A deficit proof of the triangle-free σ_t extremum

DRAFT • 1 November 2026 • MathDB 365710
AI-assisted research manuscript; no human authorship or affiliation is asserted. Prepared with OpenAI Codex from the frozen proof and review dossier. Zero accepted credit; root alone accepts. Campaign status remains 6/20 beyond 39.

Abstract. For each integer $n \geq 2$, we prove that the maximum of the sum of squared degree differences over all unordered vertex pairs of a connected triangle-free graph is attained by a complete bipartite graph. Moreover, every maximizer is complete bipartite. The proof strengthens the classical triangle-free Zagreb inequality by a connected non-complete-bipartite deficit of at least $n - 2$. An induced four-vertex path supplies an injection into induced one-edge triples; a discrete polynomial bracketing argument controls the rounding loss. A Lean certificate proves the full finite-order assertion, including attainment and classification. The deficit identity is credited to Zhou (2004). Novelty is bounded by the disclosed literature review; no worldwide-priority claim is made.

1. Definitions and exact question. All graphs here are finite, simple and undirected. Write $n = |V(G)|$, $m = |E(G)|$, $d(v) = |N(v)|$ and $M_1 = \sum_v d(v)^2$. Following equation (4) of [FDKŠ], define

$$\sigma_t(G) = \sum_{\{u,v\} \subseteq V(G),\, u \neq v} (d(u) - d(v))^2.$$

Each unordered pair is included once, whether or not it is an edge. Thus σ_t is n^2 times the population variance of the degrees. It is distinct from both the edge-based squared-difference index and total absolute irregularity.

For $1 \leq a < n$ put $F(n,a) = a(n-a)(n-2a)^2$ and $B(n) = \max_{1 \leq a < n} F(n,a)$. The finite domain is nonempty when $n \geq 2$. An “extremal complete bipartite graph” means $K_{a,n-a}$ with $F(n,a) = B(n)$; it does not mean every complete bipartite graph. Proposition 11 of [FDKŠ], for $n \geq 7$, reduces the smaller part to the floor or ceiling of $n(2-\sqrt{2})/4$. Comparing their objective values is still necessary; neither “nearest integer always wins” nor “both choices win” is asserted.

Conjecture (literal quantified form). For every integer $n \geq 2$ and every connected triangle-free graph G on n vertices, $\sigma_t(G) \leq B(n)$, with equality attained by an extremal complete bipartite graph. This is arXiv:2411.04881v1 Conjecture 24 and the corresponding journal Conjecture 6.1 [FDKŠ-J]. MathDB 365710 records this same class and invariant. We treat the singleton separately: $\sigma_t = 0$, without asserting two nonempty parts.

mathdb:2611.00007v1 03 Nov 2026

2. Theorem and the classical deficit identity

Theorem. For every $n \geq 2$, the maximum of σ_t over connected triangle-free n-vertex graphs is $B(n)$. A graph in this class attains that maximum if and only if it is complete bipartite with positive part sizes a and $n-a$ satisfying $F(n,a) = B(n)$. In particular every maximizing graph is complete bipartite, up to its actual vertex partition; no labeling assumption is imposed.

Lemma 1 (variance and deficit). Set $D = nm - M_1$. For any graph the expansion of the all-pair sum and the handshake lemma give

$$\sigma_t = nM_1 - (\sum_v d(v))^2 = nM_1 - 4m^2$$
$$= m(n^2 - 4m) - nD. \quad (1)$$

Indeed, in expanding the squared differences each $d(v)^2$ occurs $n-1$ times, while the cross terms recover the remaining terms in $(\sum d(v))^2$. This is the usual pair-variance identity, and is Proposition 1 of [FDKŠ].

If G is triangle-free, D is the number of unordered three-vertex sets inducing exactly one edge. To see this, fix an edge uv. Its endpoint neighborhoods are disjoint: a common neighbor would complete a triangle. There are exactly $n - d(u) - d(v)$ vertices adjacent to neither endpoint. Summing over edges counts each induced one-edge triple exactly once, at its unique edge. Since $\sum_{uv \in E}(d(u)+d(v)) = \sum_v d(v)^2$, this sum is $nm - M_1 = D$. In particular D is a nonnegative integer.

Attribution. Zhou [Z], Theorem 5 and equation (1), already establish the triangle-free inequality $M_1 \leq nm$ and imply precisely this deficit identity. In his notation f_i counts edges among nonneighbors of vertex i; the triangle-free relation $d_i + e_i + f_i = m$ yields $nm - M_1 = \sum_i f_i$. We claim no novelty for that identity or its nonnegativity. The connected strict gap in Section 4 is the additional ingredient used here.

Why the classical equality case is insufficient. Knowing $M_1 \leq nm$ with equality only for complete bipartite graphs gives at most $D \geq 1$ for another graph. Completing the square gives $\sigma_t \leq n^4/16 - nD$. At $n = 10$, the relaxation $D \geq 1$ yields $625 - 10 = 615$, whereas $B(10) = 576$ (parts 1/9 or 2/8). The relaxed estimate therefore fails to dominate the discrete complete-bipartite optimum. This is an arithmetic relaxation, not a claim that a graph realizing m or D in the relaxation exists. A stronger structural loss is needed.

In the Lean implementation all degree differences and the subtraction defining D take place in $\mathbb{Z}$. The directed-edge, or dart, count of vertices adjacent to neither endpoint equals $2D$; dividing the two orientations out recovers the unordered interpretation. This avoids silently introducing a factor of two into (1).

mathdb:2611.00007v1 03 Nov 2026

3. Structure: an induced path from a boundary edge

Lemma 2. A connected triangle-free graph which is not complete bipartite contains an induced path on four vertices.

Proof, in the walk form used by Lean. Suppose there is no induced P_4. First, if $u-v-w$ is a two-edge walk, u and w have identical neighborhoods. If $u = w$ this is immediate. Otherwise, suppose x is adjacent to u but not w. The vertices x,u,v,w are distinct: $x \neq v$ follows because v is adjacent to w; $x \neq w$ follows from triangle-freeness of u,v,w; the other inequalities follow from edges and $u \neq w$. The chords $u-w$ and $x-v$ are forbidden by triangle-freeness; $x-w$ is absent by choice. Hence $x-u-v-w$ is an induced P_4, a contradiction. Reversing u and w proves equality of their neighborhoods.

Fix a vertex v and let $A = N(v)$, $B = V(G)$ minus A. Call y a twin of v when $N(y) = N(v)$. Along any walk starting at v, maintain the assertion that its current endpoint y either lies in A or is a twin of v. The assertion holds initially because v is its own twin. For the next edge $y-z$, if $y \in A$ then $v-y-z$ makes z a twin of v by the preceding argument. If y is a twin of v, the edge $y-z$ implies $v-z$, so $z \in A$. Connectivity supplies a walk to each vertex; therefore every vertex of B is a twin of v.

This also describes the boundary-edge mechanism explicitly: an edge leaving a vertex already classified relative to v forces the next vertex into the opposite class. A failure at such an edge would contradict the two-step twin lemma, producing the induced P_4 above. The checked code uses induction on the walk and the twin lemma; it does not import a separate boundary-edge or distance theorem.

Triangle-freeness makes A independent. Every $b \in B$ has neighborhood exactly A, so all cross pairs are edges and there are no edges within B. Thus A,B are a complete bipartition. If $n \geq 2$, connectedness supplies an edge, so A and B are both nonempty (B contains v). This contradicts the assumed failure of complete bipartiteness and proves the lemma. □

Formal boundary and small orders. The Lean predicate HasCompleteBipartition asserts the existence of a set S such that adjacency is exactly opposite membership in S. It permits empty parts internally, which is convenient for the structural alternative. The public attainer theorem separately proves both parts nonempty under $n \geq 2$ and connectedness. Thus the singleton is not conflated with the positive-part theorem.

Four distinct vertices are part of the definition of InducedP4, together with all three edges and all three nonconsecutive nonedges. A mere path inside $K_{2,2}$ would not suffice: it has the fourth edge closing a cycle and has $D = 0$. Lemma 2 implies that any connected triangle-free graph outside the complete-bipartite class has $n \geq 4$. Consequently orders 2 and 3 need no separate non-bipartite analysis.

mathdb:2611.00007v1 03 Nov 2026

4. The injection and the strict deficit gap

Lemma 3 (gap). If G is connected, triangle-free and not complete bipartite, then $D \geq n - 2$.

Proof. By Lemma 2 choose an induced path $a-b-c-d$. Its four vertices contain exactly two one-edge triples: {a,b,d} and {a,c,d}. The other two triples contain two consecutive path edges. For each vertex x outside this path we show that there is a one-edge triple consisting of x and two path vertices.

Let S be the set of path vertices adjacent to x. Triangle-freeness makes S independent in the path. Assume, for contradiction, that no required triple exists. Every path edge must have an endpoint in S; otherwise that edge together with x is already a one-edge triple. Thus S is an independent vertex cover of the path. Each path edge has exactly one endpoint in S, so membership alternates along a,b,c,d: S is either {a,c} or {b,d}. In both cases exactly one of a,d is in S. Since a and d are nonadjacent, {x,a,d} now induces exactly one edge, the required contradiction.

Choose one such pair of path vertices for every exterior x; this is finite choice. The selected supports are injective as x varies: each support has exactly one vertex outside {a,b,c,d}, and that vertex recovers x. They are disjoint from the two seed supports, which lie wholly in the path; the seed supports are distinct. We have therefore injected a set of size $2 + (n-4) = n-2$ into the family of induced one-edge triples. Lemma 1 identifies the size of this family with D, proving the gap. □

The counting bridge in Lean. For each edge orientation (u,v), a good ordered triple records a third vertex adjacent to neither endpoint. ConfigurationCount constructs this finite subtype and proves its cardinality is twice the deficit. A family of distinct unordered supports is doubled by the two orientations of its unique edge. Distinct supports cannot collide, and opposite orientations of one edge cannot coincide because the graph is loopless. Hence twice the family cardinality is at most 2D, which yields the desired inequality in $\mathbb{Z}$.

The proof uses no assumption on edges between different exterior vertices: a selected triple has just one such vertex. This is why the local extension argument can be summed without an extension induction, and why it applies to non-bipartite triangle-free graphs as well as bipartite ones. The only role of connectedness is in Lemma 2. Without it, the complete-bipartition-or-induced-path alternative would fail, for example on a disjoint union of complete bipartite components.

Sanity checks, not proof substitutes. The dossier includes the reviewer's actual four-vertex path and exact kernel reductions $\sigma_t(P_4) = 4$, $\sigma_t(K_{1,3}) = 12$ and $\sigma_t(K_{2,2}) = 0$. P_4 has $D = 2 = n-2$, consistent with the injection. These finite checks test the literal invariant and common convention errors. The universal result is supplied by the unbounded lemmas, not enumeration.

mathdb:2611.00007v1 03 Nov 2026

5. Discrete bracketing and assembly

Lemma 4. For every integer $n \ge 7$ there is a positive proper integer part a with $F(n,a) \ge n^4/16 - (n-3)^2$.

Proof. Let $h(t) = t(n-t)$ on integers $1 \le t \le \lfloor n/2 \rfloor$. Its consecutive increment $h(t+1)-h(t) = n-2t-1$ is positive before the last point. The first value $n-1$ is at most $n^2/8$: n^2-8n+8 is 1 at $n=7$ and increases thereafter. The last value $\lfloor n^2/4 \rfloor$ is at least $n^2/8$. Hence either h hits $n^2/8$ or consecutive values bracket it. For a bracketing pair $t,t+1$ with $t \ge 1$, its gap is at most $n-3$. One endpoint a is within $(n-3)/2$ of the target. The exact polynomial identity

$$F(n,a) = n^4/16 - 4(h(a)-n^2/8)^2$$

gives the claimed bound. Equivalently $16F(n,a) = n^4 - (8a(n-a)-n^2)^2$. All comparisons can thus be performed over integers after clearing denominators. □

The formal bracketing lemma selects the least t for which $8t(n-t) \ge n^2$ and chooses t or $t-1$. The lower endpoint is positive except in the exact first-point case, which is handled directly. This is the same discrete interval argument; no square roots, floating-point rounding or real optimization oracle is used. Exact hits, equal endpoint distances and the last interval cause no exceptions.

Proof of the theorem. Let G lie outside the complete-bipartite class. Then $n \ge 4$ by Lemma 2. Completing the square in (1) and applying Lemma 3 gives

$$\sigma_t(G) \le n^4/16 - n(n-2),$$

since $m(n^2-4m) = n^4/16 - 4(m-n^2/8)^2$. For $n \ge 7$, $n(n-2)-(n-3)^2 = 4n-9 > 0$. Lemma 4 therefore proves $\sigma_t(G) < B(n)$. For $n = 4,5,6$, the displayed upper bounds are 8, 385/16 and 57. They are strictly below the star values $F(n,1) = (n-1)(n-2)^2$, respectively 12,36 and 80. The non-complete-bipartite case has now been strictly excluded at every possible order.

For any $1 \le a < n$, $K_{a,n-a}$ is connected and triangle-free. Its vertices have degrees $n-a$ and a in their respective parts. All within-part contributions vanish, while the $a(n-a)$ cross pairs contribute $(n-2a)^2$ each. Thus $\sigma_t(K_{a,n-a}) = F(n,a)$, proving attainment of $B(n)$. More generally any graph with a complete bipartition of size $k,n-k$ has the same formula, independently of labeling. This proves the upper bound also inside that class. Finally a maximizer cannot be outside the class by the strict inequality, and inside it the formula forces $F(n,k) = B(n)$. Conversely that formula proves each such graph maximizes. □

mathdb:2611.00007v1 03 Nov 2026

6. Lean correspondence and trust boundary

The certificate works with SimpleGraph (Fin n), actual finite degrees and decidable adjacency, under Connected and CliqueFree 3. The mathematical namespace is Sigma365710. Decidable adjacency is computational data, not an extra graph restriction; the reviewer proves that σ_t is independent of which decider is supplied. Integer casts precede every subtraction. The public classification gives an actual set and its nonempty complement, with adjacency equivalent to opposite membership.

Binding or proof step	Lean declaration(s)
Unordered invariant; finite objective	sigmaT; F (Basic)
Actual K; complementary parts	completeBipartite; HasCompleteBipartition
Variance / deficit / dart identities	sigmaT_degree_identity; sigmaT_deficit_identity; orderedDefect_eq_twice_deficit
Structural alternative	two_step_twins; inducedP4_of_not_completeBipartition
Local extension / injection / gap	path_extension_witness; deficit_ge_of_support_injection; inducedP4_deficit_gap; non_completeBipartition_deficit_gap
Integer bracketing	F_square_identity; exists_bracketing_part
Bipartition values	completeBipartite_sigmaT; bipartition_value
Strict exclusion / domination	non_completeBipartition_strict; extremal_completeBipartite
Classification / exact endpoint	attainer_classification; mathdb_365710

The frozen packet contains 11 Sigma/*.lean modules plus NegativeControls.lean: 12 mathematical/control modules in total. All are copied byte for byte, including the control module. Audit.lean and PrintAxioms.lean are inspection drivers, not dependencies of the endpoint. ReviewerSanity.lean is an explicitly labeled supplement; ReviewerAudit.lean extends the full declaration inventory to it. No mathematical source is edited to accommodate the manuscript.

The replay compiles each dependency in order with -j1 -DwarningAsError=true and runs leanchecker on every compiled mathematical/control and sanity module. The full audit includes generated local declarations, records module origin, full explicit and raw types, actual value? true body presence, safety, partiality and transitive axioms. It rejects absent bodies, unsafe or partial declarations and any axiom outside propext, Classical.choice and Quot.sound. The JSON declaration inventory and original logs are part of the receipts.

The trust boundary includes the pinned Lean elaborator and kernel/checker binaries, their runtime, and the cached upstream dependency objects. The dossier pins Lean 4.34.1, commit 5045d0056413266e57c625dcd7c365b10e377c52, and mathlib d13f23b723b8a846827a245b89c10fc7d3f11612; pins.json records binary and root-object hashes and dependency revisions. It does not rebuild every upstream object, prove the compiler correct, or infer statement fidelity or novelty from kernel success. The source reviews supply the semantic comparison; this paper spells out the mathematical correspondence.

mathdb:2611.00007v1 03 Nov 2026

7. Related work, limitations and reproduction

Zhou [Z] supplies the classical triangle-free Zagreb bound and the deficit identity, as detailed above. The reviewers read his six-page body. Caro–Yuster [CY], whose full twelve-page body was reviewed, study unrestricted degree variance. Their unrestricted maximum does not solve the restricted problem: at n=8 its σ_t value is 300, whereas B(8)=252. Their degree-increasing switch can create triangles, so filtering their unrestricted maximizers does not establish the triangle-free assertion.

Snijders [S] studies degree variance; the reviewer read journal pp.163–169 and the opening of the TW-229 companion. The bipartite variance there concerns degrees of one part, not the all-vertex invariant here; the undirected extremal results inspected are unrestricted. The remainder of that journal article and companion were not fully read. Bell [B] remains unread. Body-level restatements in Zhou, Remark 4, and Elphick–Wocjan [EW], p.52, characterize the cited Bell results as unrestricted quasi-clique/quasi-star squared-degree or variance bounds. Those restatements do not clear every theorem in Bell's original paper.

The actual origin papers Réti [R] and Dimitrov–Stevanović [DS] also remain unread. Related arXiv papers were compared in the focused review but were not established to be equivalent versions of those bodies. We do not exclude relevant partial lemmas there. The refereed 2026 successor [FDKŠ-J] retains the target as Conjecture 6.1. This is useful context, not proof that no later solution exists. No fresh external search is claimed in this manuscript lane.

The independent reviewers retained bounded novelty HOLDs; the supplied root adjudication assigns QUALIFIED_PASS while explicitly accepting these residual risks. The strict-gap literature screen found no established subsumer in the bodies inspected, but was not exhaustive. This draft preserves that distinction. It makes no worldwide-first or unconditional new-to-literature assertion and awards zero credit. Root alone may accept; the campaign stays at 6/20 beyond 39.

Reproduction. REPRODUCE.md describes safe extraction and the LEAN_BIN and LEAN_PACKAGES overrides. The source-only ZIP contains no compiled objects or upstream cache. Its manifest is checked before a fresh sequential replay; existing oleans are refused. Each substantive command has a two-CPU affinity cap, CPU limit 120 s, wall limit 140 s, 64 GiB virtual-address limit, MALLOC_ARENA_MAX=2 and sampled 6 GiB private-RSS guard. Total VmRSS is also reported, and a systemd user scope adds MemoryMax=6G where available. Three negative controls test specific false factor, induced-path and triangle-free mutations; their rejection is not a general soundness test. Frozen input hashes are compared before and after. Rendered pages and a visual QA receipt accompany the full dossier.

References and evidence locators. [FDKŠ] S. Filipovski, D. Dimitrov, M. Knor and R. Škrekovski, Some results on sigma_t-irregularity, arXiv:2411.04881v1 (2024), Eq.4, Props.1,10,11, Conj.24. [FDKŠ-J] Same work, Ars Math. Contemp. 26 (2026), #P4.01, DOI 10.26493/1855-3974.3268.60f. [Z] B. Zhou, Zagreb indices, MATCH 52 (2004), 113–118. [CY] Y. Caro and R. Yuster, Graphs with large variance, Ars Combin. 57 (2000), 151–162. [S] T.A.B. Snijders, The degree variance: an index of graph heterogeneity, Social Networks 3 (1981), 163–174. [B] F.K. Bell, A note on the irregularity of graphs, Linear Algebra Appl. 161 (1992), 45–54 (body unread). [EW] C. Elphick and P. Wocjan, New measures of graph irregularity (2014), inspected passages identified in the frozen reviews. [R] T. Réti, Appl. Math. Comput. 344–345 (2019), 107–115, DOI 10.1016/j.amc.2018.10.010 (body unread). [DS] D. Dimitrov and D. Stevanović, Appl. Math. Comput. (2023), article 127709, DOI 10.1016/j.amc.2022.127709 (body unread). Detailed access limits, body-reading coverage and receipts are preserved in sources/*review*.md and the root adjudication.

mathdb:2611.00007v1 03 Nov 2026