THE RATIONAL DECISION STRUCTURE OF BARANWAL'S CERTIFIED DEPTH-ERROR INSTANCE

KOREA SUPERINTELLIGENCE LABS

Abstract. Baranwal (arXiv:2607.16676v2) certifies, for the pairwise message-passing classifier on the broadcast-labelled Poisson Galton–Watson tree with a two-point feature law at $(\Delta, \gamma, t_0) = (3, 0.55, 0.95)$, the first four values of the depth-error curve, $\mathcal{E}(0) = 0.025$, $\mathcal{E}(1) = 0.02261938\ldots$, $\mathcal{E}(2) \in [0.03870641, 0.03870642]$ and $\mathcal{E}(3) \in [0.07589622, 0.07589623]$, by a never-renormalised lattice computation in which lattice points within 10^{-8} of a tie are counted as errors. We observe that at this instance every message is the logarithm of a rational number, $2\,\mathrm{artanh}(\gamma^k t_0) = \log r_k$ with $r_k = (1+\gamma^k t_0)/(1-\gamma^k t_0)$, so that the sign of the depth-ℓ statistic is an exact comparison of two integers. We derive the seven reduced ratios $r_0, \ldots, r_6$, exhibit a distinguishing-prime certificate (29, 3433, 23, 59629, 41, 317) for their multiplicative independence—so the statistic never vanishes at depths $\ell \le 6$ and the tie convention is never active there—verify exhaustively that no tie occurs at depth ≤ 2 on the box $m_0 = \pm 1$, $|m_1| \le 30$, $|m_2| \le 70$ (17,202 exact comparisons of integers of up to 367 digits), and identify the exact depth-one decision thresholds $m_1 \le -4$ (root sign $+1$) and $m_1 \le 3$ (root sign -1), the general half by induction, which express $\mathcal{E}(1)$ as a two-term Skellam sum. The two-point law has atoms, so at this instance the source's tie clause is not a null event a priori; the certificate shows it is vacuous. The ratios, the certificate, the exhaustive check, the threshold inequalities and five controls are machine-checked in Lean 4 with no axiom beyond `propext` and `Quot.sound`; the step from the certificate to independence is a one-line valuation argument. As uncertified evidence we report $\mathcal{E}(4) \in [0.10714, 0.10719]$ and $\mathcal{E}(5) \in [0.13364, 0.13375]$ from a scalar sub-probability lattice in floating point, and price a certified version.

1. Introduction

1.1. **The objects.** The sparse contextual stochastic block model with average degree $\Delta = O(1)$ has as local weak limit a broadcast-labelled Poisson Galton–Watson tree: the root o carries a label $y_o \in \{\pm 1\}$, every vertex has $\mathrm{Poi}(\Delta)$ children, each child keeps its parent's label with probability $(1+\gamma)/2$ and flips it with probability $(1-\gamma)/2$, and each vertex v carries a feature X_v whose law $\mathbb{P}_\pm$ depends only on the label of v, the features being conditionally independent given the labels. Baranwal [1] studies, on this limit tree, the *pairwise message-passing classifier* of depth ℓ, which is the rule derived in [2] (Theorem 1 and Corollary 1.1 of the updated arXiv version, as [1] specifies; see also [3, Ch. 5]). Writing N_k for the set of vertices at distance k from the root and

$$t(x) = \frac{\rho_+(x) - \rho_-(x)}{\rho_+(x) + \rho_-(x)} \in [-1, 1] \tag{1}$$

for the bounded likelihood-ratio transform of the feature (equation (3) of [1], $\rho_\pm$ being the densities of $\mathbb{P}_\pm$), the rule and its error are, in the source's own notation (equations (4) and (5) of [1], quoted verbatim),

$$h_\ell = \mathrm{sgn}(T_\ell), \qquad T_\ell = \sum_{k=0}^{\ell} \sum_{v \in N_k} M_k(X_v), \qquad M_k(x) = 2\,\mathrm{artanh}\big(\gamma^k\, t(x)\big), \tag{2}$$

$$\mathcal{E}(\ell) = \mathbb{P}\big(y_o\, T_\ell \le 0\big), \tag{3}$$

"counting ties conservatively as errors (under Assumption 3 ties are null events)". The single number governing the value of depth is the Kesten–Stigum ratio $\kappa = \gamma^2 \Delta$. Below the threshold ($\kappa < 1$) the

source proves (Theorem 3.1 of [1], under its Assumption 3, a bounded density for $\log\psi(X)$ where $\psi = \rho_+/\rho_-$) that the error sequence is Cauchy at a geometric rate, $|\mathcal{E}(\ell) - \mathcal{E}(\ell')| \le C\kappa^{(\ell+1)/3}$ for all $\ell' > \ell$, so that $\mathcal{E}(\infty) = \lim_\ell \mathcal{E}(\ell)$ exists; above the threshold it proves a depth amplification bound (Theorem 3.3). The pairwise rule adds messages from vertices sharing ancestry as if they were independent, and its error curve is not monotone in ℓ (Observation 3.9 of [1]). The source makes this exact in its Appendix C, Proposition C.1, which we quote verbatim:

> Consider the two-point feature law $t(X) \in \{\pm t_0\}$ with $\mathbb{P}_+\big(t(X) = t_0\big) = \frac{1+t_0}{2}$ (which satisfies Assumptions 1–2), at $(\Delta, \gamma, t_0) = (3,\, 0.55,\, 0.95)$, so that $\kappa = 0.9075 < 1$. Then
>
> $$\mathcal{E}(0) = 0.025, \qquad \mathcal{E}(1) = 0.02261938\ldots,$$
>
> $$\mathcal{E}(2) \in [0.03870641,\, 0.03870642], \qquad \mathcal{E}(3) \in [0.07589622,\, 0.07589623],$$
>
> and in particular
>
> $$\mathcal{E}(3) - \mathcal{E}(2) \ge 0.0371898 > 0 \qquad \text{and} \qquad \mathcal{E}(2) - \mathcal{E}(1) \ge 0.0160 > 0\,:$$
>
> the depth-error curve of the pairwise rule strictly increases from depth 1 to depth 3. The proof is the finite, certified computation described below; the two-point law has atoms, so this instance also illustrates the discrete-feature regime outside Assumption 3.

The proof in [1] is a finite computation on the lattice law of the signed level counts (recalled in Section 2), carried out under the following accounting rule, again quoted verbatim:

> The computation evaluates these finite convolutions on truncated integer boxes with a strict accounting rule: every distribution is kept as a *sub*-probability and never renormalized, so each array's mass deficit is exactly the truncation error it carries; the Poisson weights beyond the group-size caps ($i \le 16$, $j \le 9$, level-2 groups $\le 20/\le 12$) are added to the error ledger as-is; lattice points with $|T_\ell| < 10^{-8}$ are counted as errors (the tie-conservative side); and a conservative floating-point budget covers rounding. The resulting certified brackets are those displayed in Proposition C.1, with total ledger below 5×10^{-9}—six orders of magnitude below the certified gap.

The tie clause is not idle here. The source's remark after (3) that ties are null events is made under its Assumption 3, and the last sentence of Proposition C.1 records that the two-point law lies outside that assumption: at the instance nothing in the source excludes $T_\ell = 0$, and its computation pays for that with the 10^{-8} guard.

1.2. **What this paper adds.** Nothing in the source's treatment of the instance uses that $\gamma = 11/20$ and $t_0 = 19/20$ are rational. Since $2\operatorname{artanh} x = \log\frac{1+x}{1-x}$, every message at the instance is $\pm\log r_k$ for the *rational* number

$$r_k = \frac{1+\gamma^k t_0}{1-\gamma^k t_0} \in \mathbb{Q}_{>0},$$

so the decision $T_\ell \le 0$ is the exact rational comparison $\prod_k r_k^{m_k} \le 1$, m_k being the signed count of level-k feature signs, and $T_\ell = 0$ is the exact equation $\prod_k r_k^{m_k} = 1$. We work out this structure and its consequences for the source's computation.

(1) The seven reduced ratios $r_0, \ldots, r_6$ (Proposition 3.1), derived from γ and t_0 rather than asserted.
(2) A distinguishing-prime certificate for the multiplicative independence of $r_0, \ldots, r_6$ (Theorem 4.1): the primes $29, 3433, 23, 59629, 41, 317$ each divide exactly one of the fourteen numerators and denominators, to the first power, and that one belongs to $r_1, \ldots, r_6$ respectively. Consequently (Corollary 4.3) $T_\ell \ne 0$ for every realisation of the tree at every depth $\ell \le 6$, so at those depths the tie clause of (3) is never active and $\mathcal{E}(\ell) = \mathbb{P}(y_o T_\ell < 0)$.

(3) An exhaustive verification, independent of the valuation argument behind Corollary 4.3, that no tie occurs at depth ≤ 2 on the exponent box $m_0 = \pm 1$, $|m_1| \le 30$, $|m_2| \le 70$ (Theorem 5.1): 17,202 exact comparisons of integers of up to 367 decimal digits. We relate the box to the truncation boxes of the source's computation in Section 5.
(4) The exact depth-one thresholds (Proposition 6.1): with root sign $+1$ the rule errs if and only if $m_1 \le -4$, with root sign -1 if and only if $m_1 \le 3$. The half of this that concerns all large exponents is an induction, not a finite check. The thresholds turn $\mathcal{E}(1)$ into a two-term Skellam sum (Corollary 6.2), which reproduces the source's $0.02261938\ldots$.
(5) Five controls (Proposition 7.1) that refute the neighbouring wrong thresholds and wrong primes.

Items (1)–(5) are machine-checked in Lean 4 (Section 9). Beyond depth 3 the source computes nothing at this instance; in Section 8 we report, as evidence and not as a theorem, the brackets $\mathcal{E}(4) \in [0.10714, 0.10719]$ and $\mathcal{E}(5) \in [0.13364, 0.13375]$ obtained by a scalar never-renormalising lattice scheme in floating point, agreeing with a Monte Carlo run, and we price a certified version.

1.3. **Scope.** The source leaves open, for $\kappa > 1$, the true rate at which $\mathcal{E}(\ell)$ approaches its large-ℓ behaviour and even whether $\lim_{\ell\to\infty} \mathcal{E}(\ell)$ exists at all (Section 6 of [1]: "the existence of $\lim_{\ell\to\infty} \mathcal{E}(\ell)$ itself, which our Cauchy argument establishes only below the threshold"). The certified instance has $\kappa = 0.9075 < 1$. There the source's Theorem 3.1 gives the limit under Assumption 3, which the two-point law does not satisfy (the source says so in Proposition C.1 and, after its Assumption 3, that the assumption "excludes atomic features (e.g. discrete node attributes) from Theorem 3.1 as stated"). For such laws the source records in its Appendix A the form of the bound that its proof yields without Assumption 3,

$$\big|\mathcal{E}(\ell) - \mathcal{E}(\ell')\big| \le \inf_{s>0} \Big\{ \mathcal{Q}_{T_{\ell'}}(s) + \frac{C_1 \kappa^{\ell+1}}{s^2} \Big\}, \qquad \mathcal{Q}_X(w) := \sup_{a\in\mathbb{R}} \mathbb{P}\big(X \in (a, a+w]\big),$$

and calls it "the correct statement of Theorem 3.1" for atomic laws. At the instance that form yields no limit: the root has no children with probability e^{-3}, and on that event $T_{\ell'} = c_0 s_o$ for every ℓ', so $T_{\ell'}$ carries an atom of mass at least $e^{-3}/2$ at c_0 or at $-c_0$, whence $\mathcal{Q}_{T_{\ell'}}(s) \ge e^{-3}/2 > 0.024$ for all $s > 0$ and all ℓ', and the right-hand side of the display never falls below that. No statement of the source—Theorem 3.1, the informal restatement of its conclusion in the contributions paragraph of Section 1.1 (which refers to Theorem 3.1 for the statement), Appendix A, or Proposition C.1—asserts that $\lim_{\ell} \mathcal{E}(\ell)$ exists at the instance, and neither do we. Nothing in this paper bears on the existence or the value of that limit; what is uncomputed at the instance is $\mathcal{E}(\ell)$ for $\ell \ge 4$, and Section 8 is evidence about the next terms only.

1.4. **Numbering of the source.** Theorem, proposition and equation numbers of [1] are those of the arXiv v2 PDF (27 pages), which we read and cross-checked against the auxiliary file of a fresh compilation of the v2 e-print (`main.tex`) with `tectonic`; both give Appendix C, Proposition C.1, equations (3)–(5) in Section 2.2, Theorem 3.1, Theorem 3.3, Observation 3.9, Assumptions 1–3 and Appendix A. The founding record of this work cited the same environments by their LaTeX labels (`prop:nonmonocert`, `app:certified`, `eq:classifier`, `eq:error`, `eq:t-def`, `thm:below`, `obs:nonmono`, `ass:density`) because it had not compiled the source; we give the printed numbers here. The v1 and v2 e-prints differ only in the preamble (theorem-environment counters and the author line); their bodies are identical, so the numbering above is also that of v1. The arXiv comments field announces code as ancillary files: the v1 e-print carries an `anc/` directory (`README.txt`, `experiments.py`, `kappa_bp.py`, `nonmono_certify.py` and saved results), which arXiv lists for v1, while the v2 e-print carries no ancillary file although its Appendix C still names `code/nonmono_certify.py`. We read `nonmono_certify.py` from the v1 files; its parameters are those the text prints (group-size caps 16/9 and 20/12, tie window 10^{-8}, floating-point budget 5×10^{-10}), and its truncation boxes are quoted in Remark 5.2.

2. The instance

Throughout, $(\Delta, \gamma, t_0) = (3, \frac{11}{20}, \frac{19}{20})$ and the feature law is the two-point law of Proposition C.1. Following the source's Appendix C we condition on $y_o = +1$; by the symmetry of the model under simultaneous negation of all labels and all features (Assumption 1 of [1], which the two-point law satisfies) this loses nothing, and $\mathcal{E}(\ell) = \mathbb{P}(T_\ell \le 0 \mid y_o = +1)$. Write $\sigma_v \in \{\pm 1\}$ for the label of v relative to the root and $s_v \in \{\pm 1\}$ for the sign of $t(X_v) = t_0 s_v$. The source records that

$$\mathbb{P}(s_v = +1 \mid \sigma_v) = \frac{1 + t_0 \sigma_v}{2},$$

that the message of a vertex at distance k is $c_k s_v$ with $c_k = 2\,\mathrm{artanh}(\gamma^k t_0)$, and hence that

$$T_\ell = \sum_{k=0}^{\ell} c_k\, m_k, \qquad m_k = \sum_{v \in N_k} s_v \in \mathbb{Z}, \qquad m_0 = s_o, \tag{4}$$

so that $\mathcal{E}(\ell)$ is determined by the lattice law of $(s_o, m_1, \ldots, m_\ell)$. The level-$k$ vertices with $s = +1$ and with $s = -1$ have additive counts $m_k^{\pm}$ and $m_k = m_k^{+} - m_k^{-}$.

Lemma 2.1. *For $x \in (-1, 1)$, $2\,\mathrm{artanh}\, x = \log \dfrac{1+x}{1-x}$. Consequently, at the instance,*

$$c_k = \log r_k, \qquad r_k := \frac{1 + \gamma^k t_0}{1 - \gamma^k t_0} = \frac{20^{k+1} + 19 \cdot 11^k}{20^{k+1} - 19 \cdot 11^k} \in \mathbb{Q}_{>0},$$

and, writing $r_k = a_k / b_k$ in lowest terms,

$$T_\ell \le 0 \iff \prod_{k=0}^{\ell} r_k^{m_k} \le 1 \iff \prod_{m_k \ge 0} a_k^{m_k} \prod_{m_k < 0} b_k^{-m_k} \;\le\; \prod_{m_k \ge 0} b_k^{m_k} \prod_{m_k < 0} a_k^{-m_k}, \tag{5}$$

with $T_\ell = 0$ if and only if the two integers in (5) are equal.

Proof. The identity is [5, eq. 4.37.24]. With $\gamma^k t_0 = 19 \cdot 11^k / 20^{k+1}$ the rest is the substitution of (4): $T_\ell = \sum_k m_k \log r_k = \log \prod_k r_k^{m_k}$, and clearing denominators is legitimate because all a_k, b_k are positive. □

Thus the decision of the pairwise rule at this instance, at every depth, is a comparison of two positive integers, with no floating point anywhere and no tie threshold. This lemma is elementary and is not part of the formal development (Section 9); everything that follows in Sections 3–7 is.

3. The reduced ratios

Proposition 3.1. *At the instance, the ratios r_k of Lemma 2.1, computed from $\gamma = 11/20$ and $t_0 = 19/20$ and reduced to lowest terms, are*

k	0	1	2	3	4	5	6
a_k	39	609	10299	185289	3478179	67059969	1313659659
b_k	1	191	5701	134711	2921821	60940031	1246340341

and $\gcd(a_k, b_k) = 1$ for $0 \le k \le 6$.

Proof. In the formal development a_k and b_k are *defined* as $(20^{k+1} + 19 \cdot 11^k)/g_k$ and $(20^{k+1} - 19 \cdot 11^k)/g_k$ with $g_k = \gcd(20^{k+1} + 19 \cdot 11^k,\, 20^{k+1} - 19 \cdot 11^k)$; the table and the coprimality are then evaluated by the proof kernel from these definitions (fourteen exact integer computations). The table is a theorem about the derived values, not a definition. □

Remark 3.2. For orientation, the prime factorisations are

$$39 = 3 \cdot 13, \quad 609 = 3 \cdot 7 \cdot 29, \quad 191, \quad 10299 = 3 \cdot 3433, \quad 5701,$$
$$185289 = 3 \cdot 13 \cdot 4751, \quad 134711 = 23 \cdot 5857, \quad 3478179 = 3 \cdot 1159393, \quad 2921821 = 7^2 \cdot 59629,$$
$$67059969 = 3 \cdot 41 \cdot 545203, \quad 60940031 = 107 \cdot 569533,$$
$$1313659659 = 3 \cdot 13 \cdot 33683581, \quad 1246340341 = 317 \cdot 467 \cdot 8419,$$

where an unfactored entry is prime. These factorisations were obtained by trial division outside the formal development; the formal development contains exactly the divisibility statements of Theorem 4.1 and Proposition 7.1. Numerically, $c_0 = \log 39 \approx 3.66356$, $c_1 \approx 1.15954$, $c_2 \approx 0.59141$, $c_3 \approx 0.31879$, $c_4 \approx 0.17430$, $c_5 \approx 0.09570$, $c_6 \approx 0.05261$.

4. Multiplicative independence

Theorem 4.1. *Let* $(p_1, \dots, p_6) = (29, 3433, 23, 59629, 41, 317)$. *Then each* p_k *is prime, and for each* $k \in \{1, \dots, 6\}$*:*

(i) p_k *divides exactly one of the two integers* a_k, b_k*;*
(ii) p_k^2 *divides neither* a_k *nor* b_k*;*
(iii) p_k *divides none of the twelve integers* a_j, b_j *with* $0 \le j \le 6$, $j \ne k$.

Concretely, $29 \mid a_1$, $3433 \mid a_2$, $23 \mid b_3$, $59629 \mid b_4$, $41 \mid a_5$ *and* $317 \mid b_6$.

Proof. Primality is by trial division by every $d < 245$ (since $245^2 = 60025 > 59629$), and (i)–(iii) are $6 \cdot (2 + 2 + 14)$ divisibility tests on the integers of Proposition 3.1; all of it is evaluated by the proof kernel. □

The following two corollaries are the valuation argument that the certificate is designed for. They are ordinary mathematics and are not part of the formal development.

Corollary 4.2. $r_0, \dots, r_6$ *are multiplicatively independent in* $\mathbb{Q}_{>0}$*: if* $m \in \mathbb{Z}^7$ *and* $\prod_{j=0}^{6} r_j^{m_j} = 1$, *then* $m = 0$. *Equivalently,* $c_0, \dots, c_6$ *are linearly independent over* $\mathbb{Q}$.

Proof. Fix $k \in \{1, \dots, 6\}$ and apply the p_k-adic valuation to the relation. By Theorem 4.1, $v_{p_k}(r_j) = 0$ for $j \ne k$ and $v_{p_k}(r_k) = \pm 1$, so $0 = \sum_j m_j v_{p_k}(r_j) = \pm m_k$ and $m_k = 0$. With $m_1 = \cdots = m_6 = 0$ the relation reads $39^{m_0} = 1$, whence $m_0 = 0$. The equivalent form follows by taking logarithms. □

Corollary 4.3. *At the instance, for every depth* $\ell \le 6$ *and every realisation of the labelled tree,* $T_\ell \ne 0$. *Hence for* $\ell \le 6$ *the tie clause of* (3) *is never active:* $\mathcal{E}(\ell) = \mathbb{P}(y_o T_\ell < 0)$, *and the classifier* $h_\ell = \operatorname{sgn}(T_\ell)$ *never outputs* 0.

Proof. By Lemma 2.1, $T_\ell = 0$ means $\prod_{k \le \ell} r_k^{m_k} = 1$ with $m_0 = s_o \in \{\pm 1\}$, an integer vector that is not zero; Corollary 4.2 excludes it. □

We do not claim anything for $\ell \ge 7$: the certificate covers $r_0, \dots, r_6$ only, and extending it is a matter of exhibiting a distinguishing prime for each further r_k.

5. Tie-freeness at depth two, exhaustively

Corollary 4.3 rests on the valuation argument written out above. The following theorem is an independent, purely computational confirmation of the same fact at depth ≤ 2, on an explicit box of exponents; it is the statement the formal development carries about ties.

Theorem 5.1. *For every* $m_0 \in \{1, -1\}$, *every integer* m_1 *with* $|m_1| \le 30$ *and every integer* m_2 *with* $|m_2| \le 70$,

$$39^{m_0} \Big(\frac{609}{191}\Big)^{m_1} \Big(\frac{10299}{5701}\Big)^{m_2} \ne 1.$$

Equivalently, on this box $T_2 \ne 0$.

Proof. For each of the $2 \cdot 61 \cdot 141 = 17{,}202$ exponent triples the two integers of (5) are formed (they have up to 367 decimal digits, the extreme being $39 \cdot 609^{30} \cdot 10299^{70}$) and compared for equality; every comparison is carried out by the proof kernel by reduction of the defining function (no compiled evaluator is involved). At founding the check took 5.4 s and 1.86 GB of memory. □

Remark 5.2 (relation to the source's accounting rule). The source's rule counts lattice points with $|T_\ell| < 10^{-8}$ as errors, and its proof of Proposition C.1 says so twice ("counting ties conservatively as errors"; "lattice points with $|T_\ell| < 10^{-8}$ are counted as errors (the tie-conservative side)"). Corollary 4.3 and Theorem 5.1 show that no *exact* tie exists at depths ≤ 6, respectively on the box. For the near-tie guard, a computation outside the formal development gives

$$\min\{|T_2| :\ m_0 = \pm 1,\ |m_1| \le 30,\ |m_2| \le 70\} = 0.0011974\ldots,$$

attained at $(m_0, m_1, m_2) = (1, 5, -16)$, seven orders of magnitude above 10^{-8}; so on the box neither clause of the rule is ever triggered at depth 2, and there $\mathcal{E}(2) = \mathbb{P}(T_2 < 0)$ holds point by point. How the box compares with the region the source's computation visits can be read from the script `nonmono_certify.py` attached to the v1 e-print (Section 1.4): the root's group sizes are capped at $i \le 16$, $j \le 9$, so $m_1 \in [-(i+j), i+j] \subseteq [-25, 25]$; the per-child (m_2, m_3) kernels are truncated to $|m_2| \le 34$, $|m_3| \le 90$ and the root arrays to $|m_2| \le 64$, $|m_3| \le 150$ (the script's `B2K, B3K = 34, 90` and `B2, B3 = 64, 150`), mass outside a box being charged to the ledger. So every lattice point (s_o, m_1, m_2) at which the source's depth-2 computation evaluates the sign of T_2 lies in $\{\pm 1\} \times [-25, 25] \times [-64, 64]$, strictly inside the box of Theorem 5.1. The group-size caps alone would allow $|m_2| \le 32(i+j) \le 800$; it is the truncation boxes, not the caps, that confine m_2. Independently of the script, the unconditional probability that the tree has more than 70 level-2 vertices is 5.1×10^{-10} (a computation outside the formal development, $N_1 \sim \mathrm{Poi}(3)$ and $N_2 \mid N_1 \sim \mathrm{Poi}(3N_1)$), below the source's stated total ledger of 5×10^{-9}.

6. Depth one: the exact thresholds and the Skellam formula

At depth one, $T_1 = c_0 s_o + c_1 m_1 = \log\left(39^{s_o}(609/191)^{m_1}\right)$, so the decision depends on the position of m_1 relative to a threshold that depends on s_o.

Proposition 6.1. (a) *For every $a \in \mathbb{N}$, $39 \cdot 191^{\,a+4} < 609^{\,a+4}$.*

(b) *For $a \in \{0, 1, 2, 3\}$, $609^{\,a} < 39 \cdot 191^{\,a}$.*

Consequently, at the instance: if $s_o = +1$ then $T_1 \le 0$ if and only if $m_1 \le -4$; if $s_o = -1$ then $T_1 \le 0$ if and only if $m_1 \le 3$; and in both cases the inequality $T_1 \le 0$, when it holds, is strict.

Proof. (a) is proved by induction on a: the base case $39 \cdot 191^4 < 609^4$ is a kernel evaluation, and the step multiplies the hypothesis $39 \cdot 191^{n+4} + 1 \le 609^{n+4}$ by $191 \le 609$. It is an induction covering every exponent, not a finite check. (b) is four kernel evaluations. Both (a) and (b) are in the formal development; the passage to the thresholds is the following elementary argument, which is not. For $s_o = +1$, $T_1 \le 0$ means $39 \cdot 609^{m_1} \le 191^{m_1}$, impossible for $m_1 \ge 0$; for $m_1 = -a < 0$ it means $39 \cdot 191^a \le 609^a$, which holds for $a \ge 4$ by (a) and fails for $a \le 3$ by (b). For $s_o = -1$, $T_1 \le 0$ means $609^{m_1} \le 39 \cdot 191^{m_1}$, which holds for $m_1 < 0$ trivially, for $0 \le m_1 \le 3$ by (b), and fails for $m_1 \ge 4$ by (a). Since (a) and (b) are strict, so is $T_1 < 0$ whenever $T_1 \le 0$, in agreement with Corollary 4.3. □

Corollary 6.2. *Conditionally on $y_o = +1$, the counts m_1^+ and m_1^- of level-one vertices with $s = +1$ and $s = -1$ are independent Poisson variables with means*

$$\lambda_+ = \frac{\Delta(1 + \gamma t_0)}{2} = 2.28375, \qquad \lambda_- = \frac{\Delta(1 - \gamma t_0)}{2} = 0.71625,$$

independent of s_o, so that $m_1 = m_1^+ - m_1^- \sim \mathrm{Skellam}(\lambda_+, \lambda_-)$ [4], and

$$\mathcal{E}(1) = \frac{1 + t_0}{2}\,\mathbb{P}(m_1 \le -4) + \frac{1 - t_0}{2}\,\mathbb{P}(m_1 \le 3), \tag{6}$$

while $\mathcal{E}(0) = \mathbb{P}(s_o = -1) = \frac{1 - t_0}{2} = 0.025$.

Proof. A child of the root has $s = +1$ with probability $\frac{1+\gamma}{2}\cdot\frac{1+t_0}{2} + \frac{1-\gamma}{2}\cdot\frac{1-t_0}{2} = \frac{1+\gamma t_0}{2}$, independently across children and of s_o (which depends on y_o alone); thinning the $\mathrm{Poi}(\Delta)$ offspring gives the two independent Poisson counts, and (6) is Proposition 6.1 weighted by $\mathbb{P}(s_o = \pm 1 \mid y_o = +1) = \frac{1\pm t_0}{2}$. For $\mathcal{E}(0)$, $T_0 = s_o \log 39$ and $\log 39 > 0$. □

Remark 6.3. Evaluating (6) in floating point (a computation outside the formal development) gives $\mathcal{E}(1) = 0.022619380242067\ldots$, in agreement with the source's printed $0.02261938\ldots$. The source states the value without the thresholds; the thresholds -4 and 3 are the content of Proposition 6.1. The source's own Appendix C uses the same Skellam structure one level down (its step 1, a $\mathrm{Skellam}(3q, 3(1-q))$ variable with $q = \frac{1+\gamma t_0 \sigma_2}{2}$).

7. Controls

A finite verification that refutes nothing is easily satisfied by a misplaced quantifier. The formal development therefore also contains the following statements, each of which rules out a neighbouring wrong claim.

Proposition 7.1. (i) $609^3 < 39 \cdot 191^3$*: the depth-one threshold for* $s_o = +1$ *is not at* $m_1 = -3$ *(a too-small threshold is refuted).*

(ii) $39 \cdot 191^4 < 609^4$*: it is not at* $m_1 = -5$ *either, since* $m_1 = -4$ *already gives* $T_1 < 0$ *(a too-large threshold is refuted).*

(iii) *The tie predicate of Theorem 5.1 evaluated at* $(m_0, m_1, m_2) = (0, 0, 0)$ *is true, all three exponents being zero: the predicate is not false by its shape, and Theorem 5.1 is not vacuous.*

(iv) $3 \mid a_k$ *for every* $0 \le k \le 6$*: the prime* 3 *distinguishes nothing.*

(v) $13 \mid a_0$*,* $13 \mid a_3$ *and* $13 \mid a_6$*: the prime* 13 *cannot serve as a distinguishing prime for any* r_k*.*

Proof. Kernel evaluations, as in Proposition 3.1. Items (iv) and (v) show that the small primes occurring in the a_k genuinely fail the certificate and that Theorem 4.1 needs its large primes. □

8. Beyond depth three: evidence and a price

Nothing in this section is a theorem; the numbers are reported because they are the first values of $\mathcal{E}(\ell)$ at the instance beyond the source's, and because a certified version is within reach and can be priced.

8.1. **A scalar never-renormalising lattice.** The source's computation lives on the joint lattice of the vector (s_o, m_1, m_2, m_3), whose dimension grows with the depth. A scheme that scales to any depth replaces T_ℓ by two scalar surrogates. Fix a scale $D \in \mathbb{N}$ and put $A_k = \lfloor c_k D \rfloor$ and $B_k = \lceil c_k D \rceil$; then

$$V^{\mathrm{lo}} = \sum_k \left(A_k m_k^+ - B_k m_k^-\right) \;\le\; D\,T_\ell \;\le\; \sum_k \left(B_k m_k^+ - A_k m_k^-\right) = V^{\mathrm{hi}}.$$

Both surrogates are integer-linear in the additive counts $m_k^\pm$, so both propagate through the branching recursion exactly as T_ℓ does, but on a one-dimensional integer lattice: with $L_d^{\pm,\mathrm{lo}}$ the law of the lower surrogate of a depth-d subtree whose root has relative label ± 1, global sign symmetry gives $L_d^{-,\mathrm{lo}} = \mathrm{reflect}(L_d^{+,\mathrm{hi}})$, and the two laws are computed as a coupled pair by a compound-Poisson step (the $\mathrm{Poi}(\Delta)$ children, each of relative label $+1$ with probability $(1+\gamma)/2$ followed by the root's own message $\pm A_d/\pm B_d$ with probabilities $\frac{1\pm t_0}{2}$. Every step is monotone in each sub-probability argument, so truncation only removes mass, and

$$\nu^{\mathrm{hi}}(\{v \le 0\}) \;\le\; \mathcal{E}(\ell) \;\le\; 1 - \nu^{\mathrm{lo}}(\{v > 0\})$$

for the resulting sub-probability laws $\nu^{\mathrm{lo}}, \nu^{\mathrm{hi}}$ of the two surrogates at the root—the source's never-renormalise rule, carried by a different route.

8.2. **Measurements.** With $D = 2^{16}$, the compound-Poisson step done by a fast Fourier transform in double precision (length 2^{24}, 27–31 s per depth), the scheme gives the brackets in the following table, against the source's values and against a Monte Carlo run of 4×10^5 trees grown to depth 6 (half-widths are 1.96 standard errors):

ℓ	scalar lattice bracket	source [1]	Monte Carlo
0	[0.025000000, 0.025000000]	0.025	0.024903 ± 0.000483
1	[0.022619380, 0.022619380]	0.02261938 . . .	0.022600 ± 0.000461
2	[0.038706416, 0.038706418]	[0.03870641, 0.03870642]	0.038703 ± 0.000598
3	[0.075406246, 0.075937117]	[0.07589622, 0.07589623]	0.075370 ± 0.000818
4	[0.107141742, 0.107187204]	—	0.106393 ± 0.000956
5	[0.133642479, 0.133742878]	—	0.133485 ± 0.001054
6	—	—	0.155765 ± 0.001124

Rows 0–2 reproduce the source to its printed digits and row 3 contains the source's bracket; rows 4 and 5 are new, and both lie inside the Monte Carlo intervals. Taking the source's values for $\ell \le 3$ (midpoints of its brackets), the midpoints of our brackets for $\ell = 4, 5$ and the Monte Carlo value at $\ell = 6$, the increments $\mathcal{E}(\ell) - \mathcal{E}(\ell - 1)$ for $\ell = 2, \ldots, 6$ are 0.0161, 0.0372, 0.0313, 0.0265, 0.0221: the strict increase that Proposition C.1 certifies from depth 1 to depth 3 continues, on this evidence, through depth 6, with increments that peak at $\ell = 3$ and then decay. This is consistent with, and no proof of, convergence of $\mathcal{E}(\ell)$ at the instance, which the source does not establish there (Section 1.3); we claim nothing about that limit either way.

Two caveats. First, the floating-point Fourier step is not rigorous, so these brackets are evidence, not certified enclosures, even though the scheme itself is. Second, the bracket at $\ell = 3$ is loose (5×10^{-4} wide while the others are below 10^{-4}) because the width is essentially $\mathbb{P}(|T_\ell| \lesssim U/D)$ with U the number of vertices, and T_3 has a lattice atom within 6×10^{-4} of zero carrying about 5×10^{-4} of mass, which the coefficient quantisation straddles at every D; this is exactly the phenomenon that the exact rational comparison of Lemma 2.1 removes, and it is the reason the source works on the joint lattice rather than on a scalar grid.

8.3. **The price of a certified $\mathcal{E}(4)$.** The scalar scheme can be run in exact integer arithmetic. Its bracket width is about $1.9/D$; to separate $\mathcal{E}(4)$ from the source's certified $\mathcal{E}(3) \le 0.07589623$ it suffices to certify $\mathcal{E}(4) \ge 0.09$, i.e. a width of about 0.015, i.e. $D \approx 256$: arrays of about 2×10^4 cells, per-level kernels of about 10^4 cells, and about 20 convolution steps per level over four levels, roughly 1.6×10^{10} integer multiply-adds. In C this is about 30 s. Inside a proof assistant it is a compiled extension rather than an interpreted evaluation (at an interpreted penalty of several hundred times it would exceed a few CPU-hours), with fixed-point integers rounded down so that the mass deficit remains the ledger; the estimated loss from the fixed-point floor at that scale is about 4×10^{-4}, comfortably inside a 0.015 bracket. The engineering of that extension, not the compute, is the cost, and it is where a certified $\mathcal{E}(4)$—and, at about four times the size, $\mathcal{E}(5)$—stands. The source's joint-lattice route, by contrast, needs a three-dimensional array of about 5×10^6 cells at depth 4 and 170 three-dimensional convolutions, and a fourth dimension at depth 5 (about 10^9 complex cells), which is why we do not price it.

9. Verification

Formalised and machine-checked in Lean 4 [6] (toolchain v4.32.0) are Proposition 3.1 (two declarations, the table and the coprimality), Theorem 4.1 (two declarations, primality and the divisibility pattern), Theorem 5.1, Proposition 6.1(a) and (b), and the five items of Proposition 7.1; the formal statements are reproduced verbatim in Appendix A. The module imports nothing from Mathlib and nothing beyond the repository's attribute definitions: the ratios are natural-number arithmetic with the built-in gcd, primality is trial division (a Boolean function sound below $245^2 = 60025$), and the tie predicate of Theorem 5.1 forms and compares the two integers of (5) directly. All finite

evaluations, including the 17,202 comparisons of Theorem 5.1, are performed by the kernel (`decide`, respectively `decide +kernel`); no compiled evaluator (`native_decide`) is used anywhere. Reading the axioms of every declaration with `#print axioms` gives the following.

declaration	axioms
`ratio_table`, `ratio_reduced`, `dprime_prime`, `thr_above`	none
`no_tie_depth_two`, `control_trivial_tie`	none
`control_threshold_not_three`, `control_threshold_is_four`	none
`dprime_distinguishes`, `control_three_divides_all`	`propext`
`control_thirteen_not_distinguishing`	`propext`
`thr_below` (the induction)	`propext`, `Quot.sound`

There is no `sorry`, no axiom declared by the development, and no use of `Classical.choice`. Not formalised, and written out above as ordinary mathematics, are Lemma 2.1, Corollaries 4.2, 4.3 and 6.2, the passage from Proposition 6.1(a),(b) to the thresholds, and the computations of Remarks 3.2, 5.2 and 6.3; nothing in Section 8 is formalised or claimed. The repository is private; repository reference to be added at submission.

Appendix A. The formal statements

The definitions and the pinned theorem statements of the Lean module, verbatim (proofs omitted; each is `decide`, `decide +kernel` for `no_tie_depth_two`, and the induction described in Proposition 6.1 for `thr_below`).

```
def xn (k : Nat) : Nat := 19 * 11 ^ k
def xd (k : Nat) : Nat := 20 ^ (k + 1)
def rG (k : Nat) : Nat := Nat.gcd (xd k + xn k) (xd k - xn k)
def rNum (k : Nat) : Nat := (xd k + xn k) / rG k
def rDen (k : Nat) : Nat := (xd k - xn k) / rG k

theorem ratio_table :
    (List.range 7).map (fun k => (rNum k, rDen k)) =
      [(39, 1), (609, 191), (10299, 5701), (185289, 134711),
       (3478179, 2921821), (67059969, 60940031), (1313659659, 1246340341)]

theorem ratio_reduced :
    (List.range 7).all (fun k => Nat.gcd (rNum k) (rDen k) == 1) = true

def dprime : Nat → Nat
  | 1 => 29
  | 2 => 3433
  | 3 => 23
  | 4 => 59629
  | 5 => 41
  | 6 => 317
  | _ => 0

def isPrimeB (n : Nat) : Bool :=
  2 ≤ n && (List.range 245).all (fun d => d < 2 || n < d * d || n % d != 0)

theorem dprime_prime :
    ((List.range 6).map (· + 1)).all (fun k => isPrimeB (dprime k)) = true

theorem dprime_distinguishes :
    ((List.range 6).map (· + 1)).all (fun k =>
      let p := dprime k
      -- it hits r_k on exactly one side
      (((p ∣ rNum k) && !(p ∣ rDen k)) || (!(p ∣ rNum k) && (p ∣ rDen k)))
```

```
      -- to the first power only
      && !(p * p ∣ rNum k) && !(p * p ∣ rDen k)
      -- and it misses every other r_j on both sides
      && (List.range 7).all (fun j =>
            (j == k) || (!(p ∣ rNum j) && !(p ∣ rDen j)))) = true

theorem thr_below : ∀ a : Nat, 39 * 191 ^ (a + 4) < 609 ^ (a + 4)

theorem thr_above : (List.range 4).all (fun a => 609 ^ a < 39 * 191 ^ a) = true

def zpowND (n d : Nat) (m : Int) : Nat × Nat :=
  if 0 ≤ m then (n ^ m.toNat, d ^ m.toNat) else (d ^ (-m).toNat, n ^ (-m).toNat)

def tieAt (m0 m1 m2 : Int) : Bool :=
  let a := zpowND 39 1 m0
  let b := zpowND 609 191 m1
  let c := zpowND 10299 5701 m2
  a.1 * b.1 * c.1 == a.2 * b.2 * c.2

def noTie2 : Bool :=
  (List.range 61).all fun i =>
    (List.range 141).all fun j =>
      let m1 : Int := (i : Int) - 30
      let m2 : Int := (j : Int) - 70
      !(tieAt 1 m1 m2) && !(tieAt (-1) m1 m2)

theorem no_tie_depth_two : noTie2 = true

theorem control_threshold_not_three : 609 ^ 3 < 39 * 191 ^ 3
theorem control_threshold_is_four : 39 * 191 ^ 4 < 609 ^ 4
theorem control_trivial_tie : tieAt 0 0 0 = true
theorem control_three_divides_all :
    (List.range 7).all (fun k => 3 ∣ rNum k) = true
theorem control_thirteen_not_distinguishing :
    (13 ∣ rNum 0) ∧ (13 ∣ rNum 3) ∧ (13 ∣ rNum 6)
```

References

[1] A. R. Baranwal, *The value of depth in message passing on sparse graphs: a Kesten–Stigum dichotomy*, arXiv:2607.16676v2 [math.ST] (cross-listed to cs.LG, math.PR, stat.ML), v1 18 July 2026, v2 29 August 2026, 27 pages; doi:10.48550/arXiv.2607.16676. No journal version is recorded for it in the arXiv metadata. The ancillary files announced in the arXiv comments (among them `anc/nonmono_certify.py`) are attached to v1 only; the v2 e-print carries none. Numbering above is that of the v2 PDF, cross-checked against a compilation of the v2 e-print: the certified instance is Proposition C.1 of Appendix C, the classifier and its error are equations (4) and (5) of Section 2.2, the below-threshold theorem is Theorem 3.1, and the non-monotonicity observation is Observation 3.9.

[2] A. Baranwal, K. Fountoulakis and A. Jagannath, *Optimality of message-passing architectures for sparse graphs*, Advances in Neural Information Processing Systems 36 (NeurIPS 2023); arXiv:2305.10391 (v3, 9 January 2025). The theorem numbers quoted in Section 1 are those given by [1] for the updated arXiv version.

[3] A. Baranwal, *Statistical foundations for learning on graphs*, PhD thesis (Computer Science), University of Waterloo, 2024; UWSpace, https://hdl.handle.net/10012/21209 (issued 27 November 2024).

[4] J. G. Skellam, *The frequency distribution of the difference between two Poisson variates belonging to different populations*, J. Roy. Statist. Soc. **109** (1946), no. 3, p. 296; doi:10.2307/2981372.

[5] NIST Digital Library of Mathematical Functions, https://dlmf.nist.gov/, equation 4.37.24 ($\operatorname{arctanh} z = \frac{1}{2}\ln\frac{1+z}{1-z}$).

[6] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: A. Platzer and G. Sutcliffe (eds.), Automated Deduction — CADE 28, Lecture Notes in Computer Science **12699**, Springer, Cham, 2021, pp. 625–635; doi:10.1007/978-3-030-79876-5_37.