Machine-checked rooted-tree majorants for polymer expansions with holes

Lluis Eriksson*

July 2026

Abstract

We present a machine-checked quantitative toolkit for cluster expansions of polymer systems with excluded regions (*holes*), in the discrete cube geometry of Balaban–Dimock renormalization-group analyses. Five Lean 4 theorems, checked against a pinned Mathlib revision, provide: (i) the identity $\sum_T \prod_v c_T(v)! = n!\,C_n$ for child factorials over spanning trees of the complete graph K_{n+1}, with the rooted-tree majorant 4^n as corollary; (ii) a marked-root leaf summation for the tree-graph majorant of an Ursell-type expansion with holes, with the moment constant M paid once at the root and closed leaf ratio $4M^2$ per additional vertex, together with its Catalan-sharpened form $M^{2n+1}C_n$ (a gain of order $n^{3/2}$ in the nth coefficient); and (iii) a target-preserving orderwise bound in which the target union itself survives until the modified-metric exponential is extracted. A certified companion (interval arithmetic, 120-bit precision, committed transcript with a committed reproduction witness) tabulates the smallness gate and encloses every derived constant. Non-vacuity is machine-checked: a concrete hole family satisfying every hypothesis is exhibited in Lean, and the two distinct hypothesis sets among the polymer-facing theorems are both instantiated at it with a strictly positive weight. Each claim is labelled with its verification layer: exact (Lean theorem), certified (interval transcript), or paper-level.

1 Introduction

Cluster (polymer) expansions convert log-partition functions into sums over connected geometric objects; their convergence proofs hinge on combinatorial majorants for tree sums and on geometric smallness gates [1, 2, 3]. When the expansion must respect *holes* — excluded regions that the expansion sees as atomic units, as in renormalization-group treatments of large-field regions [6, 8] — the bookkeeping is notoriously error-prone: root normalizations, child factorials, and hole-separation hypotheses interact in ways that invite wrong constants.[1]

This paper contributes, in machine-checked form:

1. an exact Catalan evaluation of the rooted child-factorial tree sum (Theorem 3.1), of which the classical 4^n tree-shape budget (Theorem 3.2) is a corollary;

*This manuscript and parts of the formal development were produced with AI assistance under the repository's audit protocol. Every theorem below is machine-checked in Lean 4 against a pinned Mathlib; the constant tables of Section 5 are enclosed by committed interval arithmetic; remaining numerics are labelled paper-level where they occur. Repository: `https://github.com/lluiseriksson/THE-ERIKSSON-PROGRAMME`.

[1] A wrong constant of exactly this kind occurred in our own pre-registered prediction of the smallness-gate crossover; the certified run caught it (Remark 5.1).

2. finite, volume-free bounds for the marked-root and fixed-target-union tree-graph majorants of an Ursell-type expansion with holes (Theorems 3.3, 3.4, 3.5), with one smallness gate as the only inequality hypothesis and every constant an explicit function of (d, κ_0);

3. a certified constant table (Section 5) and a machine-checked non-vacuity witness (Section 7).

Statuses are kept separate throughout: **exact** means a Lean theorem in the pinned core with axiom oracle [`propext`, `Classical.choice`, `Quot.sound`]; **certified** means a committed interval-arithmetic transcript with ball-plus-boolean output; **paper-level** means ordinary mathematical prose not formalized here, and is always flagged. The formal development lives in a larger repository whose top-level namespace (`YangMills/`) reflects its original motivation in lattice gauge theory; the present results are self-contained combinatorial and expansion-theoretic tooling, and no reduction to any continuum or gauge-theoretic problem is claimed or used.

2 Setting and definitions

Throughout, $d, L \in \mathbb{N}$ with $L \neq 0$.

Definition 2.1 (cubes and adjacency)**.** *A* cube *is a point of the discrete torus* $(\mathbb{Z}/L)^d$ *(Lean:* `Cube d L`*). Two cubes are* adjacent *(*`cubeAdj`*) iff they are distinct and each coordinate differs by* 0 *or* ± 1*; this is the sup-metric neighbor graph, in which every cube has* $3^d - 1$ *neighbors. A cube set is* connected *when it is walk-connected in this graph.*

Definition 2.2 (hole families and skeletons)**.** *A* hole family $\mathcal{H}$ *(Lean:* `HoleFamily d L`*) is a finite set of finite cube sets — the connected components of the excluded region. A cube set* X respects $\mathcal{H}$ *iff every hole* $H \in \mathcal{H}$ *satisfies* $H \subseteq X$ *or* $H \cap X = \emptyset$*: the expansion sees each hole as an atomic unit [8, eq. (149)]. The* skeleton $\mathrm{sk}(X)$ *is the set of cubes of* X *lying in no hole. The bounds below hypothesize: pairwise disjointness of holes, no adjacency edge between distinct holes, and non-emptiness of each hole.*

Definition 2.3 (polymers)**.** *Given* $\mathcal{H}$ *and an activity* $z\colon \mathcal{P}((\mathbb{Z}/L)^d)_{\mathrm{fin}} \to \mathbb{C}$*, a* polymer *is a nonempty, connected, hole-respecting cube set with nonempty skeleton;* Ω *denotes the (finite) type of polymers (Lean:* `OmegaPolymerType`*). The activity* z *parameterizes the polymer type for compatibility with the upstream expansion; no result below depends on it, and the non-vacuity witness of Section 7 takes* $z \equiv 0$*. Two polymers are* incompatible*, written* $X \nsim X'$*, iff their skeletons intersect. For a tuple* $\mathbf{X} = (X_0, \ldots, X_n) \in \Omega^{n+1}$*, its* incompatibility graph $G(\mathbf{X})$ *has vertex set* $\{0, \ldots, n\}$ *and an edge* $\{i, j\}$ *iff* $X_i \nsim X_j$*.*

Definition 2.4 (modified metric and hole weight)**.** *The* discrete modified metric $\rho(X)$ *(Lean:* `discreteModifiedMetric`*) is the Steiner length of the skeleton inside* X*:*

$$\rho(X) \;=\; \min\bigl\{\, |S| - 1 \;:\; \mathrm{sk}(X) \subseteq S \subseteq X,\; S \text{ connected} \bigr\},$$

with $\rho(X) = 0$ *if no such* S *exists. For a rate* $\kappa \geq 0$ *the* hole weight *is* $W_\kappa(X) = e^{-\kappa\,(\rho(X)+1)}$ *(Lean:* `appendixFHoleExpWeight`*).*

Definition 2.5 (the two tree-graph majorants)**.** *Fix $\mathcal{H}$, an activity z, a nonnegative weight $w\colon \Omega \to \mathbb{R}_{\geq 0}$, and $n \in \mathbb{N}$. Write $\tau(\mathbf{X})$ for the number of spanning trees of $G(\mathbf{X})$. The* marked-root sum *at a cube r is*

$$S^{\#}_{w,r}(n) \;=\; \frac{1}{(n+1)!} \sum_{\substack{\mathbf{X}\in\Omega^{n+1} \\ r\in \mathrm{sk}(X_0)}} \tau(\mathbf{X}) \prod_{j=0}^{n} w(X_j)$$

(Lean: `appendixFHoleHsharpWeightedTreeMarkedRootSum`*), and the* fixed-target sum *at a finite cube set Y is*

$$T^{\#}_{w,Y}(n) \;=\; \frac{1}{(n+1)!} \sum_{\substack{\mathbf{X}\in\Omega^{n+1} \\ \bigcup_j X_j = Y}} \tau(\mathbf{X}) \prod_{j=0}^{n} w(X_j)$$

(Lean: `appendixFHoleHsharpWeightedTreeTerm`*). These are the standard tree-graph majorants of the order-$(n+1)$ coefficients of an Ursell (Mayer) expansion: by tree-graph inequalities of Penrose type [4, 5], Ursell coefficients are dominated by spanning-tree counts of incompatibility graphs, and it is the majorant objects $S^{\#}, T^{\#}$ — not the signed coefficients — that the theorems below bound. The passage from signed coefficients to these majorants lives upstream in the repository's Kotecký–Preiss/Mayer–Ursell development and is not re-proved here. (The word "second" in the Lean names refers to the second cluster-expansion pass of the upstream development; nothing in this paper depends on the first.)*

Definition 2.6 (gate and derived constants)**.** *The* smallness gate *at dimension d and rate $\kappa_0 > 0$ is*

$$G(d,\kappa_0) \;=\; (3^d)^2\, e^{-\kappa_0}\, 2^{3^d+1} \;<\; 1 \qquad (\textit{Lean hypothesis } h_{Cq}), \tag{1}$$

and the derived constants are

$$\mathrm{RS} = (1-G)^{-1}, \qquad M = \max\!\big(1,\ \kappa_0^{-1},\ \mathrm{RS}\big), \qquad \Lambda = 4M^2 \tag{2}$$

(Lean: `appendixFSecondUrsellMomentConstant`*,* `appendixFSecondUrsellLeafConstant`*). All three are explicit functions of (d,κ_0) alone — in particular* volume-free*: no constant below depends on the torus size L, the target Y, or on n beyond the displayed powers. In applications where the order-n term additionally carries an activity factor ε^n, the resulting geometric series has ratio $\Lambda\varepsilon$, which is below one exactly for $\varepsilon < \varepsilon^* := \Lambda^{-1}$ (strict; nothing is claimed at equality). Status: the Lean endpoints prove the coefficient bounds $M\Lambda^n$ and $M^{2n+1}C_n$ (exact); the enclosures of Λ^{-1} are certified; the passage from coefficient bound to a convergent activity series is paper-level algebra conditional on the stated ε^n structure.*

3 Main results

All five statements carry stable, publication-facing names in module `MarkedRootedClosure`, file `PaperTheorems.lean`; each is a direct application of a pinned upstream theorem — no new axiom, no opaque assumption. Here $\mathcal{T}(K_{n+1})$ is the set of spanning trees of the complete graph on the vertices $\{0,\ldots,n\}$, each tree rooted at 0, and $c_T(v)$ is the number of children of v in the rooted tree T.

Theorem 3.1 (exact Catalan value; `normalizedRootedChildFactorialTreeCatalanIdentity`). *For every $n \in \mathbb{N}$,*

$$\frac{n+1}{(n+1)!}\sum_{T\in\mathcal{T}(K_{n+1})}\prod_{v}(c_T(v))! \;=\; C_n, \qquad \textit{equivalently} \qquad \sum_{T}\prod_{v} c_T(v)! \;=\; n!\,C_n,$$

where $C_n = \binom{2n}{n}/(n+1)$ is the nth Catalan number.

Theorem 3.2 (rooted-tree majorant; `normalizedRootedChildFactorialTreeBound`). *For every n, the left-hand side of Theorem 3.1 is at most 4^n.*

Theorem 3.2 follows from Theorem 3.1 via $C_n \le 4^n$; the classical budget overcounts by exactly the factor $4^n/C_n$, which is $\sim \sqrt{\pi}\,n^{3/2}$.

Theorem 3.3 (marked-root leaf summation; `markedRootLeafGeometricBound`). *Let $\mathcal{H}$ satisfy the separation hypotheses of Definition 2.2, let $w \ge 0$ be dominated by the hole weight at doubled rate, $w \le W_{2\kappa_0}$ pointwise, let r be any cube, and assume the gate (1). Then for every n,*

$$(n+1)\cdot S^{\#}_{w,r}(n) \;\le\; M\cdot\Lambda^n.$$

Theorem 3.4 (Catalan-sharpened marked-root bound; `markedRootCatalanBound`). *Under exactly the hypotheses of Theorem 3.3,*

$$(n+1)\cdot S^{\#}_{w,r}(n) \;\le\; M^{2n+1}\cdot C_n.$$

Since $M\Lambda^n = M^{2n+1}4^n$, this improves Theorem 3.3 by the factor $4^n/C_n$: the tree-shape factor is paid at its exact Catalan value.

Theorem 3.5 (target-preserving orderwise bound; `targetPreservingWeightedTreeBound`). *Let $\mathcal{H}$ satisfy the separation hypotheses of Definition 2.2 and assume the gate (1). Let $w, u \ge 0$ with $u \le W_{2\kappa_0}$ pointwise, suppose $w \le W_{\mathrm{rate}}\cdot u$ pointwise for some rate ≥ 0, and let Y be a finite cube set whose skeleton contains some cube r. Then*

$$T^{\#}_{w,Y}(n) \;\le\; M\cdot W_{\mathrm{rate}}(Y)\cdot\Lambda^n.$$

Here "orderwise" means: for each fixed order n separately.

4 Proof architecture

The formal proofs are the authority; this section gives the mathematical mechanism at a level where a specialist can reconstruct the logic. Module pointers are to the repository.

4.1 The Catalan identity

The sum $\sum_T\prod_v c_T(v)!$ counts pairs (T,σ) where T is a spanning tree of K_{n+1} rooted at 0 and σ assigns to each vertex a linear order on its children — i.e., it counts *labelled rooted plane trees* on the vertex set $\{0,\dots,n\}$ with root 0. Plane trees are rigid (a plane structure admits no nontrivial automorphism fixing the root), so this count factors as

$$\#\{\text{plane shapes with } n+1 \text{ nodes}\}\times\#\{\text{labellings of the } n \text{ non-root nodes}\} \;=\; C_n\cdot n!,$$

the first factor by the classical bijection between plane trees and binary trees, formalized in `YangMills/KP/PlaneTreeCatalan.lean` through Mathlib's

```
Tree.treesOfNumNodesEq_card_eq_catalan;
```

the identity then normalizes to Theorem 3.1 (`YangMills/KP/RootedCatalanExact.lean`).

Example 4.1. *$n = 2$: K_3 has three spanning trees. Rooted at 0, the two paths with 0 as an endpoint contribute $\prod_v c_T(v)! = 1$ each, and the path through 0 — the star at 0 when rooted — contributes $2! = 2$; the sum is $4 = 2!\,C_2$.*

4.2 Leaf summation and the $2n + 1$ moment budget

Theorems 3.3 and 3.4 are proved by splitting the tuple sum by the (rooted) shape of a spanning tree and peeling leaves. Three mechanisms produce the constants:

One-pin polymer sums. The basic estimate controls the sum of $W_{2\kappa_0}(X)$ over polymers X whose skeleton contains a fixed cube. A polymer with $\rho(X) + 1 = m$ has a skeleton of at most m cubes carried by a minimal connected Steiner set S; enumerating rooted connected Steiner sets cube by cube costs $(3^d)^2$ per step (degree-squared branching in the adjacency graph), and each cube of S carries a fiber entropy 2^{3^d+1} (the factor $2^{|S|}$ for skeleton-subset choices times $2^{3^d|S|}$ for admissible hole fillings over S). The sum over polymers at Steiner length m is therefore dominated by G^m with G the gate (1), and the total by the geometric sum $\mathrm{RS} = (1-G)^{-1} \le M$. The gate is the only inequality *hypothesis*; the constants are then explicit functions of (d, κ_0), with the κ_0^{-1} component of M paid by the spare κ_0 of decay left over when the doubled-rate weight $W_{2\kappa_0}$ surrenders only W_{κ_0} to the geometric mechanism.

Moment paid once at the root. In the marked-root sum the coordinate X_0 is pinned at r, so the root contributes one one-pin sum: one factor M. Each tree edge propagates a pin: a child polymer must have skeleton intersecting its parent's, so summing the child costs one one-pin sum (M) *after* choosing the pin inside the parent, whose cost is accounted to the parent's own moment. The bookkeeping closes with each of the n non-root vertices paying two factors (M for its own one-pin sum, M for the pin it offers its children) and the root paying one: M^{2n+1} total. The formal count is the fixed-parent kernel lemma

```
appendixFHoleHsharpWeightedTreeMarkedRootFixedParentKernelSum
        _le_prod_childMomentBudget_mul_rootMoment
```

and its consumers in `YangMills/RG/AppendixFSecondUrsellLeafSummation.lean`.

Shape sum. What remains is the sum over rooted tree shapes with child multiplicities — exactly the object of Section 4.1. Paying it at its Catalan value gives Theorem 3.4; paying it at the 4^n budget and absorbing $4^n M^{2n} = \Lambda^n$ gives Theorem 3.3. The $(n + 1)$ prefactor in both statements is the marked-coordinate count: pinning coordinate 0 rather than an arbitrary coordinate costs a factor $n + 1$ against the $(n + 1)!$ normalization.

4.3 Target preservation

In Theorem 3.5 the tuple union is constrained to a fixed target Y. Extracting the metric exponential *before* summing would discard this constraint and with it the geometry of Y. Instead, the hypothesis $w \le W_{\mathrm{rate}} \cdot u$ is applied per polymer, and the extracted factors combine, for each contributing tuple (connected in the incompatibility graph; disconnected tuples contribute zero to $T^\#$), into at least the full-target factor $W_{\mathrm{rate}}(Y)$: the superadditivity $\rho(Y) + 1 \le \sum_i (\rho(X_i) + 1)$ is proved by stitching minimal Steiner sets along a spanning tree of the incompatibility graph (Lean: `omegaClusterUnion_discreteModifiedMetric_add_one_le` `_sum_of_spanningTree`). Only after this factor is secured is the remainder u routed through the marked-root bound at a root $r \in \mathrm{sk}(Y)$. The order of operations is the content of the theorem's name and the reason the constant is volume-free in Y.

5 The certified companion

Script `scripts/c1_admissibility_arb.py` (SHA-256

`f114fc4bcf2193c47603fb595b91b4e3ea62addc3a1bfebe0aa0c147c447f1f9`

`python-flint` 0.9.0, precision 120 bits; transcript committed alongside it, exit 0) encloses G, M, Λ (printed as `Lf` in the transcript), and ε^* with midpoint–radius enclosures plus provable booleans; the gate booleans are tri-state and provable-only (PASS only when $G < 1$ holds for every point of the enclosing ball, FAIL only when $G \geq 1$ does, UNDECIDED otherwise — no UNDECIDED occurs in the table). The run was re-executed from scratch on 2026-07-11 and reproduced the committed transcript byte-identically; the comparison log (SHA-256 of both transcripts) is committed as `scripts/c1_admissibility_rerun_20260711_COMPARISON.txt`. One header note: the script docstring's phrase "certified activity radius" predates the status correction of Definition 2.6, which governs. Selected rows (midpoints shown; radii in the transcript, all below 3×10^{-6} relative):

d	κ_0	gate G	M	Λ	ε^*
2	11	1.3853059	FAIL (provably ≥ 1)		
2	11.5	0.84023048	6.2590162	156.70114	0.0063815746
2	12	0.50962555	2.0392580	16.634292	0.060116775
2	**16**	0.0093341175	1.0094221	4.0757316	**0.24535472**
2	24	3.13×10^{-6}	1.0000031	4.0000251	0.24999843
3	25	2.7177241	FAIL (provably ≥ 1)		
3	26	0.99979481	4873.5020	9.50×10^{7}	1.05×10^{-8}
3	28	0.13530751	1.1564805	5.3497886	0.18692327
3	**33**	0.00091169486	1.0009125	4.0073035	**0.24954436**

The certified crossovers of the gate are

$$\kappa^*(2) \in 11.32592096 \pm 2.72 \times 10^{-10}, \qquad \kappa^*(3) \in 25.99979479 \pm 2.32 \times 10^{-9},$$

and $\varepsilon^* \uparrow 1/4$ as $\kappa_0 \to \infty$ (the pure tree-majorant limit $\Lambda \downarrow 4$). The row $(d, \kappa_0) = (3, 26)$ sits just above $\kappa^*(3)$ and displays the blow-up of the derived constants as $G \to 1^-$: admissible, but with $M \approx 4.9 \times 10^3$ and $\varepsilon^* \approx 1.05 \times 10^{-8}$.

Scope of the certificates: the transcript's band booleans certify the gate at the two registered band points $\kappa_0 = 16$ and $\kappa_0 = 33$; the extension to all larger κ_0 is the elementary monotonicity of $\kappa_0 \mapsto e^{-\kappa_0}$ (paper-level, one line). Admissibility everywhere means the strict gate $G < 1$, as in the Lean hypothesis.

Remark 5.1 (a measured registration failure)**.** *Before the certified run, we committed to the repository (`docs/C1-CHARTER.md`) a predicted crossover for the gate:* 15.72 *(d=2) and* 32.60 *(d=3), from a hand-assembled expression carrying* $(3^d)^4$ *where the gate (1) carries* $(3^d)^2$*. The certified run measured the discrepancy (transcript, "crossover audit":* REGISTERED PREDICTION

DOES NOT RE-DERIVE); the pre-committed admissibility bands $\kappa_0 \geq 16$ (d=2) and $\kappa_0 \geq 33$ (d=3) remain true — sufficient, not sharp — and only certified values appear in this paper. We record the failure as evidence for the correction protocol: constants are computed, never assembled by hand.

6 What the Catalan sharpening buys

The gain of Theorem 3.4 over Theorem 3.3 at order n is the integer ratio $4^n/C_n$ (computed in integer arithmetic, displayed rounded; paper-level, independently recomputable):

n	C_n	4^n	$4^n/C_n$
1	1	4	4.00
2	2	16	8.00
5	42	1,024	24.38
10	16,796	1,048,576	62.43
20	6,564,120,420	$\approx 1.10 \times 10^{12}$	167.50
30	$\approx 3.81 \times 10^{15}$	$\approx 1.15 \times 10^{18}$	302.21

Being $\sim \sqrt{\pi}\, n^{3/2}$, the improvement is polynomial: it does *not* move the exponential threshold. In an application with per-vertex activity ε, both bounds give the same radius requirement $4M^2\varepsilon < 1$ asymptotically; what changes is every finite-order coefficient, hence every truncation-error estimate at fixed order. Concretely, at $(d, \kappa_0) = (2, 16)$ (certified row: $M = 1.0094221$, ball radius $\leq 4 \times 10^{-8}$), truncating so that order $n = 10$ is the last term retained, the first omitted coefficient is smaller by the factor $4^{11}/C_{11} \approx 71.4$ under the Catalan form — almost two decimal digits of truncation error at no analytic cost. For the near-critical certified row $(3, 26)$, where $M \approx 4.9 \times 10^3$ inflates every coefficient through M^{2n+1}, the same polynomial factor directly divides the order-n term of any error budget. These consequences are paper-level arithmetic on top of the machine-checked theorems.

7 Non-vacuity, machine-checked

A characteristic failure mode of formalized bounds — one this programme has itself committed and repaired in earlier work — is hypotheses that no concrete geometry satisfies. Module `MarkedRootedClosure/NonVacuity.lean` therefore exhibits — as instance terms, not existentials — the hole family $\mathcal{H}^w$ in $d = 2$, $L = 8$ with two singleton holes at sup-distance 4, and proves by kernel decision (`decide`) that the holes are disjoint, edge-separated, and nonempty; the gate (1) at $\kappa_0 = 16$ is proved in Lean via $e^{16} > 2.7^{16} > 82944$ — the Lean-exact counterpart of the certified row $G \approx 0.00933$. The two distinct hypothesis sets among Theorems 3.3–3.5 are then both instantiated at $\mathcal{H}^w$ with the *strictly positive* weight $W_{32} = e^{-32(\rho+1)}$ (no zero-weight degeneracy; Theorem 3.4 shares its hypotheses verbatim with Theorem 3.3, so its satisfiability follows):

`nonvacuous_markedRootLeafGeometricBound`, `nonvacuous_targetPreservingWeightedTreeBound`.

What this witness proves: both hypothesis sets among Theorems 3.3–3.5 are satisfiable by a concrete geometry with positive weights, so the theorems have non-trivial instances and the bounds are statements about nonzero quantities. What it does not prove: that the instantiated bounds are sharp, or that the witness geometry models any physical large-field configuration; it is a consistency certificate, not an application.

8 Related work and novelty

Cluster expansions. The abstract polymer framework and its convergence criteria go back to Kotecký–Preiss [1], with the modern sharp form of the criterion due to Fernández–Procacci [2]; pedagogical accounts include Brydges [3]. Tree-graph inequalities dominating Ursell coefficients by spanning-tree counts originate with Penrose [4]; the family of tree and forest formulas is surveyed by Abdesselam–Rivasseau [5]. The upstream repository formalizes a Kotecký–Preiss criterion and a Mayer–Ursell expansion, and the majorants $S^{\#}, T^{\#}$ of Definition 2.5 are the standard objects of this tradition.

Holes and large fields. The specific geometry — hole-respecting polymers, skeletons, and the modified (Steiner) metric — is that of renormalization-group analyses of lattice gauge theory in the Balaban school [6], in the concrete form of Dimock's expository series [7, 8, 9]; the hole-component families of Definition 2.2 follow [8, §3.1.2]. We are not aware of published *machine-checked* bounds for expansions in this geometry prior to the present programme.

The Catalan identity. The count of labelled rooted plane trees underlying Theorem 3.1 is classical combinatorics at the level of [10], and we make no novelty claim for the enumeration itself. The identity in this normalized child-factorial form, and its machine-checked proof, were first announced in the author's own preprint [11] with its companion repository; the present paper subsumes that announcement into the hole-respecting expansion toolkit (Theorem 3.4 and the certified constants). We are not aware of an *independent* prior source for the identity in exactly this form, for its use to replace the 4^n shape budget inside an Ursell tree majorant with holes, or for a formalization of either; readers aware of one are invited to point us to it.

What is new here. To our knowledge: (i) the formalized exact evaluation (Theorem 3.1) and its formalized corollary chain into hole-respecting leaf summation (Theorem 3.4); (ii) the machine-checked volume-free constant structure M^{2n+1} with the single-gate discipline (1); (iii) the machine-checked non-vacuity witness; and (iv) the certified-constants companion with committed pre-registered bands and a documented registration failure (Remark 5.1). The mathematical ingredients are classical; the contribution is the machine-checked integration.

Formalization context. The development builds on Mathlib [13]; the upstream repository's formalized cluster expansion (sharp Kotecký–Preiss criterion, Mayer–Ursell identity, and a volume-uniform Wilson-loop area law) is described in the author's preprint [12]. The Catalan/plane-tree layer contributes upstream API (a `PlaneTree` companion prepared for separate submission to Mathlib).

9 Reproducibility

Toolchain: Lean `leanprover/lean4:v4.29.0-rc6`; Mathlib pinned to commit

`07642720480157414db592fa85b626dafb71355b`

(lakefile and manifest agree); `python-flint` 0.9.0 on CPython 3.12.6; `tectonic` 0.15.0.

Theorem-to-artifact map (statuses as in Section 3; all Lean files in module `MarkedRootedClosure` unless noted):

Result	Lean name	File	Status
Thm. 3.1	`...TreeCatalanIdentity`	`PaperTheorems.lean`	exact
Thm. 3.2	`...TreeBound`	`PaperTheorems.lean`	exact
Thm. 3.3	`markedRootLeafGeometricBound`	`PaperTheorems.lean`	exact
Thm. 3.4	`markedRootCatalanBound`	`PaperTheorems.lean`	exact
Thm. 3.5	`targetPreservingWeightedTreeBound`	`PaperTheorems.lean`	exact
§7 witnesses	`nonvacuous_*`	`NonVacuity.lean`	exact
§5 table	—	`scripts/c1_admissibility_*`	certified
§6 ratios	—	integer arithmetic in text	paper-level
§6 consequences	—	—	paper-level

From a clean clone of the repository at the release revision:

```
lake exe cache get           # Mathlib oleans
lake build MarkedRootedClosure
lake env lean MarkedRootedClosure/Oracle.lean   # axiom oracle
python scripts/c1_admissibility_arb.py          # certified table
tectonic -X compile \
  papers/c1-rooted-tree-majorants/c1_rooted_tree_majorants.tex
```

Release facts (recorded from the run logs): `lake build MarkedRootedClosure` completed with exit code 0 in **8244 jobs** at the Lean-integration commit

`4f95a5ad8154deef0ef38716073810072332fd37`

(release tag `c1-v1.0` on the final manuscript commit). The axiom oracle (`lake env lean MarkedRootedClosure/Oracle.lean`; log committed as `scripts/c1_oracle_output.txt` at the same commit) reports, for each of the seven theorems — the five of Section 3 and the two non-vacuity instantiations of Section 7 — exactly

`[propext, Classical.choice, Quot.sound]`,

the three standard classical axioms of Mathlib; no project axiom, no `sorry`. The raw build transcripts (all three recorded runs, including the two measured failures on the way to green), the canonical-hash manifest distinguishing the integration commit, the manuscript commit, and the tag, and a dual-hash (LF blob / CRLF worktree) statement of the reproduction witness are committed under `scripts/c1_lake_build_*` and `papers/c1-rooted-tree-majorants/RELEASE-MANIFEST.md` in the documentation patch `c1-v1.0.1`.

Pin-drift audit (measured): relative to the module's original `archive/UPSTREAM.lock` (upstream commit `4e45246a…`), the blob of `KP/RootedLeafSummation.lean` is byte-identical at the build revision, while `RG/AppendixFSecondUrsellLeafSummation.lean` drifted (+144/−15 lines: the `treeShape` refactor and the Catalan route; the geometric theorem's statement is byte-unchanged). The Catalan chain (`KP/RootedCatalanExact.lean`, `KP/PlaneTreeCatalan.lean` and their dependencies) postdates the lock and is therefore not audited against it; for those files, and for everything else, the authority is the build itself: the theorems are compiled against the release revision, not the lock.

10 Conclusion

Five machine-checked theorems give the exact combinatorial value and the volume-free analytic envelope of rooted tree-graph majorants for polymer expansions with holes, with certified constants and a machine-checked consistency witness. The limitations are stated where they arise: the signed-coefficient-to-majorant passage is upstream and the activity-series algebra is paper-level; the Catalan gain is polynomial and moves no radius; the witness certifies satisfiability, not sharpness. Reasonable next uses are (a) consuming Theorem 3.4 inside the upstream programme wherever the 4^n budget is currently paid, and (b) upstreaming the plane-tree API.

References

[1] R. Kotecký and D. Preiss, *Cluster expansion for abstract polymer models*, Comm. Math. Phys. **103** (1986), 491–498. doi:10.1007/BF01211762.

[2] R. Fernández and A. Procacci, *Cluster expansion for abstract polymer models. New bounds from an old approach*, Comm. Math. Phys. **274** (2007), 123–140. doi:10.1007/s00220-007-0279-2, arXiv:math-ph/0605041.

[3] D. C. Brydges, *A short course on cluster expansions*, in: Phénomènes critiques, systèmes aléatoires, théories de jauge (Les Houches 1984), North-Holland, Amsterdam, 1986, pp. 129–183.

[4] O. Penrose, *Convergence of fugacity expansions for classical systems*, in: Statistical Mechanics: Foundations and Applications (T. A. Bak, ed.), Benjamin, New York, 1967, pp. 101–109.

[5] A. Abdesselam and V. Rivasseau, *Trees, forests and jungles: a botanical garden for cluster expansions*, in: Constructive Physics (V. Rivasseau, ed.), Lecture Notes in Physics **446**, Springer, Berlin, 1995, pp. 7–36. doi:10.1007/3-540-59190-7_20.

[6] T. Bałaban, *Renormalization group approach to lattice gauge field theories. I. Generation of effective actions in a small field approximation and a coupling constant renormalization in four dimensions*, Comm. Math. Phys. **109** (1987), 249–301.

[7] J. Dimock, *The renormalization group according to Balaban, I. Small fields*, Rev. Math. Phys. **25** (2013), 1330010. arXiv:1108.1335.

[8] J. Dimock, *The renormalization group according to Balaban, II. Large fields*, J. Math. Phys. **54** (2013), 092301. arXiv:1212.5562.

[9] J. Dimock, *The renormalization group according to Balaban, III. Convergence*, Ann. Henri Poincaré **15** (2014), 2133–2175. arXiv:1304.0705.

[10] R. P. Stanley, *Catalan Numbers*, Cambridge University Press, New York, 2015.

[11] L. Eriksson, *A machine-verified bijective proof of the rooted child-factorial Catalan identity over spanning trees of the complete graph*, preprint, ai.viXra.org:2607.0001, July 2026.

[12] L. Eriksson, *A machine-checked volume-uniform Wilson-loop area law via a formalized cluster expansion*, preprint, ai.viXra.org:2607.0005, July 2026.

[13] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. on Certified Programs and Proofs (CPP 2020), ACM, 2020, pp. 367–381. doi:10.1145/3372885.3373824.