A counterexample to Gao and Gu's 2-adic conjecture for weighted Morse link numbers

Qianghao Wu

November 26, 2026

Abstract

Let $L_n^{(k)}$ be the weighted Catalan number with weight $b(x) = (2x+1)^{2k}$, and let C_n be the Catalan number. The case $k = 1$ counts the combinatorial types of Morse links. Gao and Gu conjectured that for every positive integer k there are a 2-adic integer α_k and an integer $c_k \geq 0$ such that $\xi_2(L_n^{(k)} - C_n) = s_2(n) + \xi_2(n - \alpha_k) + c_k$ for all $n \geq 2$. This generalizes a conjecture of Postnikov for $k = 1$. We show that the conjecture is false for $k = 2$. The exact values $\xi_2(L_n^{(2)} - C_n) = 4$, 13 and 11 at $n = 2$, 3 and 19 are incompatible with every choice of α_2 and c_2. Lean 4 with Mathlib checks the theorem from the definition of weighted Catalan numbers by Dyck paths, with only the standard axioms. In our data, the failure comes from the index $n = 3$ alone. The version of the conjecture for $n \geq 4$ remains open: it fits our data for $k \leq 24$ and $n \leq 200$. In the same data, the version for $n \geq 2$ fails exactly for $k \equiv 2 \pmod 4$. These two statements are computed, not proved. This work has not yet been reviewed by experts. The computations, the proofs and the text were produced by AI models (from Anthropic and OpenAI), working under the author's direction.

1 Introduction

A *Dyck path* of semilength n is a lattice path from $(0, 0)$ to $(2n, 0)$ with steps $(1, 1)$ (up) and $(1, -1)$ (down) that never goes below the x-axis. Let $b : \mathbb{Z}_{\geq 0} \to \mathbb{Z}$. Gao and Gu [1, Definition 1.1] give each up step the weight $b(y)$, where y is the height at the start of the step, and each down step the weight 1. The weight of a path is the product of the weights of its steps. The *weighted Catalan number* C_n^b is the sum of the weights of all Dyck paths of semilength n. For $b \equiv 1$ this is the Catalan number C_n. For $b(x) = (2x+1)^2$, C_n^b is the number L_n of combinatorial types of Morse links of order n [6, 1].

For an integer $m \neq 0$, $\xi_2(m)$ is the exponent of the highest power of 2 that divides m, and $s_2(n)$ is the sum of the binary digits of n [1, p. 2]. For a 2-adic integer $x \neq 0$, $\xi_2(x)$ is the usual 2-adic valuation.

History. Gao and Gu [1, Section 1] summarize the work on the 2-adic valuations of weighted Catalan numbers. It is well known that $\xi_2(C_n) = s_2(n+1) - 1$. Postnikov and Sagan gave a sufficient condition on b for $\xi_2(C_n^b) = \xi_2(C_n)$ for all n [1, Theorem 1.3]. For polynomial b, Konvalinka conjectured a necessary and sufficient condition and An proved it; Konvalinka also extended the result of Postnikov and Sagan to weighted q-Catalan numbers (all as described in [1, Section 1]). Gao and Gu relaxed the conditions of Postnikov and Sagan [1, Theorem 2.1], extended the method to weighted q-Catalan numbers, and proved some periodicity conjectures of Postnikov on Morse link numbers modulo m [1, Section 4]. They also stated a number of conjectures. Inagaki and Pramatarova [3] recently gave infinitely many counterexamples to one of them, Conjecture 2.14

of [1], which is about when $\xi_2(C_n^b) = \xi_2(C_n)$. The same authors study multidimensional weighted Catalan numbers and their periodicity [4, 5].

The conjecture. Postnikov [6, Conjecture 3.4] stated a 2-adic pattern for $L_n - C_n$. Gao and Gu rephrase it as follows [1, Section 4.3]:

> Conjecture 4.11. There exists a 2-adic integer $\alpha = \ldots 010111_2$ such that $\xi_2(L_n - C_n) = s_2(n) + \xi_2(n - \alpha) + 2$ for all $n \geq 2$.

They then state [1, Conjecture 4.12], quoted word for word from arXiv:1908.03914v1:

> Conjecture 4.12. Let k be a positive integer and $L_n^{(k)}$ denote the weighted Catalan numbers with weight $b(x) = (2x+1)^{2k}$. There exists a 2-adic integer α_k and nonnegative integer c_k such that
>
> $$\xi_2(L_n^{(k)} - C_n) = s_2(n) + \xi_2(n - \alpha_k) + c_k$$
>
> for all $n \geq 2$.

So $L_n^{(1)} = L_n$, and Conjecture 4.11 is the case $k = 1$. The paper also appeared in the Journal of Number Theory [2]. We did not read the journal text. We quote the arXiv version, which is the only arXiv version.

Result. Conjecture 4.12 is false for $k = 2$ (Theorem 1). The proof uses three exact values. In our data, the obstruction comes from $n = 3$ only: if the range is changed to $n \geq 4$, the data for $k \leq 24$ and $n \leq 200$ fit the conjecture (Section 5). So the result suggests a change of the range from $n \geq 2$ to $n \geq 4$ for some k. The changed statement is open. It says nothing about the 2-adic pattern for large n, and it does not touch Postnikov's case $k = 1$.

Prior work on this question. We searched the two public collections of AI-produced results that we know of, GitHub and Zenodo archives, the citations of [1] in Semantic Scholar and OpenAlex, and the web. We found no earlier settlement of Conjecture 4.12. The work of Inagaki and Pramatarova [3] treats only Conjecture 2.14. Their weights $b_t(x) = (4t + 1)x^3 - x^2 + 1$ are not of the form $(2x + 1)^{2k}$, so their construction does not apply to Conjecture 4.12. Zhu and Huang [7] study only the case $k = 1$, Conjecture 4.11. In their computations the 2-adic integer α is replaced by the integer 25431; Postnikov's digits of α give $\alpha \equiv 549719 \pmod{2^{20}}$ [6, Conjecture 3.4].

2 The result

Theorem 1. *Let $b(x) = (2x + 1)^4$. There is no 2-adic integer α and no integer $c \geq 0$ such that $\xi_2(L_n^{(2)} - C_n) = s_2(n) + \xi_2(n - \alpha) + c$ for all $n \geq 2$. Hence Conjecture 4.12 of [1] is false.*

Table 1 gives the values that the proof uses. Write $d(n) = \xi_2(L_n^{(2)} - C_n) - s_2(n)$. If the conjecture holds for $k = 2$, then $d(n) = \xi_2(n - \alpha) + c$ for all $n \geq 2$.

For $n = 2$: the Dyck paths are $UUDD$ and $UDUD$, with weights $b(0)b(1) = 81$ and $b(0)^2 = 1$. So $L_2^{(2)} = 82$ and $C_2 = 2$. For $n = 3$, the five Dyck paths give

$$L_3^{(2)} = b(0)^3 + 2\,b(0)^2 b(1) + b(0)b(1)^2 + b(0)b(1)b(2) = 1 + 162 + 6561 + 50625 = 57349,$$

and $C_3 = 5$. The values at $n = 4$ and $n = 19$ come from the exact computation in Section 4. At $n = 19$, $L_{19}^{(2)}$ is the 90-digit integer

n	$L_n^{(2)} - C_n$	$\xi_2(L_n^{(2)} - C_n)$	$s_2(n)$	$d(n)$
2	$80 = 2^4 \cdot 5$	4	1	3
3	$57344 = 2^{13} \cdot 7$	13	2	11
4	$2^4 \cdot 10127819$	4	1	3
19	$2^{11} \cdot q$, q odd, 87 digits	11	3	8

Table 1: Exact values for $k = 2$. The row $n = 4$ is not needed for the proof. We list it because the first proof that we found used it.

```
26202022122521055679991677002458745285538308258168778866450164384695813822405545969142677 4
```

(one number in two lines), and $C_{19} = 1767263190$.

3 Proof

Lemma 2. *Let x be a nonzero 2-adic integer and $0 \le j \le \xi_2(x)$. Then 2^j divides x in $\mathbb{Z}_2$.*

Proof. $x = 2^{\xi_2(x)}u$ with u a unit of $\mathbb{Z}_2$. □

Proof of Theorem 1. Suppose that $\alpha \in \mathbb{Z}_2$ and $c \ge 0$ satisfy the identity for all $n \ge 2$. Each $n - \alpha$ is nonzero, because the left side is finite. By Table 1:

$$\xi_2(2-\alpha) + c = 3, \qquad \xi_2(3-\alpha) + c = 11, \qquad \xi_2(19-\alpha) + c = 8.$$

The first equation gives $c \le 3$, so $j = 8 - c \ge 5$. The second and third give $\xi_2(3-\alpha) = 11 - c \ge j$ and $\xi_2(19-\alpha) = 8 - c = j$. By Lemma 2, 2^j divides $3-\alpha$ and $19-\alpha$, so 2^j divides their difference 16 in $\mathbb{Z}_2$. Since $j \ge 5$, 2^5 divides 16 in $\mathbb{Z}_2$, and so in $\mathbb{Z}$. This is false. □

Remark 3. The argument shows more: the values at $n = 2, 3$ force $c = 3$ and $\alpha \equiv 3 + 2^8 \pmod{2^9}$. Then $\xi_2(19-\alpha) = \xi_2(16) = 4$, while the identity needs $8 - 3 = 5$. Every other $n \le 18$ agrees with these α and c. So $n = 19$ is the first index that shows the contradiction. If $n = 3$ is removed, no contradiction appears for $k \le 24$ and $n \le 200$ (Section 5).

4 How the result was checked

Lean. The file `lean/GaoGu412.lean` states and proves Theorem 1 in Lean 4 with Mathlib [8, 9] (Lean and Mathlib v4.36.0). The statement uses these definitions:

- A path is a list of Booleans: `true` is an up step, `false` is a down step. The height after a list is the number of up steps minus the number of down steps.
- `IsDyckPath n s`: s has length $2n$, the height of every prefix of s is ≥ 0, and the height of s is 0.
- `weight b s`: the product over the up steps of b at the height of the prefix before the step.
- `weightedCatalan b n`: the sum of `weight b s` over all s of length $2n$ with `IsDyckPath n s`. Then `L k n` uses $b(x) = (2x+1)^{2k}$ and `C n` uses $b \equiv 1$.

- `Conjecture412 k`: there are $\alpha \in \mathbb{Z}_2$ (`ℤ_[2]`) and $c \in \mathbb{N}$ such that for all $n \geq 2$, $n - \alpha \neq 0$ and `padicValInt 2 (L k n - C n)` equals the binary digit sum of n plus `(n - α).valuation` plus c.

The source statement implies the condition $n - \alpha \neq 0$: every weight is at least 1 and $b(1) > 1$, so $L_n^{(k)} - C_n > 0$ for $n \geq 2$; the left side is finite, so $\xi_2(n-\alpha)$ must be finite. So the source conjecture for k implies `Conjecture412 k`, and a disproof of `Conjecture412 2` disproves the source. Mathlib's `PadicInt.valuation` gives the value 0 at 0, which is why the condition is written out. The main theorems are

```
theorem conjecture_4_12_false_k2 : ¬ Conjecture412 2
theorem conjecture_4_12_false : ¬ ∀ k : ℕ, 0 < k → Conjecture412 k
```

The proof shows by induction that `weightedCatalan` satisfies the height recursion (1) below. It then evaluates the recursion with a list of 40 heights, and checks the valuations at $n = 2, 3, 19$ of Table 1 by kernel reduction (`decide +kernel`). It does not use `native_decide`. The command `#print axioms` prints `[propext, Classical.choice, Quot.sound]` for both theorems. The file `lean/Challenge.lean` repeats the definitions and the two statements, and `lake comparator` confirms that the proof file proves exactly these statements. On one core of an Apple M1 Pro, the proof file compiles in about 3 minutes.

Python. Two exact checkers in Python use only integers. The first computes C_n^b in three ways: brute-force enumeration of Dyck paths for $n \leq 10$, the height recursion

$$f_0(h) = [h = 0], \qquad f_{m+1}(h) = b(h)\, f_m(h+1) + f_m(h-1) \ (h \geq 0), \qquad C_n^b = f_{2n}(0), \qquad (1)$$

where $f_m(h)$ is the weighted number of paths with m steps that start at height h, stay at height ≥ 0 and end at height 0 (with $f_m(-1) = 0$), and the first-return recursion that corresponds to the continued fraction of [1, Proposition 1.2]. The three agree for $n \leq 40$. A second checker, written separately, gives the same values of $d(n)$. The certificate script compares the values with `data/expected.json`. It also tests all residues of α modulo 2^{12} and all $c \leq 3$ at $n = 2, 3, 19$, and finds no solution; this test is exhaustive, because the identity needs $\xi_2(n - \alpha) \leq 11$ at these n. As a control, the same test for $k = 1$ and $c = 2$ on $2 \leq n \leq 40$ accepts the residues of α that end in 010111_2, as Conjecture 4.11 predicts.

5 Computations for other k

For $1 \leq k \leq 24$ and $2 \leq n \leq 200$ we computed $d_k(n) = \xi_2(L_n^{(k)} - C_n) - s_2(n)$ exactly and tested whether some α_k and c_k fit. The test uses $\xi_2(n - \alpha) = m$ if and only if $\alpha \equiv n + 2^m \pmod{2^{m+1}}$. It is exact for the finite range. The data show the following. These are observations for the stated range only. We do not prove them.

- For $n \geq 2$, no (α_k, c_k) fits exactly when $k \equiv 2 \pmod 4$, that is, $k = 2, 6, 10, 14, 18, 22$. For the other 18 values of k, a fit exists.
- For $n \geq 4$, a fit exists for every $k \leq 24$, with exactly one value of c_k. For $k = 2$, the fit is $c_2 = 3$ and $\alpha_2 \equiv 1010110011_2 \pmod{2^{10}}$.
- In every case with a fit, $c_k = \xi_2(k) + 2$.
- For $k = 1$, the data fit with $c_1 = 2$ and $\alpha_1 \equiv 101010111_2 \pmod{2^9}$. This agrees with the digits $\ldots 010111_2$ of Conjecture 4.11.

6 Limits

- Proved (in Lean and by hand): Conjecture 4.12 as stated, with "for all $n \geq 2$", is false, because it fails for $k = 2$.
- In our data, the failure comes from the single index $n = 3$. A reader can regard the result as a correction of the range for some k. The conjecture with "for all $n \geq 4$" remains open. Our data for $k \leq 24$ and $n \leq 200$ support it.
- Not proved: that the conjecture fails for every $k \equiv 2 \pmod 4$, and that $c_k = \xi_2(k) + 2$. Both are observations for $k \leq 24$ and $n \leq 200$.
- Our result does not apply to Conjecture 4.11 (the case $k = 1$). Our data fit it for $2 \leq n \leq 200$.
- We read the arXiv version [1] of the source. We did not read the text of the journal version [2]. If the journal version changes the range of Conjecture 4.12, then our result applies to the arXiv statement only.
- The Python programs are not formally verified. The Lean proof does not depend on them.

References

[1] Y. Gao and A. Gu, Arithmetic of weighted Catalan numbers, arXiv:1908.03914v1 (2019). https://arxiv.org/abs/1908.03914

[2] Y. Gao and A. Gu, Arithmetic of weighted Catalan numbers, J. Number Theory 226 (2021), 213–242. doi:10.1016/j.jnt.2021.03.007

[3] R. Inagaki and D. Pramatarova, Counterexamples to a conjecture of Gao and Gu on weighted Catalan numbers, arXiv:2611.26953v1 (2026). https://arxiv.org/abs/2611.26953

[4] R. Inagaki and D. Pramatarova, On weighted and bounded multidimensional Catalan numbers, arXiv:2510.14956v1 (2025). https://arxiv.org/abs/2510.14956

[5] R. Inagaki and D. Pramatarova, On semisymmetric height and a multidimensional generalization of weighted Catalan numbers, arXiv:2604.04900v1 (2026). https://arxiv.org/abs/2604.04900

[6] A. Postnikov, Counting Morse curves and links, preprint. https://math.mit.edu/~apost/papers/morse-brief.pdf

[7] X. Zhu and X. Huang, Computational code and data for the partial resolution of the Gao-Gu Conjecture 4.11 (weighted Catalan 2-adic valuation), Zenodo (2026). doi:10.5281/zenodo.24046429

[8] The Lean 4 theorem prover. https://github.com/leanprover/lean4

[9] The Mathlib Community, Mathlib. https://github.com/leanprover-community/mathlib4