

A Capacity Criterion for Four-State Potts Reconstruction on Trees

OpenAI

October 5, 2026

Abstract

For the ferromagnetic four-state broadcast model with $0 < \lambda < 1$, we prove that reconstruction on a bounded-degree deterministic rooted tree occurs exactly when its L^3 capacity with edge resistances $\lambda^{-2|e|}$ is positive. This gives an exact criterion without regularity or growth-rate assumptions on the tree, including at the exponential critical boundary.

1 Introduction

Let T be an infinite locally finite rooted tree with root ρ . Assume that every vertex has at most B children, where B is a fixed positive integer; vertices with zero or one child are allowed. Write $|v|$ for the depth of a vertex, $\text{ch}(v)$ for its children, and $L_n = \{v : |v| = n\}$. The tree is deterministic and known to the observer.

Remark (notation). $L_n = \{v \in V(T) : |v| = n\}$ for $n \in \mathbb{N}_0$.

Fix $0 < \lambda < 1$. Give the root a uniform spin in $[4]$, and transmit spins independently along edges conditional on their parent spins, using

$$P_\lambda(j \mid i) = \lambda \mathbf{1}_{\{i=j\}} + \frac{1-\lambda}{4}. \quad (1.1)$$

For the uniform probability vector $\mathbf{u} = (1/4, 1/4, 1/4, 1/4)$, put

$$a_n(T, \lambda) = \mathbb{E} \text{TV}(\mathbb{P}(\sigma_\rho \in \cdot \mid \sigma_{L_n}), \mathbf{u}), \quad \text{TV}(p, \mathbf{u}) = \frac{1}{2} \sum_{i=1}^4 |p_i - 1/4|.$$

Data processing makes a_n nonincreasing. We say that *reconstruction occurs* if its limit $a_\infty(T, \lambda)$ is positive.

Remark (notation). $P_i = 4p_i q_i / Z$ on $\{Z > 0\}$; its combined posterior law is the image under P of the measure $Z d(\mu \otimes \nu)$, and its value on $\{Z = 0\}$ may be any probability vector. In Lemma 4.1: $Z = 1 + s \geq 1 - \lambda > 0$, so the displayed coordinate formula defines P throughout the product sampling law.

A nonnegative flow θ to infinity assigns a nonnegative number to each edge, with incoming flow equal to the sum of outgoing flows at every nonroot vertex. Its mass is

$$|\theta| = \sum_{w \in \text{ch}(\rho)} \theta(\rho, w).$$

For an edge e , let $|e|$ be the depth of its child endpoint, and let ∂T denote the infinite rays from the root. Define

$$V_\lambda(\theta, \xi) = \sum_{e \in \xi} (\lambda^{-2|e|} \theta(e))^2, \quad \xi \in \partial T, \quad (1.2)$$

$$\text{Cap}_{3,\lambda}(T) = \sup \left\{ |\theta| : \sup_{\xi \in \partial T} V_\lambda(\theta, \xi) \leq 1 \right\}. \quad (1.3)$$

The zero flow is allowed. This is the L^3 nonlinear capacity with edge resistances $\lambda^{-2|e|}$: in the standard L^p normalization, the path potential has exponent $p - 1$ [8, Section 2].

Theorem 1.1. *Let T be an infinite rooted tree with at most $B < \infty$ children per vertex, and let $0 < \lambda < 1$. For the broadcast model (1.1),*

$$a_\infty(T, \lambda) > 0 \iff \text{Cap}_{3,\lambda}(T) > 0.$$

Remark. For $x \in \mathbb{R}^3$, $|x|^2 = x_1^2 + x_2^2 + x_3^2$. The combined posterior law is the law of P under $Z d(\mu \otimes \nu)$, with $P_i = 4p_i q_i / Z$ on $\{Z > 0\}$ and any measurable simplex-valued extension on $\{Z = 0\}$.

There is no regularity or growth-rate assumption on the tree. In particular, the criterion applies at the exponential Kesten–Stigum boundary, where an exponential-growth test alone does not settle the question. The exclusion of $\lambda = 1$ is essential: an infinite ray then reconstructs perfectly, whereas its capacity in (1.3) is zero.

History and relation to earlier work

The second-eigenvalue scale in tree broadcasting originates in the branching-process work of Kesten and Stigum [3]. For a regular tree with d children per vertex, the corresponding bound guarantees reconstruction when $d\lambda^2 > 1$ [5]. Evans, Kenyon, Peres, and Schulman [1] developed the relation between broadcasting, the Ising model, and electrical capacity. For Potts channels, Sly [9] proved sharpness of the second-eigenvalue bound for three states on regular trees of sufficiently large degree, including nonreconstruction at equality, and proved nonsharpness for at least five states. Mossel, Sly, and Sohn [6] established three- and four-state nonreconstruction at and below the bound for Galton–Watson families of sufficiently large mean degree, under hypotheses covering regular and Poisson offspring. At four states the quadratic correction vanishes, making higher-order terms decisive [9, 6].

Pemantle and Peres [8] formulated critical Ising criteria on arbitrary trees using nonlinear capacities and concave recursions. Their free-boundary criterion uses L^2 capacity with squared channel resistances, whereas their plus-boundary criterion uses L^3 capacity with unsquared channel resistances [8, Theorems 2.1–2.2]. The present four-state reconstruction criterion has the L^3 exponent but the squared channel resistance $\lambda^{-2|e|}$. The nonlinear recursion used below is the $p = 3$ instance of their capacity recursion [8, Lemma 3.1 and Corollary 3.4]; we give the needed flow construction directly.

Our essential probabilistic input is the preserved class of posterior laws and its moment-saving inequality from *The Reconstruction Threshold for the Ferromagnetic Four-State Potts Model* [7, Proposition 3.2]. That companion studies the threshold on regular and Poisson Galton–Watson trees. Its posterior-law statement allows arbitrary channel parameters and conditionally independent products, which is the scope needed here. We state this input precisely in Section 2; its polynomial-certificate proof is not repeated. All remaining arguments for Theorem 1.1 are proved below.

Proof strategy

We measure a posterior p 's information by the quadratic statistic $m = 4\mathbb{E}|p - \mathbf{u}|^2$. The companion inequality saves a cubic term when two informative branches are combined. For the implication from reconstruction to capacity, a change of variables $m \mapsto m + Km^3$ turns this branching saving into a nonlinear inequality that also has a strict saving at one-child vertices. A direct telescoping construction then produces an admissible flow of positive mass.

The converse starts from a flow and scales it down. At each observation level we add independent noise, chosen separately at each vertex v so that its terminal information equals the scaled flow entering v divided by $\lambda^{2|v|}$. The resulting information stays below this depth-weighted flow profile. To bound what is lost, we track a signed third moment in addition to m . In the four-state coordinates, the mixed quadratic term in the product formula cancels. The remaining loss is controlled by the third moment, whose propagation along an edge has factor λ^3 , rather than λ^2 .

Sampling a ray according to the flow converts the accumulated loss into sums along a single path. The two propagation factors leave a summable kernel λ^{j-i} between depths i and j . The capacity bound controls its quadratic path sum, and a single choice of noise scale works at every observation depth. This flow-calibrated degradation and third-moment estimate avoid any need for comparable branch sizes or a strict exponential growth inequality.

Section 2 introduces the posterior calculus and the companion input. Section 3 constructs the flow from reconstruction. Section 4 proves the local moment estimates, and Section 5 uses them to prove the converse. Throughout, set

$$r = \lambda^2 \in (0, 1).$$

2 Posterior experiments and the companion input

An experiment consists of an observation about a uniform input spin in [4]. Let p be its posterior vector and μ the law of p under the marginal law of the observation. A posterior law is *symmetric* if it is invariant under all permutations of the four entries. Define

$$L = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}, \quad x = Lp, \quad A(p) = |x|^2,$$

and put

$$m(\mu) = \mathbb{E}_\mu A, \quad \tau(\mu) = \mathbb{E}_\mu [x_1 x_2 x_3], \quad h(\mu) = |\tau(\mu)|. \quad (2.1)$$

Here and below, vector norms are Euclidean. The rows of L are orthogonal and span the orthogonal complement of the constant vector. Consequently

$$A(p) = 4|p - \mathbf{u}|^2, \quad 0 \leq A(p) \leq 3, \quad \frac{1}{8}A(p) \leq \text{TV}(p, \mathbf{u}) \leq \frac{1}{2}\sqrt{A(p)}. \quad (2.2)$$

The lower bound follows from $|p_i - 1/4|^2 \leq |p_i - 1/4|$; the upper bound is Cauchy–Schwarz. In expectation,

$$\frac{1}{8}m(\mu) \leq \mathbb{E} \text{TV}(p, \mathbf{u}) \leq \frac{1}{2}\sqrt{m(\mu)}. \quad (2.3)$$

Under degradation of an observation, the new posterior is the conditional expectation of the old one given the new observation. Convexity of A therefore makes m nonincreasing under data processing.

Remark (notation). μ denotes the distribution of the posterior vector p for an experiment about a uniform input spin in [4], taken under the marginal law of the observation. In Theorem 1.1: The posterior laws of finite-depth subtree experiments, including descendant-level observations, independently channel-degraded terminal observations, and partial products at a vertex, belong to $\mathcal{C}$. In Proposition 2.1: $\mu \in \mathcal{C}$, where $\mathcal{C}$ is the class of symmetric posterior laws specified in the statement. In Lemma 4.1: μ is symmetric, meaning invariant under all permutations of its four posterior entries; membership in $\mathcal{C}$ is not required.

2.1 Channels and products

In a channel step with parameter $t \in [0, 1]$, first transmit a uniform input I to J through P_t , then observe an experiment about J . The uniform distribution is stationary and the channel is symmetric, so the posterior about I is $u + t(p - u)$. Thus

$$x \mapsto tx, \quad m \mapsto t^2m, \quad \tau \mapsto t^3\tau, \quad h \mapsto t^3h. \quad (2.4)$$

The completely revealing experiment has $m = 3$, while the uninformative experiment has $m = h = 0$.

Remark (notation). In a channel step, $t \in [0, 1]$ denotes the parameter of the channel P_t . In Theorem 1.1: In the tangent inequality $(1 + 2t)^{-1/2} \geq 1 - t$ for $t \geq 0$, the letter t is a local dummy variable ranging over $[0, \infty)$, independent of its use as a channel parameter.

If two observations are independent conditional on the same input spin, write $\mu \star \nu$ for their combined posterior law. The normalized-product rule below is standard in posterior recursions for reconstruction; see Mézard and Montanari [4, Sections 2.2–2.4]. Sample p, q independently from their marginal posterior laws μ, ν . The joint marginal law of the observations is obtained by weighting this product law by

$$Z = 4 \sum_{i=1}^4 p_i q_i, \quad P_i = \frac{4p_i q_i}{Z}. \quad (2.5)$$

The value of P at $Z = 0$ is irrelevant. Indeed, the likelihood of each observation conditional on input i , relative to its marginal law, is $4p_i$ or $4q_i$. Formula (2.5) follows by averaging their product over the uniform input. In particular, $\mathbb{E}Z = 1$. Symmetry is preserved by both channel steps and products.

Remark (notation). ν is the second marginal posterior law for an experiment about a uniform input spin in [4]. In Proposition 2.1: $\nu \in \mathcal{C}$, where $\mathcal{C}$ is the class of symmetric posterior laws specified in the statement. In Lemma 4.1: ν is a symmetric posterior law obtained by a channel step of parameter λ , with $0 < \lambda < 1$. $\mu \star \nu$ is the law of P under $Z d(\mu \otimes \nu)$, where $Z = 4 \sum_i p_i q_i$ and $P_i = 4p_i q_i / Z$ on $\{Z > 0\}$, with arbitrary simplex-valued extension on $\{Z = 0\}$. q is the second marginal posterior vector, with law ν , for an experiment about a uniform input spin in [4]. In Lemma 4.1: $q_i \geq (1 - \lambda)/4$ for every $i \in \{1, 2, 3, 4\}$, since its law is obtained by a channel step of parameter λ . $Z = 4 \sum_{i=1}^4 p_i q_i$ is the density of the joint marginal observation law relative to the product of its marginal laws, with $\int Z d(\mu \otimes \nu) = 1$. In Lemma 4.1: $Z = 1 + s \geq 1 - \lambda > 0$.

The columns of L are the four sign triples whose product is one. Permuting these columns acts on x by coordinate permutations and even sign changes. For any symmetric posterior law, it follows that

$$\mathbb{E}x_i = 0, \quad \mathbb{E}x_i x_j = \frac{m}{3} \mathbf{1}_{\{i=j\}}, \quad (2.6)$$

and every degree-three coordinate monomial except $x_1x_2x_3$ has zero expectation. Also $|x_i| \leq 1$. Any coordinate monomial of degree at least two is therefore bounded in absolute value by $|x|^2$: retain two factors and use $|x_ix_j| \leq (x_i^2 + x_j^2)/2$, or $|x_i|^2$ when the indices agree. In particular,

$$h(\mu) \leq m(\mu). \quad (2.7)$$

2.2 The moment-saving input

We use the following statement from the companion manuscript. The existence and preservation of the class, not only symmetry, are essential to its upper bound.

Proposition 2.1 (Companion input, [7, Proposition 3.2 and Equations (3.8)–(3.9)]). *There is a class $\mathcal{C}$ of symmetric posterior laws containing the revealing and uninformative laws and preserved by channel steps with any parameter in $[0, 1]$ and by conditionally independent products. For $\mu, \nu \in \mathcal{C}$,*

$$m(\mu \star \nu) \leq m(\mu) + m(\nu) - c \left(\mathbb{E}_\mu[A^2]m(\nu) + m(\mu)\mathbb{E}_\nu[A^2] \right), \quad c = \frac{1}{1000}. \quad (2.8)$$

Every finite-depth subtree experiment used here has law in $\mathcal{C}$. This includes observations of a descendant level and observations in which each terminal spin is independently passed through its own channel P_t . At the terminals, membership follows from the revealing law and channel closure; an empty observation is uninformative. Working toward the subtree root uses channel steps and products, with conditional independence supplied by the broadcast process. The same argument covers every partial product at a vertex. No assertion about a limiting posterior law belonging to $\mathcal{C}$ will be needed.

3 From reconstruction to capacity

The nonlinear capacity recursion of Pemantle and Peres [8, Lemma 3.1 and Corollary 3.4] has a particularly direct flow interpretation in the present normalization. We first give that construction, then derive its hypothesis from the posterior information.

Lemma 3.1 (A flow from recursive labels). *Suppose nonnegative finite labels $(z_v)_{v \in T}$ satisfy*

$$z_v \leq \sum_{w \in \text{ch}(v)} \frac{rz_w}{\sqrt{1 + z_w^2}}. \quad (3.1)$$

Then there is a flow of mass z_ρ whose potential (1.2) is at most one on every ray. In particular, $\text{Cap}_{3,\lambda}(T) \geq z_\rho$.

Proof. At each vertex choose coefficients

$$0 \leq d_{vw} \leq \frac{r}{\sqrt{1 + z_w^2}} \quad \text{such that} \quad z_v = \sum_{w \in \text{ch}(v)} d_{vw}z_w.$$

If the upper-bound sum is positive, multiply all its coefficients by the ratio of z_v to that sum; otherwise all labels in the displayed equality are zero and take all coefficients to be zero. Set $\ell_\rho = 1$, $\ell_w = \ell_v d_{vw}$ recursively, and define

$$\theta(v, w) = \ell_w z_w.$$

The displayed equality gives conservation at every nonroot vertex and mass z_ρ at the root. Put $g_v = r^{-|v|}\ell_v$. Then

$$g_w^2(1 + z_w^2) \leq g_v^2, \quad (r^{-|w|}\theta(v, w))^2 = (g_w z_w)^2 \leq g_v^2 - g_w^2.$$

Summing along any finite initial segment of a ray gives a bound by $g_\rho^2 = 1$. Monotone convergence of the nonnegative sums proves the assertion for every infinite ray. $\square$

For a vertex v , let $m_v^{(n)}$ be the quadratic information about σ_v obtained from its descendants at distance n , with $m_v^{(0)} = 3$. Data processing gives limits

$$m_v = \lim_{n \rightarrow \infty} m_v^{(n)} \in [0, 3].$$

If reconstruction occurs, (2.3) implies $m_\rho > 0$. We show that these limiting numbers yield labels satisfying Lemma 3.1 with a positive root label.

For a fixed vertex v , put

$$b_w = rm_w, \quad S = \sum_{w \in \text{ch}(v)} b_w, \quad D = \max_{w \neq w'} b_w b_{w'} (b_w + b_{w'}),$$

where $D = 0$ if fewer than two children are present. Then

$$m_v \leq S - cD. \tag{3.2}$$

At finite depth the parent information is $m_v^{(n+1)}$, and the child values after their edge channels are $b_w = rm_w^{(n)}$. Choose any pair of children and combine their branch experiments first. Jensen's inequality gives $\mathbb{E}A^2 \geq (\mathbb{E}A)^2$, so Proposition 2.1 saves at least $cb_w b_{w'} (b_w + b_{w'})$. Combine the remaining children using the subadditivity part of the same proposition. Maximize over the pair and pass to the limit. All sums and maxima are finite. With zero or one child, the empty observation or channel scaling gives the assertion directly.

There are at most B^3 ordered triples of children, and every mixed product in S^3 is at most the largest b_w squared times the second largest, hence at most D . Therefore

$$S^3 \leq \sum_w b_w^3 + B^3 D. \tag{3.3}$$

Choose $K > 0$ with $KB^3 \leq c$ and set $y_v = m_v + Km_v^3$. Since $m_v \leq S$, Equations (3.2) and (3.3) yield

$$\begin{aligned} y_v &\leq S - cD + KS^3 \leq S + K \sum_w b_w^3 \\ &= r \sum_w y_w - Kr(1 - r^2) \sum_w m_w^3. \end{aligned} \tag{3.4}$$

This correction supplies a saving even at a one-child vertex, where the original pair term vanishes: the positive factor is $r - r^3$.

The bound $m_w \leq 3$ gives $y_w \leq (1 + 9K)m_w$. With

$$\delta = \frac{K(1 - r^2)}{(1 + 9K)^3} > 0,$$

Equation (3.4) implies

$$y_v \leq r \sum_w y_w (1 - \delta y_w^2) \leq r \sum_w \frac{y_w}{\sqrt{1 + 2\delta y_w^2}}.$$

The second inequality is the tangent bound $(1 + 2t)^{-1/2} \geq 1 - t$ for $t \geq 0$. Thus $z_v = \sqrt{2\delta} y_v$ satisfies (3.1), with $z_\rho > 0$. Lemma 3.1 proves positive capacity and completes the first implication of Theorem 1.1.

4 Local estimates for information loss

For the converse, subadditivity alone is not enough: we must bound the information lost when branches are combined. The next lemma uses only symmetry and one channel-smoothed input. Unlike Proposition 2.1, it does not require membership in $\mathcal{C}$.

Lemma 4.1. *Fix $0 < \lambda < 1$. Let μ, ν be symmetric posterior laws, with ν obtained by a channel step of parameter λ . There is a finite constant $C = C(\lambda)$ such that*

$$m(\mu \star \nu) \geq m(\mu) + m(\nu) - C(h(\mu)m(\nu) + m(\mu)h(\nu)), \quad (4.1)$$

$$h(\mu \star \nu) \leq h(\mu) + h(\nu) + Cm(\mu)m(\nu). \quad (4.2)$$

The first inequality holds with $C = 7$.

Proof. All expectations in this proof are over independent marginal posterior vectors p, q of laws μ, ν . Put

$$x = Lp, \quad y = Lq, \quad s = x \cdot y, \quad a = m(\mu), \quad b = m(\nu), \quad \alpha = \tau(\mu), \quad \beta = \tau(\nu).$$

Since $p = \mathbf{u} + L^\top x/4$ and similarly for q , the product rule (2.5) becomes

$$Z = 1 + s, \quad LP = \frac{N}{1 + s}, \quad N = x + y + R, \quad R_i = x_j y_k + x_k y_j, \quad (4.3)$$

where $\{i, j, k\} = \{1, 2, 3\}$. Each coordinate of q is at least $(1 - \lambda)/4$, so

$$1 + s \geq 1 - \lambda > 0, \quad -\lambda \leq s \leq 3. \quad (4.4)$$

Remark (notation). $s = x \cdot y$ satisfies $-\lambda \leq s \leq 3$ and $1 + s \geq 1 - \lambda > 0$.

The quadratic statistic. Using the marginal weight Z in (2.5),

$$m(\mu \star \nu) = \mathbb{E} \frac{|N|^2}{1 + s}.$$

The second moments in (2.6) give

$$\mathbb{E}|R|^2 = \frac{2}{3}ab, \quad \mathbb{E}s^2 = \frac{1}{3}ab.$$

In the expansion of $\mathbb{E}[|N|^2(1 - s)]$, the term $\mathbb{E}|R|^2$ therefore cancels $2\mathbb{E}s^2$. Terms of degree one in either vector have zero expectation. For the remaining terms,

$$\mathbb{E}[sx \cdot R] = 2\alpha b, \quad \mathbb{E}[sy \cdot R] = 2a\beta, \quad \mathbb{E}[s|R|^2] = 6\alpha\beta.$$

For example, $x \cdot R = 2(x_1x_2y_3 + x_1x_3y_2 + x_2x_3y_1)$; after multiplication by s , only terms proportional to $x_1x_2x_3$ survive expectation over x . The last identity follows in the same way by retaining only terms with all three indices in each vector. Thus the expansion is exactly

$$\mathbb{E}[|N|^2(1-s)] = a + b - 4(ab + a\beta) - 6\alpha\beta. \quad (4.5)$$

As $(1+s)^{-1} = 1 - s + s^2/(1+s)$, we obtain

$$m(\mu \star \nu) = a + b - 4(ab + a\beta) - 6\alpha\beta + \mathbb{E} \frac{|N|^2 s^2}{1+s}. \quad (4.6)$$

The final term is nonnegative. Moreover, (2.7) implies

$$|\alpha\beta| \leq \frac{1}{2}(h(\mu)b + ah(\nu)).$$

This proves (4.1) with constant seven, without any sign assumption on α or β .

The third moment. The corresponding product formula is

$$\tau(\mu \star \nu) = \mathbb{E} \frac{N_1 N_2 N_3}{(1+s)^2}.$$

Use the exact remainder

$$(1+s)^{-2} - (1-2s) = \frac{s^2(3+2s)}{(1+s)^2}. \quad (4.7)$$

The numerator $N_1 N_2 N_3$ is bounded, and (4.4) bounds the multiplier in (4.7) by a constant depending only on λ times s^2 . The error in replacing the denominator factor by $1-2s$ is consequently at most $C_0(\lambda)ab$.

In the finite polynomial $N_1 N_2 N_3(1-2s)$, the only pure terms are $x_1x_2x_3 + y_1y_2y_3$. Every mixed monomial of degree one in one vector has zero expectation. Every other mixed monomial has degree at least two in each vector, so its absolute expectation is at most ab , by independence and the monomial bound following (2.6). Summing the finitely many coefficients and taking absolute values proves (4.2). Increase the common $C(\lambda)$ to at least seven to obtain both estimates. $\square$

5 From capacity to reconstruction

Assume $\text{Cap}_{3,\lambda}(T) > 0$ and choose an admissible flow θ of positive mass. We construct degraded observations at every finite depth and bound their information from below uniformly in that depth. The noise scale will be fixed only after the total loss has been estimated. Noisy boundary observations also underlie robust reconstruction [2]. Here we choose the noise separately at each vertex from the flow and use the resulting information to bound that of the original observations.

5.1 A flow profile and terminal observations

For $0 < \epsilon \leq 1$, let $\Theta(v)$ be the flow $\epsilon\theta$ entering a nonroot vertex v , and define

$$\Theta(\rho) = M := \epsilon|\theta|, \quad X_v = r^{-|v|}\Theta(v).$$

Conservation becomes

$$X_v = r \sum_{w \in \text{ch}(v)} X_w. \quad (5.1)$$

Every positive-flow edge extends to an infinite ray: at every positive-flow vertex, conservation gives at least one positive-flow child. Hence admissibility implies $X_v \leq \epsilon$ for all nonroot vertices. In particular, $\theta(\rho, w) \leq r$ for each child of the root, and $|\theta| \leq Br$. On every ray $\xi = (v_0 = \rho, v_1, v_2, \dots)$, we therefore have

$$\sum_{j=0}^{\infty} X_{v_j}^2 \leq Q\epsilon^2, \quad Q = 1 + (Br)^2. \quad (5.2)$$

The extra term in Q is needed because $X_\rho = M$ need not be at most ϵ .

Fix $n \geq 1$. At each $v \in L_n$, observe its spin only after an additional independent channel P_{t_v} , where

$$t_v = \sqrt{X_v/3}. \quad (5.3)$$

These are valid parameters since $X_v \leq \epsilon \leq 1$. For $|v| \leq n$, let $\hat{m}_v, \hat{h}_v$ be the statistics of the experiment about σ_v using these degraded terminal observations in its subtree. The dependence on n is suppressed. These laws and their partial products belong to $\mathcal{C}$, as explained after Proposition 2.1.

By channel scaling, at the terminal level

$$\hat{m}_v = X_v, \quad \hat{h}_v \leq X_v \quad (|v| = n). \quad (5.4)$$

Subadditivity in (2.8), channel scaling, and (5.1) give inductively

$$\hat{m}_v \leq X_v \quad (|v| \leq n). \quad (5.5)$$

A vertex with no observed descendants has the uninformative experiment. Finite branches carry no flow, so early leaves and zero-flow subtrees are consistent with these bounds.

5.2 Local and accumulated losses

For $|v| < n$, define the deficit from linear addition by

$$D_v = r \sum_{w \in \text{ch}(v)} \hat{m}_w - \hat{m}_v. \quad (5.6)$$

We first prove that, with a fixed finite C_1 depending only on B, λ ,

$$\hat{h}_v \leq \lambda^3 \sum_w \hat{h}_w + C_1 X_v^2, \quad (5.7)$$

$$0 \leq D_v \leq C_1 X_v \left(\lambda^3 \sum_w \hat{h}_w + X_v^2 \right). \quad (5.8)$$

The coefficient of the propagated third moment is exactly λ^3 ; the constant C_1 multiplies only the new error.

Combine the child branches one at a time. After the edge channel, their individual statistics, now about the spin at v , are

$$b_w = r\hat{m}_w, \quad h'_w = \lambda^3 \hat{h}_w.$$

Every partial product has m at most the sum of its b_w 's, hence at most X_v . Applying (4.2) at each addition gives the partial bound

$$h \leq \sum_w h'_w + CX_v \sum_w b_w \leq H + CX_v^2, \quad H = \sum_w h'_w.$$

The sums on the right may be taken over all children, so this also proves (5.7) after increasing C_1 . Every addition has nonnegative deficit by Proposition 2.1. Its deficit is at most

$$C((H + CX_v^2)b_w + X_v h'_w)$$

by (4.1). Summing and using $\sum_w b_w \leq X_v$ gives $D_v \leq 2CX_v H + C^2 X_v^3$, proving (5.8). The second input at every use of Lemma 4.1 is a single child branch after its channel of parameter λ , as required.

Write $u \succeq v$ when u is a descendant of v , including $u = v$. Iterating (5.7) in (5.8) and using (5.4) gives, for $i = |v| < n$,

$$D_v \leq C_2 X_v \left(\sum_{\substack{u \succeq v \\ |u| < n}} \lambda^{3(|u|-i)} X_u^2 + \sum_{\substack{u \succeq v \\ |u|=n}} \lambda^{3(n-i)} X_u \right). \quad (5.9)$$

Here $C_2 = \max\{C_1, C_1^2\}$ suffices. At each internal descendant the term $C_1 X_u^2$ is inserted once and then propagated with a factor λ^3 per edge; no power of C_1 grows with depth.

Telescoping the deficits (5.6) over the finite tree above L_n now yields

$$\widehat{m}_\rho = r^n \sum_{v \in L_n} X_v - \sum_{|v| < n} r^{|v|} D_v = M - \sum_{|v| < n} r^{|v|} D_v. \quad (5.10)$$

The last equality uses conservation of Θ . It remains to show that this total loss is at most a fixed fraction of M , uniformly in n .

5.3 Averaging the losses along a flow path

Sample a vertex of L_n with probabilities $\Theta(v)/M$ and let $(V_0 = \rho, V_1, \dots, V_n)$ be its ancestors. Conservation gives $\sum_{v \in L_j} \Theta(v) = M$ for every $j \leq n$ and therefore

$$\mathbb{P}(V_j = u) = \frac{\Theta(u)}{M} \quad (u \in L_j). \quad (5.11)$$

This remains true with early leaves: their flow is zero. Every path with positive probability extends to an infinite ray, so (5.2) bounds the sum of its $X_{V_i}^2$.

For $i \leq j$ and $|u| = j$, the identity responsible for the summable path kernel is

$$r^i \lambda^{3(j-i)} X_u = \lambda^{j-i} \Theta(u). \quad (5.12)$$

The unique depth- i ancestor of u is specified once u is chosen. Thus, for $0 \leq i \leq j < n$,

$$\begin{aligned} \sum_{|v|=i} r^i X_v \sum_{\substack{u \succeq v \\ |u|=j}} \lambda^{3(j-i)} X_u^2 \\ = M \lambda^{j-i} \mathbb{E}[X_{V_i} X_{V_j}], \end{aligned} \quad (5.13)$$

whereas the terminal terms satisfy

$$\sum_{|v|=i} r^i X_v \sum_{\substack{u \succeq v \\ |u|=n}} \lambda^{3(n-i)} X_u = M \lambda^{n-i} \mathbb{E} X_{V_i}. \quad (5.14)$$

Combining these identities with (5.9) gives

$$\sum_{|v|<n} r^{|v|} D_v \leq C_2 M \mathbb{E} \left[\sum_{0 \leq i \leq j < n} \lambda^{j-i} X_{V_i} X_{V_j} + \sum_{0 \leq i < n} \lambda^{n-i} X_{V_i} \right]. \quad (5.15)$$

For each fixed difference $k = j - i \geq 0$, Cauchy–Schwarz and (5.2) give, on every path in the support,

$$\sum_{i=0}^{n-1-k} X_{V_i} X_{V_{i+k}} \leq \left(\sum_{i=0}^{n-1-k} X_{V_i}^2 \right)^{1/2} \left(\sum_{i=0}^{n-1-k} X_{V_{i+k}}^2 \right)^{1/2} \leq Q \epsilon^2.$$

Each individual X_{V_i} is at most $\sqrt{Q} \epsilon$. Summing the geometric weights in (5.15), the expectation there is at most

$$\frac{Q \epsilon^2}{1 - \lambda} + \frac{\sqrt{Q} \epsilon}{1 - \lambda}. \quad (5.16)$$

Choose $\epsilon \in (0, 1]$ so small that C_2 times this expression is at most $1/2$. The choice is independent of n . Equations (5.10) and (5.15) now give

$$\hat{m}_\rho \geq M/2 > 0 \quad \text{for every } n \geq 1.$$

The original, undegraded level observation has at least this much quadratic information by data processing. By (2.3),

$$a_n(T, \lambda) \geq M/16 \quad (n \geq 1).$$

Thus $a_\infty(T, \lambda) > 0$, proving the remaining implication of Theorem 1.1. The degraded experiments may vary with n ; the proof requires only the same positive lower bound at each depth, not their convergence or mutual monotonicity.

References

- [1] William Evans, Claire Kenyon, Yuval Peres, and Leonard J. Schulman, *Broadcasting on trees and the Ising model*, Annals of Applied Probability **10** (2000), no. 2, 410–433.
- [2] Svante Janson and Elchanan Mossel, *Robust reconstruction on trees is determined by the second eigenvalue*, Annals of Probability **32** (2004), no. 3B, 2630–2649.
- [3] Harry Kesten and Bernt P. Stigum, *Additional limit theorems for indecomposable multidimensional Galton–Watson processes*, Annals of Mathematical Statistics **37** (1966), no. 6, 1463–1481.
- [4] Marc Mézard and Andrea Montanari, *Reconstruction on trees and spin glass transition*, Journal of Statistical Physics **124** (2006), no. 6, 1317–1350.
- [5] Elchanan Mossel and Yuval Peres, *Information flow on trees*, Annals of Applied Probability **13** (2003), no. 3, 817–844.

- [6] Elchanan Mossel, Allan Sly, and Youngtak Sohn, *Exact phase transitions for stochastic block models and reconstruction on trees*, Annals of Probability **53** (2025), no. 3, 967–1018, Author preprint: [arXiv:2212.03362](#).
- [7] OpenAI, *The Reconstruction Threshold for the Ferromagnetic Four-State Potts Model*, OpenAI Math Release preprint [OAI:The-Reconstruction-Threshold-for-the-Ferromagnetic-Four-State-Potts-Model-October-5-2026](#), 2026.
- [8] Robin Pemantle and Yuval Peres, *The critical Ising model on trees, concave recursions and nonlinear capacity*, Annals of Probability **38** (2010), no. 1, 184–206.
- [9] Allan Sly, *Reconstruction for the Potts model*, Annals of Probability **39** (2011), no. 4, 1365–1406.