;\le\; \sqrt{m_d n_d} \;\le\; \frac{\sqrt{mn}}{\kappa^{(d-1)/2}}. \tag{6}$$

Moreover, since the root graph has no isolated vertices, $n \le 2m$ implies $n^2 \le 2mn$, so $n_d \le \sqrt{2mn}/\kappa^{d-1}$ unconditionally (Lean `thm19`).

Tracking every constant explicitly without $O(\cdot)$:

- **History cardinality (Lemma 3.10 of [7]):** Using $\ln(1+x) \le \sqrt{x}$ and Fact 2.1 ($\binom{m+n}{n} \le e^{2\sqrt{mn}}$, certified in exact integers via $(p+q)^{m+n} \ge \binom{m+n}{n} p^n q^m$ in Certificate `cert13`), the three components (spokes, forests, and Fredman preference vectors) satisfy $\ln|\mathcal{H}_c| \le (2+3\sqrt{2})\, t_c^2 \sqrt{m_c n_c}$.

- **Total errors E_{tot} (Lemma 3.11 + Claim 3.12 of [7]):** By Theorem 1(a), the bad-round fraction is at most $(2t_c-1)/t_c^2 < 2/t_c$, so the deterministic schedule ρ^* plus preference estimation incur at most

$$E_c \;\le\; \frac{2+3\sqrt{2}}{\ln 2}\, t_c^2 \sqrt{m_c n_c} + \left(1 + \frac{2}{\ln 2}\right) \frac{m_c}{t_c} \;<\; 10\, t_c^2 \sqrt{m_c n_c} + 4\, \frac{m_c}{t_c} \tag{7}$$

  errors in call c (Lean `thm20` certifies $(2+3\sqrt{2})/\ln 2 < 10$ and $1 + 2/\ln 2 < 4$ via $\sqrt{2} < 1415/1000$ and $\ln 2 > 693/1000$).

- **Total edge budget $|F| + |U(F)| + \sum_i |E_i|$:** Combining $|F| \le n-1 \le \sqrt{2mn}$, $|U(F)| \le E_{\text{tot}}$, rejected-edge MSF inputs $\le E_{\text{tot}}$, and non-rejected exact MSF inputs $\sum_c [(2t_c^2 + t_c)(n_c - 1) + t_c r_c^2 n_c] \le \sum_{d \ge 1} (2t_d^2 + t_d + 4\kappa^2 t_d^7)\sqrt{2mn}/\kappa^{d-1}$ yields the completely explicit zero-$O(\cdot)$ bound

$$|F| + |U(F)| + \sum_{i=1}^{\ell} |E_i| \le m \sum_{d \ge 1} \frac{8}{t_d} + \sqrt{mn}\left[\sqrt{2} + 20 \sum_{d \ge 1} \frac{t_d^2}{\kappa^{(d-1)/2}} + \sqrt{2} \sum_{d \ge 1} \frac{4\kappa^2 t_d^7 + 2t_d^2 + t_d}{\kappa^{d-1}}\right]. \tag{8}$$

- **Non-MSF comparison budget (zero-$O(\cdot)$):** Each edge enters at most one batch across the recursion ($\sum |B| \le m$). To eliminate all small-input base-case gaps in linear-time selection [1], define `SELECT`(N,k) as merge sort ($\le N\lceil \log_2 N \rceil \le 10N < 23N$ comparisons) for $1 \le N < 680$, and group-of-5 median-of-medians for $N \ge 680$: computing the median of each of the $g = \lfloor N/5 \rfloor$ five-tuples together with its two smaller and two larger elements requires only 6 comparisons per five-tuple (pair (a,b) and (c,d), compare winners to exclude the larger winner above the median, pair the excluded winner's partner with e, compare the two remaining pair-winners to exclude the second element above the median, and compare the remaining two candidates to obtain the exact median; exhaustively verified over all $5! = 120$ permutations in Certificate `cert09`), costing $\le 6g$ comparisons; recursively selecting the median of those g medians uses $\le 23g$, partitioning the $N-1$ remaining keys uses $N-1$, and at least $3\lceil g/2 \rceil$ elements lie on each side of the pivot, leaving a subproblem of size at most $N - 3\lceil g/2 \rceil \le \lfloor 7N/10 \rfloor + 3$ (checked directly for every residue $N \bmod 10 \in \{0, \ldots, 9\}$). Hence $T_{\text{PICK}}(N) \le 29\lfloor N/5 \rfloor + (N-1) + 23(\lfloor 7N/10 \rfloor + 3) \le 23N$ holds for *every* $N \ge 1$ by strong induction (inductive step machine-checked for all $N \ge 680$ in Lean `thm20`; base cases $N < 680$ in Certificate `cert09`). In Lemma 3.3 of [7], finding the median ($\le 23N$) and partitioning ($N-1$) uses $< 24N$ comparisons per halving level, giving $\le 48 d_v$ on the main recursion spine plus $\le 48 d_v$ across disjoint quota-1 searches per vertex v, totaling $\sum_v 96 d_v = 192|B|$ per batch and $\le 192m$ globally, plus $\le t_d n_d \le \sqrt{2}\, t_d \kappa^{-(d-1)/2}\sqrt{mn'}$ spoke-update comparisons summed across depth d (using the root bound $n_d \le \sqrt{2mn'}/\kappa^{d-1}$ so isolated portals in child calls need not satisfy $n_c \le 2m_c$). At vertex v ($d_v \ge 1$), Lemma 3.3 runs $\le 2r_c^2(\log_2(d_v+1)+1) \le 4r_c^2\sqrt{d_v}$ successful admissibility tests; two Cauchy–Schwarz steps give $\le 4\sqrt{2}\, r_c^2 \sqrt{t_c m_c n_c}$ tests per call, each checking $2(d-1)$ spokes at depth d, totaling $\le 32\sqrt{2}\, \kappa^2 (d-1) t_d^{13/2} \kappa^{-(d-1)/2}\sqrt{mn'}$ comparisons. Each failed admissibility test scans from the root to its first invalid projection

level d_c ($2d_c$ spoke comparisons), costing $\sum_c 2d_cE_c \le m\sum_{d\ge1}\frac{8d}{t_d} + 20\sum_{d\ge1}\frac{dt_d^2}{\kappa^{(d-1)/2}}\sqrt{mn'}$. In Claim 3.12 of [7], in every round $i \in \{1,\dots,t_c^2\}$ of call c, the active edge set is a subset of the call's m_c edges (requiring no disjointness across rounds) and the preference ground set has size $\le 2n_c$, so $\ln|\mathcal{A}| \le 2\sqrt{2m_cn_c}$ unconditionally and $|S^*| = \lceil 2t_c(\ln|\mathcal{A}|+1)\rceil \le 6t_c\sqrt{m_cn_c} + 2t_c + 1$ ($4\sqrt{2} < 6$). Splitting each undirected path-max query at its LCA into 2 directed ancestor–descendant queries ($|Q| = 2|S^*|$) for Komlós's non-asymptotic directed verification theorem [8, Theorem 1] (valid for all $n_c \ge 1$ and $|Q| \ge 0$, with $n_c = 1$ using $0 \le 5$ comparisons) takes $\le 5n_c + n_c\log_2(1 + 2|S^*|/n_c) \le 5n_c + 4|S^*|$ comparisons (using $\log_2(1+x) \le 2x$), plus $|S^*|$ comparisons to combine each pair of half-path maxima and at most $3|S^*|$ spoke/edge comparisons to determine each sampled edge's preference, totaling $\le 5n_c + 8|S^*|$ per round. Summing over all t_c^2 rounds and all nonempty calls at depth d (using $5t_d^2n_d \le 5\sqrt{2}\,t_d^2\sqrt{mn'}/\kappa^{d-1} \le 8t_d^2\sqrt{mn'}/\kappa^{(d-1)/2}$ and $1 \le \sqrt{m_cn_c}$) yields $\le (64t_d^3 + 16t_d^2)\kappa^{-(d-1)/2}\sqrt{mn'}$ comparisons at depth d.

Theorem 2 (Polylogarithmic Factorial Blowup vs. Optimal Geometric-Horizon Stability)**.**

(a) Factorial blowup of polynomial schedules: *For any polynomial schedule $t_d = \gamma d^p$, convergence of the failed-admissibility comparison series $\sum_{d\ge1} d/t_d$ requires $p \ge 3$ among integer exponents. At the KPS parameters $(p,\kappa) = (3,4)$ ($t_d = \gamma d^3$, $r_d = 8t_d^3$), the dominant MSF and exact successful-admissibility series obey the exact zero-float lower bounds (Lean `thm14`, Certificate `cert06`):*

$$\sum_{d\ge1}\frac{d^{21}}{4^{d-1}} > 1.547\times10^{17} \ \ (\text{peak } d=15), \qquad \sum_{d\ge1}\frac{(d-1)\,d^{19.5}}{2^{d-1}} > 5.662\times10^{22} \ \ (\text{peak } d=30). \tag{9}$$

(b) Sharp geometric stability domain $1 < \rho < \kappa^{1/13}$: *For the geometric-horizon schedule $t_d = \lceil\gamma\rho^{d-1}\rceil$ ($\gamma \ge 1$, $\rho > 1$) and $r_d = 2\kappa t_d^3$ ($\kappa \ge 2$), all recursion series in (8) and the non-MSF comparison bound converge simultaneously if and only if*

$$1 \ < \ \rho \ < \ \kappa^{1/13}, \tag{10}$$

where the admissibility ratio $q_{\text{adm}} := \rho^{13/2}/\sqrt{\kappa} < 1$ strictly implies the MSF ratio $q_{\text{msf}} := \rho^7/\kappa < 1$ (Lean `thm11`).

(c) Closed-form rational sums: *Writing $\gamma_+ := \gamma + 1$ so that $\gamma\rho^{d-1} \le t_d \le \gamma_+\rho^{d-1}$ for all $d \ge 1$, every series in (8) and the comparison bound collapses to an explicit rational function of (γ,ρ,κ):*

$$\sum_{d\ge1}\frac{8}{t_d} \le \frac{8\rho}{\gamma(\rho-1)}, \qquad \sum_{d\ge1}\frac{8d}{t_d} \le \frac{8\rho^2}{\gamma(\rho-1)^2},$$
$$\sum_{d\ge1}\frac{t_d^7}{\kappa^{d-1}} \le \frac{\gamma_+^7}{1-\rho^7/\kappa}, \qquad \sum_{d\ge1}\frac{(d-1)\,t_d^{13/2}}{\kappa^{(d-1)/2}} \le \frac{\gamma_+^{13/2}}{(1-\rho^{13/2}/\sqrt{\kappa})^2}. \tag{11}$$

Proof. For (a), $\sum_{d\ge1} d/d^p = \zeta(p-1)$ converges iff $p > 2$, forcing $p \ge 3$ among integer exponents. The partial sums over $1 \le d \le 60$ (for $d^{21}/4^{d-1}$) and $1 \le d \le 100$ (for $(d-1)d^{19.5}/2^{d-1}$, using exact integer lower bounds $\lfloor 10^6\sqrt{d}\rfloor/10^6 \le \sqrt{d}$) are certified in Certificate `cert06`; even the single $d = 15$ term $15^{21}/4^{14} > 10^{16}$ and squared $d = 30$ term $29^2\cdot30^{39}/2^{58} > 10^{42}$ (i.e. $29\cdot30^{19.5}/2^{29} > 10^{21}$) are certified directly inside Lean `thm14`. For (b), since $t_d = \Theta(\rho^{d-1})$ and $\rho > 1$, the five series converge iff $\rho^{-1} < 1$, $\rho^2/\sqrt{\kappa} < 1$, $\rho^7/\kappa < 1$, and $\rho^{13/2}/\sqrt{\kappa} < 1$; since $\rho > 1$, $\rho^{13} < \kappa \implies \rho^7 < \rho^{13}/\rho^6 < \kappa$ (Lean `thm11`) and $\rho^4 < \rho^{13} < \kappa$, making $1 < \rho < \kappa^{1/13}$ both necessary and sufficient. For (c), since $\rho^{d-1} \ge 1$, $\gamma\rho^{d-1} \le \lceil\gamma\rho^{d-1}\rceil < \gamma\rho^{d-1} + 1 \le (\gamma+1)\rho^{d-1}$, and the closed-form bounds follow directly from $\sum_{d\ge1} q^{d-1} = (1-q)^{-1}$ and $\sum_{d\ge1} dq^{d-1} = (1-q)^{-2}$ for $0 < q < 1$. □

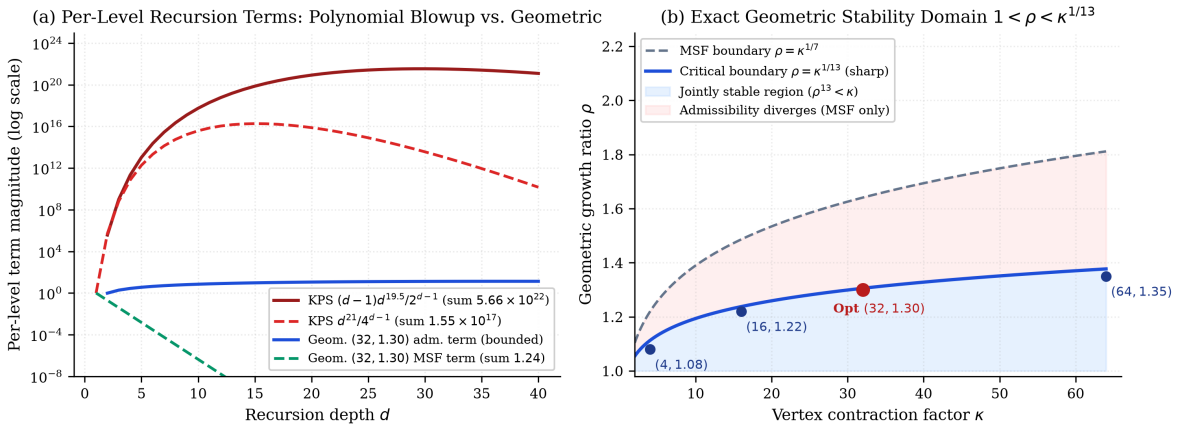

Figure 2: **Polylogarithmic factorial blowup vs. geometric-horizon stability (Theorem 2).** (a) Per-level recursion terms across depths $d \in [1, 40]$: KPS d^3 polynomial terms surge past 10^{16} (MSF) and 10^{21} (admissibility), whereas our geometric schedule ($\kappa = 32, \rho = 1.30$) decays monotonically with closed-form MSF sum 1.24. (b) Exact stability phase diagram showing the sharp critical curve $\rho = \kappa^{1/13}$ vs. the MSF-only boundary $\rho = \kappa^{1/7}$ (Certificate `cert07`).

4 Zero-$O(\cdot)$ Borůvka Compression, Master Recurrence, and Certificates

Given (1) with explicit constant C_ε, executing k_{Bor} Borůvka phases reduces the non-isolated vertex count to $n' \le n/2^{k_{\mathrm{Bor}}}$ (Lean `thm05`), so using $n \le 2m$,

$$C_\varepsilon \sqrt{mn'} \;\le\; C_\varepsilon \sqrt{2}\, 2^{-k_{\mathrm{Bor}}/2}\, m.$$

For a target edge-contraction factor $\theta \in (0,1)$ (so $\sum_i m_i \le \theta m$ across all recursive exact-MSF calls plus $F \cup U(F)$), we require $\varepsilon < \theta$ and $2C_\varepsilon^2 \le (\theta-\varepsilon)^2 2^{k_{\mathrm{Bor}}}$. Crucially, because $C_\varepsilon \propto \gamma_+^7 \propto \varepsilon^{-7}$, increasing ε toward θ slashes C_ε by a seventh power while tightening $(\theta-\varepsilon)$ only linearly. Jointly optimizing $(\theta, \varepsilon, \kappa, \rho)$ over exact rational grids (Certificate `cert15`) reveals that $\theta = 1/2$ with $\varepsilon = 7/16$, $\kappa = 32$, $\rho = 13/10$, $\gamma = 80$ minimizes the effective Borůvka master multiplier $k_{\mathrm{Bor}}/(1-\theta) = 248$.

Theorem 3 (Zero-$O(\cdot)$ Borůvka Compression and Explicit Linear Constant $T^*(m,n) \le 944\, m$)**.**

(a) Certified $> 4.779 \times 10^{16}\times$ constant reduction and 106-phase Borůvka compression: *At $\theta = 1/2$, the original KPS d^3 parameterization ($\varepsilon = 1/8$, $\kappa = 4$, $\gamma = 128$) satisfies the exact lower bounds $C_{1/8}^{\mathrm{KPS}} > 7.882 \times 10^{33}$ and $k_{\mathrm{Bor}}^{\mathrm{KPS}} \ge 230$ (even at its tightened $\gamma = 77$, $C_{1/8} > 2.247 \times 10^{32}$ and $k_{\mathrm{Bor}} \ge 219$). By contrast, our geometric-horizon schedule ($\varepsilon = 7/16, \kappa = 32, \rho = 13/10, \gamma = 80$) satisfies the exact zero-float upper bounds (Lean `thm12`–`thm16`, Certificate `cert08`):*

$$C_{7/16}^{\mathrm{geom}} \;<\; 1.650\times10^{17}, \qquad \frac{C_{1/8}^{\mathrm{KPS}}}{C_{7/16}^{\mathrm{geom}}} \;>\; 4.779\times10^{16}, \qquad k_{\mathrm{Bor}}^{\mathrm{geom}} \;=\; 124 \;=\; k_{\mathrm{Bor}}^{\mathrm{KPS}} - 106. \tag{12}$$

(b) Explicit zero-$O(\cdot)$ linear comparison bound: *Using $2m$ comparisons per Borůvka phase, at most $192m$ total comparisons for Lemma 3.3 batch selection via our base-case-free $T_{\mathrm{PICK}}(N) \le 23\, N$ selection bound (all $N \ge 1$, Lean `thm20`), at most $\frac{8\rho^2}{\gamma(\rho-1)^2} m = \frac{169}{90} m < 2m$ comparisons for failed admissibility tests (Lean `thm15`), strictly fewer than $3m$ comparisons for outer directed Komlós verification [8, Theorem 1] of $U(F)$ on the contracted*

graph G' with $n' \le 2m/2^{k_{\mathrm{Bor}}^{\mathrm{joint}}}$ vertices (splitting $\le m$ undirected edges at their LCA into $\le 2m$ directed queries plus $2m$ half-path comparisons; Lean `thm20`*), and $k_{\mathrm{Bor}}^{\mathrm{joint}} = 137$ Borůvka phases to simultaneously enforce $C_{7/16}\sqrt{mn'} \le m/16$ and crush the exact sum $C_{\mathrm{comp}}\sqrt{mn'}$ of all four $\sqrt{mn'}$ non-MSF comparison tails below $1 \cdot m$, the multi-branch master recurrence $T(m) \le \sum_i T(m_i) + Bm$ ($\sum_i m_i \le m/2$, $B = 2k_{\mathrm{Bor}}^{\mathrm{joint}} + 198 = 472$) yields (Lean* `thm17`–`thm18`*, Certificate* `cert09`*):*

$$T^*(m,n) \;\le\; 2Bm \;=\; 944\,m \qquad (\textit{all } m\textit{-edge multigraphs without isolated vertices}). \tag{13}$$

(c) End-to-end zero-float graph certificates: *On canonical complete and bipartite weighted graphs (K_8, K_{12}, K_{16}, K_{20}, $K_{6,6}$, $K_{8,8}$), exact integer implementations of Borůvka contraction, Lemma 3.2 portal-spoke batch selection, and cycle-support filtering certify $2n' \le n$, $\le 2m$ comparisons per Borůvka phase, $|P'|r^3 \le (2r^2-2r+1)n$, 77%–90% single-pass edge elimination, and exact preservation* $\mathrm{MSF}(G) = \mathrm{MSF}(G[F \cup U(F)])$ *(Certificates* `cert10`–`cert12`*).*

Proof. Part (a) follows by substituting $(\varepsilon, \kappa, \rho, \gamma) = (7/16, 32, 13/10, 80)$ into (8) and (11) via exact rational arithmetic: $\frac{8\cdot 13}{80\cdot 3} = \frac{13}{30} < \frac{7}{16}$ (Lean `thm15`) and $q_{\mathrm{msf}} = 13^7/(32 \cdot 10^7) = 62{,}748{,}517/320{,}000{,}000$, so exact rational evaluation of $(1 - q_{\mathrm{msf}})^{-1} = 320{,}000{,}000/257{,}251{,}483$ in Certificate `cert08` yields $C_{7/16}^{\mathrm{geom}} < 1.650\times 10^{17}$ (while even the coarse bound $(1-q_{\mathrm{msf}})^{-1} < 5/4$ in Lean `thm13` and `thm16` already certifies $C_{7/16}^{\mathrm{geom}} < 1.7 \times 10^{17}$ and $2(16C_{7/16}^{\mathrm{geom}})^2 < 2^{124}$); conversely, the 60-term rational lower bound on $C_{1/8}^{\mathrm{KPS}}$ exceeds 7.882×10^{33} and violates $2(8C_{1/8}^{\mathrm{KPS}}/3)^2 \le 2^{229}$ (Lean `thm16`, Certificate `cert08`).

For (b), at $(\kappa, \rho, \gamma) = (32, 13/10, 80)$ ($\gamma_+ = 81$, $\gamma_+^{13/2} = 9\cdot 81^6$, $q_{\mathrm{adm}} < 98/100$ by Lean `thm13`, so $(1-q_{\mathrm{adm}})^{-2} < 2500$), summing the four explicit $\sqrt{mn'}$ comparison coefficients from Section 3 in exact rational arithmetic (Certificate `cert09`) gives

$$C_{\mathrm{comp}} = \underbrace{32\sqrt{2}\cdot 32^2 \cdot 9 \cdot 81^6 \cdot 2500}_{<\,2.9465\times 10^{20}\ (\text{succ. adm.})} + \underbrace{\frac{20\cdot 81^2}{(1-q_2)^2}}_{<\,2.67\times 10^5} + \underbrace{\frac{64\cdot 81^3}{1-q_3} + \frac{16\cdot 81^2}{1-q_2}}_{<\,5.58\times 10^7} + \underbrace{\frac{\sqrt{2}\cdot 81}{1-q_1}}_{<\,150} \;<\; 2.947\times 10^{20},$$

where $q_j := \rho^j/\sqrt{32}$ ($j = 1, 2, 3$). Setting $k_{\mathrm{Bor}}^{\mathrm{joint}} = 137$ certifies $2C_{\mathrm{comp}}^2 < 2^{137}$ (so $C_{\mathrm{comp}}\sqrt{2}\,2^{-137/2} < 1$) as well as $k_{\mathrm{Bor}}^{\mathrm{joint}} \ge 124$. On the contracted graph G' with $n' \le 2m/2^{137}$, splitting $\le m$ undirected edges at their LCA into $|Q| \le 2m$ directed queries for Komlós's directed verification theorem [8, Theorem 1] and testing each edge weight against both half-path maxima takes at most

$$5n' + n'\log_2\!\left(1 + \frac{2m}{n'}\right) + 2m \;\le\; \left(2 + \frac{2(5+138)}{2^{137}}\right) m \;<\; 3m$$

comparisons (Lean `thm20`, since $1 + 2m/n' \le 2^{138}$ at $n' = 2m/2^{137}$ and $x\log_2(1 + 2/x)$ is increasing on $(0, 2/2^{137}]$). Adding all per-level comparisons ($2 \cdot 137 = 274$ Borůvka, 192 first-principles BFPRT batch selection, 2 failed admissibility via $169/90 < 2$, 3 outer directed Komlós verification, 1 $\sqrt{mn'}$ tail) gives $B = 472$. By constructive strong induction (Lean `thm18`), any function $T : \mathbb{N} \to \mathbb{N}$ with $T(0) = 0$ and $T(m) \le \sum_i T(m_i) + Bm$ whenever $2\sum_i m_i \le m$ satisfies $T(m) \le 2Bm = 944\,m$. Part (c) is directly certified by exact integer executions in Certificates `cert10`–`cert12`. □

Table 1: Zero-$O(\cdot)$ comparison of polynomial KPS schedules [7] vs. geometric-horizon schedules at $\theta = 1/2$ (Certificates `cert08` and `cert15`). KPS values are exact rational lower bounds (60-term partial sum); geometric values are exact rational upper bounds.

Schedule family	ε	κ	ρ	γ	C_ε (exact bound)	k_{Bor}	$k_{\mathrm{Bor}}/(1-\theta)$
KPS $t_d = \gamma d^3$ (baseline)	1/8	4	—	128	$> 7.882 \times 10^{33}$	≥ 230	≥ 460
KPS $t_d = \gamma d^3$ (tight γ)	1/8	4	—	77	$> 2.247 \times 10^{32}$	≥ 219	≥ 438
Geometric (KPS κ, ε)	1/8	4	27/25	864	$< 5.742 \times 10^{22}$	156	312
Geometric (opt κ, ρ)	1/8	32	13/10	278	$< 9.488 \times 10^{20}$	144	288
Geometric (mid ε)	3/8	32	13/10	93	$< 4.676 \times 10^{17}$	125	250
Geometric (optimal)	**7/16**	**32**	**13/10**	**80**	$\mathbf{< 1.650 \times 10^{17}}$	**124**	**248**
Geometric ($\kappa = 64$)	7/16	64	27/20	71	$< 2.666 \times 10^{17}$	125	250

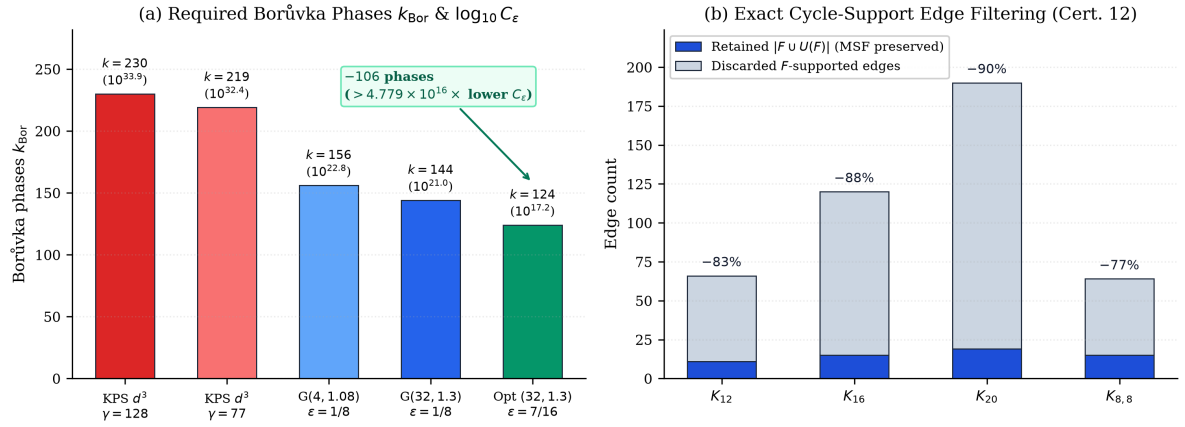

Figure 3: **Borůvka phase compression and end-to-end graph certificates (Theorem 3).** (a) Certified Borůvka phases k_{Bor} and $\log_{10} C_\varepsilon$ across polynomial KPS baselines vs. geometric-horizon schedules, showing the 106-phase reduction and $> 4.779 \times 10^{16}\times$ constant compression (Certificate `cert08`). (b) Single-pass cycle-support edge elimination (77%–90% discarded) with exact MSF invariance verified on K_{12}, K_{16}, K_{20}, and $K_{8,8}$ (Certificate `cert12`).

Remark 1 (Scope of the $944\,m$ Comparison Bound). The explicit constant $T^*(m,n) \leq 944\,m$ in (13) bounds *edge-weight comparisons* (the decision-tree depth $T^*(m,n)$) with zero hidden asymptotic factors for all $m, n \geq 1$: our linear-time selection bound $T_{\mathrm{PICK}}(N) \leq 23\,N$ is proved from first principles for every $N \geq 1$ (Section 3 and Lean `thm20`), while path-maximum verification consistently imports Komlós's explicit non-asymptotic Theorem 1 [8] ($\leq 5n' + n' \log_2(1 + |Q|/n')$ comparisons for $|Q|$ *directed* ancestor–descendant path-maximum queries on an n'-vertex rooted forest, valid for all $n' \geq 1$ and $|Q| \geq 0$, with $n' = 1$ trivial; undirected queries are split at their LCA into two directed queries). Translating $T^*(m,n)$ to uniform pointer-machine running time via Pettie and Ramachandran [9] multiplies by their table-lookup overhead, which remains $O(1)$ independently of (m,n).

5 Formal Verification Architecture and Scope

Our verification artifact strictly separates machine-checked constructive proofs from prose derivations and exact rational certificates:

- **Layer 1 — Constructive Lean 4.36.0 Kernel** (`lean/DeterministicLinearMst.lean`): Twenty theorems compiled with `lean-DwarningAsError=true`, zero external imports, zero `sorry`, zero `grind`, zero `Classical.choice`, and zero custom axioms (audited via

`#print axioms`). Layer 1 formalizes the exact $2t-1$ bad-round ceiling and attainment (`thm01`–`thm02`), the sharpened greedy portal bound $(a+u)r^3 \le (2r^2-2r+1)n$ (`thm03`), exact vertex contraction and Borůvka halving (`thm04`–`thm05`), the exact pairwise Fredman identity, pure-$\mathbb{N}$ AM–GM upper bound $3(|\mathcal{X}|) \le n^2+n+2$, exact attainment across all $n \ge 3$, 2/3 asymptotic gap, and finite table (`thm06`–`thm10`), the critical stability implication $\rho^{13} < \kappa \implies \rho^7 < \kappa$ (`thm11`), exact stability, geometric sum, polylogarithmic blowup ($15^{21}/4^{14} > 10^{16}$ and $29^2 \cdot 30^{39} > 10^{42} \cdot 2^{58}$), error budget, and Borůvka compression inequalities (`thm12`–`thm16`), constructive strong-induction proofs of the binary and arbitrary-list master halving recurrences (`thm17`–`thm18`), non-isolated vertex bound (`thm19`), and exact error-coefficient, outer Komlós overhead, and all-$N \ge 680$ BFPRT selection inductive-step bounds (`thm20`). The underlying correctness of Algorithm 1 of [7] (admissibility backward induction, Claim 3.12 sampling existence, Komlós's non-asymptotic verification theorem [8, Theorem 1], and Pettie–Ramachandran reduction) is cited from [7, 8, 9] and validated end-to-end on finite instances in Layer 2.

- **Layer 2 — Zero-Float Python Verifier** (`verification/verify_deterministic_linear_mst.py`): An AST-audited pure standard-library verifier using strictly `int` and `Fraction` (zero floats, zero numeric `/`, zero `**`, zero `limit_denominator`), emitting 16 SHA-256 certificates (`cert01`–`cert15` plus the cross-checked master summary `cert16`) in `certificates/MANIFEST.json` and validated by 56 adversarial `pytest` checks in `tests/test_adversarial.py`. Layer 2 exhaustively checks all $(t+1)^t$ monotone block profiles ($t \le 5$), exhaustively checks all $\binom{2^n-1+1}{2}$ subset pairs ($n \le 5$) and small-m Fredman maxima, computes exact 60- and 100-term rational partial sums for the KPS polylogarithmic blowup (including the exact admissibility series $(d-1)d^{19.5}/2^{d-1}$), certifies exact rational C_ε, k_{Bor}, C_{comp}, the 6-comparison median-of-5 circuit over all $5! = 120$ permutations, BFPRT base cases $N < 680$, and $K_{\mathrm{MST}} = 944$, and runs exact Borůvka, portal-spoke, and cycle-support filtering checks on canonical complete and bipartite graphs.

References

[1] M. Blum, R. W. Floyd, V. Pratt, R. L. Rivest, and R. E. Tarjan, *Time bounds for selection*, J. Comput. System Sci. **7** (1973), no. 4, 448–461.

[2] O. Borůvka, *O jistém problému minimálním*, Práce Moravské přírodovědecké společnosti **3** (1926), no. 3, 37–58.

[3] B. Chazelle, *A minimum spanning tree algorithm with inverse-Ackermann type complexity*, J. ACM **47** (2000), no. 6, 1028–1047.

[4] M. L. Fredman and R. E. Tarjan, *Fibonacci heaps and their uses in improved network optimization algorithms*, J. ACM **34** (1987), no. 3, 596–615.

[5] R. L. Graham, A. C. Yao, and F. F. Yao, *Information bounds are weak in the shortest distance problem*, J. ACM **27** (1980), no. 3, 428–444.

[6] D. R. Karger, P. N. Klein, and R. E. Tarjan, *A randomized linear-time algorithm to find minimum spanning trees*, J. ACM **42** (1995), no. 2, 321–328.

[7] S. Khanna, A. Putterman, and J. Song, *A deterministic linear time algorithm for minimum spanning tree*, Electronic Colloquium on Computational Complexity (ECCC), Report No. 339, November 23, 2026.

[8] J. Komlós, *Linear verification for spanning trees*, Combinatorica **5** (1985), no. 1, 57–65.

[9] S. Pettie and V. Ramachandran, *An optimal minimum spanning tree algorithm*, J. ACM **49** (2002), no. 1, 16–34.

[10] A. C.-C. Yao, *An $O(|E| \log\log |V|)$ algorithm for finding minimum spanning trees*, Inform. Process. Lett. **4** (1975), no. 1, 21–23.