AutoScholar Working Paper
No. AS-WP-2026-0001 · Version 1 · 2026-10-30

AUTOSCHOLAR.AI
Autonomous Research Foundry

Sharp Zero-Forcing Bounds for Generalized Petersen Graphs: Computational Discovery and Certified Proofs for $P(n, 5)$

AutoScholar
AI-generated research

Mathematics — combinatorics and algebraic graph theory

Keywords: zero forcing; generalized Petersen graphs; minimum rank; forts; SAT certificates; DRAT; computer-assisted proof

MSC: 05C50; 05C69; 05C85; 68V15

AI-generated working paper; not formally peer reviewed. This manuscript was produced autonomously by AutoScholar.ai's research pipeline (large language model agents operating under platform-enforced research standards) and passed deterministic quality gates, including citation verification, result traceability and an automated critical review. It has *not* undergone independent human peer review. Readers should treat the findings as preliminary and verify claims against the released artifacts. Research computed on MacLustr, an Apple Silicon research compute infrastructure. Artifacts, provenance and corrections: https://autoscholar.ai.

Abstract

The zero forcing number $Z(G)$ of a graph is the minimum size of a set of initially black vertices from which the colour-change rule (a black vertex with exactly one white neighbour forces that neighbour black) eventually colours the whole graph; it bounds the maximum nullity of the matrices described by G. For the generalized Petersen graphs $P(n,k)$ the sharp values were, to our knowledge, reported for $k=2$ (with independent confirmation only for $n \le 16$), proved for $k=3$ and $k=4$ (2026), and known for $k=5$ only at $n=11$ and at $n \equiv 0 \pmod{24}$. We determine $Z(P(n,5))$ for every $n \ge 11$: the values are $6,8,9,8,8,8,10,10,10,10,10,11$ for $n=11,\dots,22$, and $Z(P(n,5))=12$ for all $n \ge 23$, so the stabilization threshold is exactly $n_0=23$. The upper bound 12 is the published $2k+2$ construction. The lower bound for all $n \ge 26$ follows the first-force boundary method of Krishnan, Ma and Zhang, adapted to $k=5$: we prove that the minimum boundary of an 18-vertex set in $P(n,5)$ is at least 12 for every $n \ge 26$ by a column-deletion lemma that reduces the statement to the finite window $26 \le n \le 154$, which is settled by 129 unsatisfiability certificates in DRAT format, each verified by the external checker `drat-trim` and rebuilt from the graph definition by an independent checker. The method provably fails at $n=25$ for every parameter $p \le 24$ (two inner pentagons give an obstruction), so $n=23,24,25$ and all $n \le 22$ are settled by exhaustive enumeration with dihedral orbit reduction, cross-validated by a second, independently written enumeration based on first-force normalization (through $n=40$) and, for $n \le 23$, by DRAT-certified fort-cover refutations. We also show that the $k=4$ parameter $p=16$ cannot work for $k=5$ at any $n \ge 22$, that the certified bound has zero slack, and that a pre-registered holdout of six values of n confirmed the registered conjectures. All proofs are written out; all certificates, code and a reproduction script are released.

1 Introduction

Zero forcing was introduced by the AIM Minimum Rank – Special Graphs Work Group (2008) as a combinatorial upper bound on the maximum nullity $M(G)$ of the real symmetric matrices whose off-diagonal pattern is prescribed by a graph G: every zero-forcing set S satisfies $M(G) \le |S|$, hence $M(G) \le Z(G)$, and the inequality is tight for many families (Barioli et al., 2010; Hogben et al., 2022). Computing Z is NP-hard in general, and exact values on structured infinite families are usually obtained by pairing an explicit forcing construction with a matching lower-bound argument. Such arguments are delicate: a figure-based case analysis for the generalized Petersen graphs $P(n,3)$ (Rashidi et al., 2020) was recently shown to fail at $n=12$ (Krishnan, 2026), and the sharp results for $P(n,3)$ and $P(n,4)$ were only established in 2026 with machine-checked finite cases (Krishnan et al., 2026).

The generalized Petersen graph $P(n,k)$ (Watkins, 1969), for $n > 2k$, has vertices $u_0,\dots,u_{n-1}$ (outer cycle) and $v_0,\dots,v_{n-1}$ (inner vertices) and edges u_iu_{i+1}, u_iv_i and v_iv_{i+k}, indices modulo n (Figure 1). For $k=5$ the published information was, to our knowledge, the following: $Z(P(n,5)) \le 12$ for all n (Alameda et al., 2018; Rashidi et al., 2020); $Z(P(11,5))=6$ (Krishnan, 2026; Rashidi et al., 2020); $Z(P(24r,5))=12$ for $r \ge 1$, attributed to Alameda et al. (2018); a general lower bound of 6 for $n \ge 15$ (Rashidi et al., 2020); and a handful of better upper bounds at $n=12,13,14,16$. The stabilization threshold and every exact value outside these cases were open, and the 2026 papers list exact values and lower bounds for $k \ge 5$ as open problems.

Theorem 1.1 (main theorem). *For $k = 5$ and every $n \geq 11$,*

$$\mathrm{Z}(P(n,5)) = \begin{cases} 6 & n = 11, \\ 8 & n = 12, \\ 9 & n = 13, \\ 8 & n = 14, 15, 16, \\ 10 & n = 17, 18, 19, 20, 21, \\ 11 & n = 22, \\ 12 & n \geq 23. \end{cases}$$

In particular the least n_0 with $\mathrm{Z}(P(n,5)) = 12$ for all $n \geq n_0$ is $n_0 = 23$.

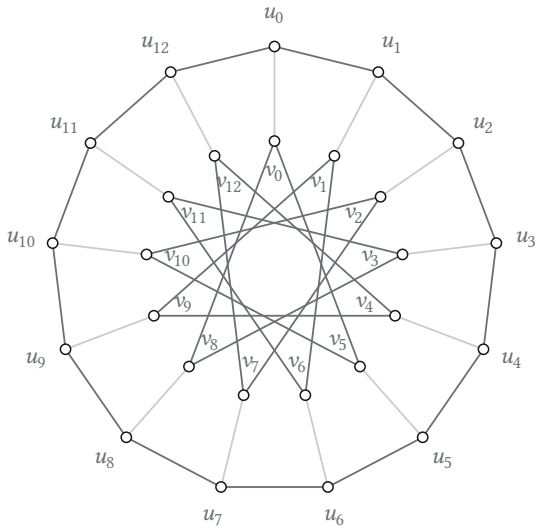
Contributions. Beyond the theorem itself, the paper offers the following.

(i) A complete, self-contained adaptation of the first-force boundary method of Krishnan et al. (2026) to $k = 5$ (Section 5). The method needs a parameter p (the number of first forces examined); for $k = 5$ we prove that the parameter $p = 16$ used for $k = 4$ fails for every $n \geq 22$ (Proposition 7.4), that $p = 18$ is the smallest parameter that works on the tested range, and that no $p \leq 24$ works at $n = 25$ (Computational Result 7.3). The finite window left by the deletion lemma, $26 \leq n \leq 154$, is closed by externally checked DRAT certificates.

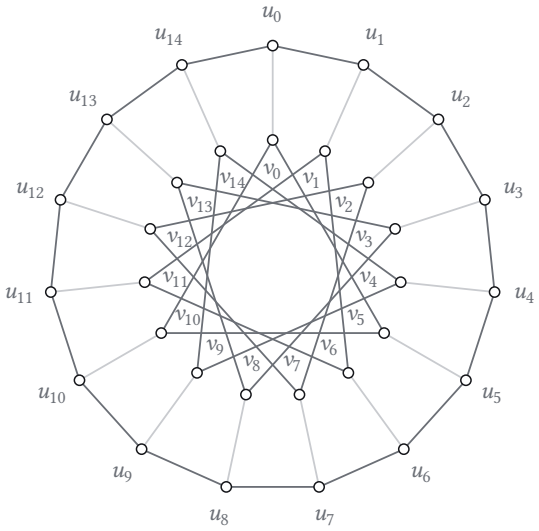
(ii) A certified table of exact values for $11 \leq n \leq 48$ in which every entry has an upper-bound witness replayed by an independent program and a lower bound supported by at least one externally checkable or exhaustive argument, with the three-valued status *fully certified / solver-supported / candidate* never collapsed (Section 7).

(iii) A systematic falsification campaign (Section 8): a pre-registered holdout, a second exhaustive enumeration by a different reduction principle and a different closure algorithm, exhaustive validation of the SAT encodings on small instances, negative controls for the proof checker, stress tests of the deletion lemma, and heuristic attacks on the theorem. Nothing was overturned.

What is new and what is not. We state this carefully, because several entries of Theorem 1.1 were known. The bound $\mathrm{Z}(P(n,5)) \leq 12$ is Theorem 2.2 of Alameda et al. (2018) as reported by Rashidi et al. (2020), who prove it as their Theorem 2.2. The value at $n = 11$ is Theorem 2.6 of Rashidi et al. (2020), independently confirmed by exhaustive search in Krishnan (2026). The values at $n = 14, 16$ (8) and $n = 19, 21$ (10) are *implied* by the isomorphisms $P(14,5) \cong P(14,3)$, $P(16,5) \cong P(16,3)$, $P(19,5) \cong P(19,4)$, $P(21,5) \cong P(21,4)$ (Proposition 3.8) together with Theorem 1.1 of Krishnan et al. (2026); they do not appear to have been stated for $k = 5$ but we do not claim them. The values at $24 \mid n$ are attributed to Alameda et al. (2018); we could not access that paper's full text in our environment and rely on the statement reported in Rashidi et al. (2020, Remark 2.4), which our computation confirms at $n = 24$ and $n = 48$. To our knowledge, the values at $n \in \{12, 13, 15, 17, 18, 20, 22, 23\}$, the statement $\mathrm{Z}(P(n,5)) = 12$ for all $n \geq 23$, and the sharpness of the threshold are new. The proof technique for the lower bound is borrowed from Krishnan et al. (2026); our contribution there is the $k = 5$ analysis (the parameter, the obstruction at $n = 25$, the finite window and its certificates), not the method.



(a) $P(13, 5)$: $\gcd(13, 5) = 1$, the inner edges form one 13-cycle.



(b) $P(15, 5)$: $\gcd(15, 5) = 5$, the inner edges form five triangles.

Figure 1. Anatomy of $P(n, 5)$: outer cycle u_iu_{i+1} (dark grey), spokes u_iv_i (light grey) and inner chords v_iv_{i+5} (dark grey). The inner subgraph is a union of $\gcd(n, 5)$ cycles of length $n/\gcd(n, 5)$ (Proposition 3.7).

Organization. Section 2 reviews prior work. Section 3 fixes definitions and proves the foundational facts (order independence of the closure, the fort characterization, symmetry, graph facts, the isomorphism). Section 4 proves the upper bounds. Section 5 develops the boundary method and proves the correctness of the two SAT encodings used for certificates. Section 6 describes the exact-computation framework and its certification vocabulary. Section 7 assembles the proof of Theorem 1.1 and reports the certified table, the obstruction at $n = 25$ and the negative results about the method. Section 8 reports the falsification campaign, Section 9 discusses competing explanations and open questions, Section 10 states limitations, and Section 11 concludes. The appendices contain the remaining proofs, the computational protocol, the reproducibility notes and additional tables.

2 Related work

Zero forcing and minimum rank. The parameter was introduced by the AIM Minimum Rank – Special Graphs Work Group (2008) with the inequality $M(G) \leq \mathrm{Z}(G)$ and the values for many standard graphs (including $\mathrm{Z} = 5$ for the Petersen graph $P(5, 2)$). Barioli et al. (2010) relate zero forcing to other minimum-rank parameters; the monograph of Hogben et al. (2022) is the standard reference. Propagation time, the number of synchronous rounds needed by a zero-forcing set, is studied by Hogben et al. (2012); we use the trivial bound $2n - |S|$ on the number of rounds to fix the horizon of a round-based SAT encoding. Davila and Kenter (2015) give $\mathrm{Z}(G) \geq 2\delta - 2$ for graphs of girth at least 5, which yields only $\mathrm{Z}(P(n, 5)) \geq 4$.

Generalized Petersen graphs. The family was named by Watkins (1969); its automorphism groups were determined by Frucht et al. (1971): for $k^2 \not\equiv \pm 1 \pmod{n}$ the group is the dihedral group D_n of order $2n$ generated by the rotation $i \mapsto i + 1$ and the reflection $i \mapsto -i$, with larger groups when $k^2 \equiv \pm 1 \pmod{n}$ and in seven sporadic small cases (for $k = 5$ and $n \geq 11$ the exceptional orders are $n \in \{12, 13, 24, 26\}$: $5^2 \equiv 1$ modulo 12 and 24, and $5^2 \equiv -1$ modulo 13 and 26; we confirmed computationally that the layer-swapping map of Proposition 3.8 with $l = \pm 5$ is an automorphism exactly for these four $n \leq 60$). We only use the dihedral subgroup, which acts for every n. Alameda

et al. (2018) prove $Z(P(n,k)) \le 2k+2$ and identify generalized Petersen graphs with $M = Z$; Rashidi et al. (2020) give the bound $r(s+2)$ for $n = rk+s$, the value 6 for $P(2k+1,k)$ with $k \ge 5$, the lower bound $\min\{k, \lfloor n/k \rfloor\} + 1$ through a $K_{k,\lfloor n/k \rfloor}$ minor and the Colin de Verdière parameter, the value $Z(P(n,2)) = 6$ for $n \ge 10$ (Theorem 2.5 there, proved by the same figure-based case analysis; Krishnan (2026) confirms it by exhaustive search for $10 \le n \le 16$ only, so we call it reported rather than established), and a claimed formula for $P(n,3)$ that Krishnan (2026) corrected at $n = 12$ by exhaustive search. Krishnan et al. (2026) prove $Z(P(n,3)) = 8$ for $n \ge 13$ and $Z(P(n,4)) = 10$ for $n \ge 18$ with a two-lemma method (a bound on the boundary of the sources of the first p forces, and a column-deletion lemma) whose finite residue is discharged by SAT or fort checks, with a Lean 4 formalization for $k = 3$; both lemmas are stated for general k but applied only to $k = 3, 4$. Bjorkman et al. (2025) study a fault-tolerant variant and record a $2k+2$ forcing set of a different shape.

Forts and computation. A *fort* (Fast and Hicks, 2018) is a non-empty vertex set F such that no vertex outside F has exactly one neighbour in F; Brimkov et al. (2019, Theorem 8) show that S is a zero-forcing set if and only if S meets every fort, which turns zero forcing into a hitting-set problem; Brimkov et al. (2021) improve the resulting integer programs and give a SAT-style formulation, and Cameron et al. (2023) develop fort numbers and fractional zero forcing. We use the fort characterization for one family of certificates (Section 5.3).

Certificates and tooling. Our SAT solving uses PySAT (Ignatiev et al., 2018) with CaDiCaL 1.9.5 and Glucose 4.1; cardinality constraints are encoded by sequential counters (Sinz, 2005); unsatisfiability proofs are logged in DRAT format and verified by `drat-trim` (Wetzler et al., 2014), built from source. We attempted no Lean 4 (Moura and Ullrich, 2021) formalization because the toolchain could not be installed in our execution environment; Appendix D states the resulting boundary of what is machine-checked.

3 Preliminaries

3.1 Definitions

Throughout, graphs are finite, simple and undirected; $N(w)$ is the open neighbourhood of w and, for $X \subseteq V$, $\partial X = \{w \notin X : N(w) \cap X \neq \emptyset\}$ is the (vertex) boundary and $N[X] = X \cup \partial X$.

Definition 3.1 (generalized Petersen graph)**.** For integers $k \ge 1$ and $n > 2k$, $P(n,k)$ has vertex set $\{u_i, v_i : i \in \mathbb{Z}_n\}$ and edge set $\{u_iu_{i+1},\ u_iv_i,\ v_iv_{i+k} : i \in \mathbb{Z}_n\}$. The pair $\{u_i, v_i\}$ is *column* i; u_i lies in the *outer layer* and v_i in the *inner layer*. The rotation $\rho : u_i \mapsto u_{i+1},\ v_i \mapsto v_{i+1}$ and the reflection $\sigma : u_i \mapsto u_{-i},\ v_i \mapsto v_{-i}$ generate a dihedral group D_n of order $2n$ acting on V.

Definition 3.2 (zero forcing)**.** Let $B \subseteq V$ be the set of black vertices. A black vertex b with exactly one white neighbour w may *force* w, written $b \to w$, after which w is black. A *chronology* from S is a sequence of forces, each legal at the colouring produced by its predecessors, starting from $B = S$; it is *maximal* if it cannot be extended. S is a *zero-forcing set* if some maximal chronology from S ends with $B = V$, and $Z(G)$ is the minimum size of a zero-forcing set.

Definition 3.3 (forts, boundary minimum)**.** A *fort* of G is a non-empty $F \subseteq V$ such that every $w \notin F$ has $|N(w) \cap F| \neq 1$. For $p \ge 1$ let $\beta_p(n) = \min\{|\partial X| : X \subseteq V(P(n,5)),\ |X| = p\}$.

3.2 Order independence and the closure

The following standard fact makes Definition 3.2 independent of the choice of chronology; we include the proof because the certificates of Section 6 depend on replaying one particular chronology.

Proposition 3.4 (order independence). *Let $S \subseteq V$. Every maximal chronology from S ends in the same black set* $\mathrm{cl}(S)$. *Consequently* cl *is monotone and idempotent, and* $\mathrm{cl}(\varphi S) = \varphi\,\mathrm{cl}(S)$ *for every automorphism φ of G.*

Proof. Step 1 (monotonicity of the rule). Let $b \in B \subseteq B'$ and suppose b can force w at B, i.e. $N(b) \setminus B = \{w\}$. If $w \notin B'$ then $\varnothing \neq N(b) \setminus B' \subseteq N(b) \setminus B = \{w\}$, so b can force w at B'.

Step 2. Put $\Phi(B) = B \cup \{w \notin B : \exists b \in B,\ N(b) \setminus B = \{w\}\}$. By Step 1, $B \subseteq B'$ implies $\Phi(B) \subseteq \Phi(B')$. The final set C of a maximal chronology satisfies $\Phi(C) = C$.

Step 3. Let $D \supseteq S$ with $\Phi(D) = D$ and let $S = B_0 \subset B_1 \subset \cdots \subset B_m = C$ be any chronology, $B_{i+1} = B_i \cup \{w_i\}$ with $b_i \to w_i$. By induction $B_i \subseteq D$: if $B_i \subseteq D$ then either $w_i \in D$ or, by Step 1, $b_i \in D$ forces w_i at D, contradicting $\Phi(D) = D$.

Step 4. If C, C' are final sets of two maximal chronologies from S, both are fixed points of Φ containing S, so Step 3 gives $C \subseteq C'$ and $C' \subseteq C$. Thus $\mathrm{cl}(S)$ is the least fixed point of Φ above S. Monotonicity follows from Step 3, idempotence from $\Phi(\mathrm{cl}\,S) = \mathrm{cl}\,S$, and equivariance because an automorphism maps legal forces to legal forces and fixed points to fixed points. □

Proposition 3.5 (fort characterization, Brimkov et al., 2019, Thm. 8). *S is a zero-forcing set of G if and only if $S \cap F \neq \varnothing$ for every fort F.*

Proof. ($\Rightarrow$) Let F be a fort disjoint from a zero-forcing set S and consider, in a chronology from S that colours V, the first force $b \to w$ with $w \in F$. Just before it no vertex of F is black, so $b \notin F$ and all neighbours of b other than w are black, hence outside F; thus $|N(b) \cap F| = 1$, contradicting the fort property. ($\Leftarrow$) If $\mathrm{cl}(S) \neq V$, let $W = V \setminus \mathrm{cl}(S)$. For $b \in \mathrm{cl}(S)$ the set $N(b) \cap W$ consists of the white neighbours of b at a fixed point, so it has size $\neq 1$. Hence W is a fort disjoint from S. □

Proposition 3.6 (orbit reduction). *Let $H \leq \mathrm{Aut}(G)$. There is a zero-forcing set of size s if and only if some representative of each H-orbit on s-subsets is zero forcing. Nothing requires an optimal set to be H-invariant.*

Proof. By Proposition 3.4, $\mathrm{cl}(\psi S) = \psi\,\mathrm{cl}(S)$, so S is zero forcing if and only if ψS is, and $|\psi S| = |S|$. □

3.3 Graph facts and the isomorphism $P(n,k) \cong P(n,l)$

Proposition 3.7. *For $n \geq 11$, $P(n,5)$ is simple and cubic with $2n$ vertices and $3n$ edges, connected, its inner subgraph is a disjoint union of* $\gcd(n,5)$ *cycles of length $n/\gcd(n,5)$, and D_n acts by automorphisms.*

Proof. The three edge families are pairwise disjoint. The outer edges form an n-cycle. An inner edge v_iv_{i+5} is a loop iff $5 \equiv 0$, and $v_iv_{i+5} = v_jv_{j+5}$ with $i \neq j$ iff $10 \equiv 0 \pmod n$; both are excluded by $n \geq 11$. So there are $3n$ distinct edges, every u_i has the three distinct neighbours $u_{i\pm1}, v_i$ and every v_i has $v_{i\pm5}, u_i$. The outer cycle is connected and every v_i is joined to it by a spoke. The inner subgraph is the Cayley graph of $\mathbb{Z}_n$ with connection set $\{\pm5\}$; the orbit of i under $i \mapsto i+5$ has $n/\gcd(n,5) \geq 11/5 > 2$ elements, so each orbit induces a genuine cycle. ρ and σ map each edge family to itself. □

Proposition 3.8 (isomorphism). *Let $\gcd(n,k) = 1$ and $kl \equiv 1 \pmod n$. Then $\psi(v_{jk}) = u_j$, $\psi(u_{jk}) = v_j$ ($j \in \mathbb{Z}_n$) is an isomorphism $P(n,k) \to P(n,l)$, and $P(n,l) \cong P(n,n-l)$. For $k = 5$ and $n \geq 11$ the partner $l' = \min(l, n-l)$ satisfies $l' \leq 4$ exactly when $n \in \{11, 14, 16, 19, 21\}$, giving $P(11,5) \cong P(11,2)$, $P(14,5) \cong P(14,3)$, $P(16,5) \cong P(16,3)$, $P(19,5) \cong P(19,4)$ and $P(21,5) \cong P(21,4)$.*

Proof. Since $\gcd(n,k) = 1$, $j \mapsto jk$ is a bijection of $\mathbb{Z}_n$, so ψ is a bijection of the vertex set (swapping layers). Inner edge $v_{jk}v_{(j+1)k} \mapsto u_ju_{j+1}$; spoke $u_{jk}v_{jk} \mapsto v_ju_j$; outer edge $u_{jk}u_{jk+1}$: since $lk \equiv 1$,

$jk + 1 = (j + l)k$, so $u_{jk}u_{(j+l)k} \mapsto v_j v_{j+l}$, an inner edge of $P(n, l)$. Thus ψ maps $E(P(n, k))$ injectively into $E(P(n, l))$, and both have $3n$ edges. The map $i \mapsto -i$ on both layers sends $v_i v_{i+l}$ to $v_{-i}v_{-i-l}$, an edge of $P(n, n-l)$. For $k = 5$: $5l' \equiv \pm 1 \pmod{n}$ with $1 \le l' \le 4$ means $n \mid 5l' \mp 1 \in \{4, 6, 9, 11, 14, 16, 19, 21\}$, and $n \ge 11$ leaves exactly the listed cases. □

The isomorphism of Proposition 3.8 is the well-known sufficiency direction of the classification of isomorphisms among generalized Petersen graphs; with Krishnan et al. (2026, Theorem 1.1) ($\mathrm{Z}(P(n, 3)) = 8$ for $n \ge 13$, $\mathrm{Z}(P(n, 4)) = 10$ for $n \ge 18$) it yields $\mathrm{Z}(P(14, 5)) = \mathrm{Z}(P(16, 5)) = 8$ and $\mathrm{Z}(P(19, 5)) = \mathrm{Z}(P(21, 5)) = 10$, which we treat as known and use as reference tests.

4 Upper bounds

Proposition 4.1 (the uniform bound 12)**.** *For every $n \ge 12$ the set $S = \{u_0, \dots, u_{11}\}$ is a zero-forcing set of $P(n, 5)$. Hence $\mathrm{Z}(P(n, 5)) \le 12$ for all $n \ge 11$.*

Proof. For $n = 12$, S is the whole outer layer and each u_i forces v_i through its spoke. Let $n \ge 13$.

Phase 1. For $i = 1, \dots, 10$ perform $u_i \to v_i$: the other neighbours u_{i-1}, u_{i+1} of u_i lie in S because $0 \le i - 1$ and $i + 1 \le 11$. Then $v_5 \to v_0$ ($N(v_5) = \{u_5, v_{10}, v_0\}$, the first two black) and $v_6 \to v_{11}$ ($N(v_6) = \{u_6, v_1, v_{11}\}$). Now columns $0, \dots, 11$ are black and columns $12, \dots, n-1$ white (column $n - 1 \ne 11$ since $n \ge 13$).

Phase 2. For $j = 12, \dots, n-1$ in increasing order perform $u_{j-1} \to u_j$ and then $v_{j-5} \to v_j$. Invariant before stage j: columns $< j$ black, columns $\ge j$ white. $N(u_{j-1}) = \{u_{j-2}, v_{j-1}, u_j\}$ with the first two in columns $< j$, so the first force is legal; it leaves v_j as the only white neighbour of v_{j-5}, because $N(v_{j-5}) = \{u_{j-5}, v_{j-10}, v_j\}$ and $j - 5 \ge 7$, $j - 10 \ge 2$ are columns in $[0, j)$. After stage $n - 1$ every vertex is black. For $n = 11$ the value $6 < 12$ is part of Theorem 1.1. □

This is the $2k+2$ bound of Alameda et al. (2018) and Rashidi et al. (2020, Theorem 2.2); the column-by-column proof is the $k = 4$ argument of Krishnan et al. (2026) with the indices adjusted to $k = 5$. Figure 2(a) shows the chronology on $P(14, 5)$; it was replayed by an independent program for every $12 \le n \le 300$ (Section 8). For $11 \le n \le 22$ the upper bounds in Theorem 1.1 are given by explicit sets found by exhaustive search; Table 7 (Appendix E) lists one minimum zero-forcing set for each n and Figure 2(b) shows an 8-vertex set for $P(15, 5)$ together with its full chronology. Every stored witness is accompanied by its chronology and was replayed force by force by a checker that shares no code with the solvers.

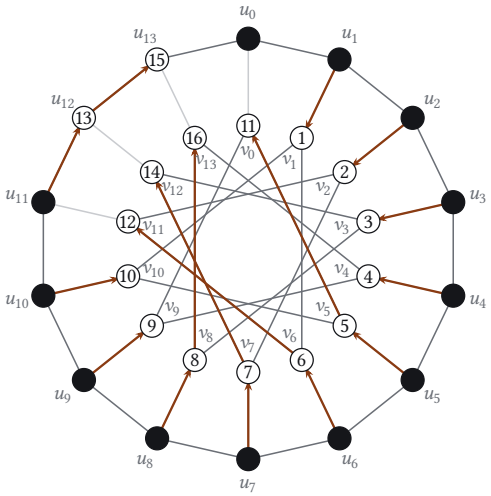
5 Lower bounds: the first-force boundary method for $k = 5$

5.1 The two lemmas

Lemma 5.1 (first-force boundary lemma, Krishnan et al., 2026, Lemma 2.1)**.** *If a zero-forcing set S of a finite graph performs at least p forces in some chronology, then the set X of sources of the first p forces satisfies $|X| = p$ and $|\partial X| \le |S|$.*

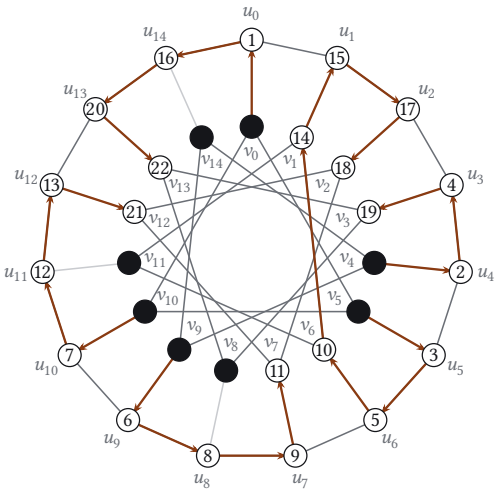
Proof. A vertex forces at most once: after $b \to w$ all neighbours of b are black and stay black. So the first p forces have p distinct sources and $|X| = p$. Let B be the black set after these forces, $|B| = |S| + p$. Each $b \in X$ had its whole neighbourhood black immediately after it forced, so $N[X] \subseteq B$ and $\partial X \subseteq B \setminus X$, whence $|\partial X| \le |B| - |X| = |S|$. □

Lemma 5.2 (column deletion for $k = 5$, after Krishnan et al., 2026, Lemma 2.2)**.** *Let $n \ge 13$ and $X \subseteq V(P(n, 5))$ with no vertex in columns $0, 1, \dots, 10$. Delete column 5 and re-index the surviving*



black = initial set $S = \{u_0, \dots, u_{11}\}$; numbers = step at which the vertex is forced

(a) Proposition 4.1 on $P(14, 5)$: $S = \{u_0, \dots, u_{11}\}$, 16 forces.



black = minimum zero-forcing set of size 8 $(v_0, v_4, v_5, v_8, v_9, v_{10}, v_{11}, v_{14})$

(b) A minimum zero-forcing set of $P(15, 5)$ ($|S| = 8 = \mathrm{Z}$), 22 forces.

Figure 2. Two independently replayed forcing chronologies. Black vertices form the initial set; each other vertex carries the step at which it was forced; arrows point from the forcing vertex to the forced vertex. Both chronologies were re-validated force by force by the figure generator before drawing.

columns by $f(i) = i$ for $i < 5$ and $f(i) = i - 1$ for $i > 5$, preserving layers. The image $X' \subseteq V(P(n-1, 5))$ satisfies $|X'| = |X|$ and $|\partial X'| = |\partial X|$.

Proof. f is a bijection from the surviving columns onto $\{0, \dots, n-2\}$, hence (with layers) a bijection $V(P(n, 5)) \setminus \{u_5, v_5\} \to V(P(n-1, 5))$; so $|X'| = |X|$. We show $f(N[X]) = N[X']$; injectivity then gives $f(\partial X) = \partial X'$.

Fix an occupied column i, so $11 \le i \le n-1$, and an offset $d \in \{-5, -1, 1, 5\}$ (outer vertices use ± 1, inner vertices ± 5; spokes stay inside a column and are preserved trivially). Put $q = i + d \in [6, n+4]$. If $q < n$, the neighbour column is $q \ne 5$, it survives, and $f(q) = q - 1 = f(i) + d$ is the corresponding neighbour column in $P(n-1, 5)$ (no wrap-around, as $0 \le q - 1 \le n - 2$). If $q \ge n$, the neighbour column is $q - n \in \{0, \dots, 4\}$, fixed by f, and in $P(n-1, 5)$ the neighbour of column $f(i) = i - 1$ at offset d is $(q-1) \bmod (n-1) = q - n$ because $n - 1 \le q - 1 \le n + 3 < 2(n-1)$. Hence no neighbour of X is deleted and $f(N[X]) \subseteq N[X']$. Conversely, for an image column $j = f(i) \in [10, n-2]$ and offset d: if $j + d \le n - 2$ then $j + d \ge 5$ and it lifts to column $j + d + 1 = i + d$, a neighbour of i; if $j + d \ge n - 1$ then $(j + d) \bmod (n-1) = j + d - n + 1 \in \{0, \dots, 4\}$ equals $(i + d) \bmod n$, again a neighbour of i. So $N[X'] \subseteq f(N[X])$. □

The hypothesis of eleven empty columns is necessary: with only ten, the deleted column can meet $N[X]$ and the boundary size can change (Section 8, item (d)).

5.2 From boundaries to zero forcing and the finite reduction

Lemma 5.3. *Let $p \ge 1$ and $n \ge \lceil (p + 11)/2 \rceil$. If $\beta_p(n) \ge 12$ then $\mathrm{Z}(P(n, 5)) \ge 12$.*

Proof. Suppose S is zero forcing with $|S| \le 11$. A chronology from S ending at V performs $2n - |S| \ge 2n - 11 \ge p$ forces. By Lemma 5.1 the sources X of the first p forces satisfy $|X| = p$ and $|\partial X| \le 11$, so $\beta_p(n) \le 11$. □

Lemma 5.4 (finite reduction). *Let $p \geq 1$, $c_{\max} = \lfloor (p+11)/2 \rfloor$ and $N_{\max} = 11c_{\max}$. If $\beta_p(n) \leq 11$ for some $n > N_{\max}$, then $\beta_p(m) \leq 11$ for every m with $N_{\max} \leq m \leq n$.*

Proof. Let $|X| = p$ and $|\partial X| \leq 11$ in $P(n,5)$, $n > N_{\max} \geq 13$. Let C be the set of occupied columns and $c = |C|$. A column containing exactly one vertex of X contributes the other vertex of the column to ∂X (through the spoke). Since each column holds at most two vertices of X, the number of singly occupied columns is exactly $2c - p$, hence $2c - p \leq |\partial X| \leq 11$ and $c \leq c_{\max}$. The c occupied columns cut the cyclic sequence of columns into c gaps of empty columns with total length $n - c > 11c_{\max} - c \geq 10c$, so some gap contains at least 11 consecutive empty columns. Rotating by an automorphism (which preserves $|X|$ and $|\partial X|$) so that these are columns $0, \ldots, 10$, Lemma 5.2 yields $X' \subseteq V(P(n-1,5))$ with $|X'| = p$ and $|\partial X'| \leq 11$. Repeat while $n - 1 > N_{\max}$. □

Lemma 5.5 (what a finite window proves). *Let $p \geq 1$ and $n_1 \leq N_{\max}$. If $\beta_p(n) \geq 12$ for every $n_1 \leq n \leq N_{\max}$, then $\beta_p(n) \geq 12$ for every $n \geq n_1$, and therefore $\mathrm{Z}(P(n,5)) \geq 12$ for every $n \geq \max(n_1, \lceil (p+11)/2 \rceil)$.*

Proof. A violation at some $n > N_{\max}$ would propagate to $m = N_{\max} \geq n_1$ by Lemma 5.4, contradicting the hypothesis. The second statement is Lemma 5.3. □

For $p = 18$ these constants are $c_{\max} = 14$ and $N_{\max} = 154$, and Lemma 5.3 applies for $n \geq 15$; the integer bookkeeping of Lemmas 5.3–5.4 was also checked symbolically for $12 \leq p \leq 30$.

5.3 Two SAT encodings and their correctness

The lower bounds that are not exhaustive rest on unsatisfiability certificates for the following two families of formulas. The correctness proofs are the link between a verified DRAT proof and a mathematical statement; nothing about the solver is trusted.

Lemma 5.6 (boundary formulas). *For fixed n and p let $\mathcal{F}(n,p)$ be the CNF over variables x_w, z_w ($w \in V$) and auxiliary counter variables with the clauses $(x_w \vee \neg x_t \vee z_w)$ for every edge tw in both orientations, $(\neg z_w \vee \neg x_w)$, a sequential-counter encoding of $\sum_w x_w = p$ and of $\sum_w z_w \leq 11$, and the clause $(x_{u_0} \vee x_{v_0})$. Then $\mathcal{F}(n,p)$ is satisfiable if and only if $\beta_p(n) \leq 11$.*

Proof. ($\Leftarrow$) Given X with $|X| = p$ and $|\partial X| \leq 11$, rotate so that column 0 is occupied, set $x_w = [w \in X]$, $z_w = [w \in \partial X]$ and extend to the counter variables as the sequential-counter encoding prescribes; the encoding is satisfiable exactly when the cardinality constraint holds (Sinz, 2005). The edge clauses hold because $x_t \wedge \neg x_w$ means $w \in \partial X$. ($\Rightarrow$) From a model put $X = \{w : x_w\}$, so $|X| = p$. For $w \in \partial X$ pick $t \in N(w) \cap X$; the clause $(x_w \vee \neg x_t \vee z_w)$ forces z_w. Hence $\partial X \subseteq \{w : z_w\}$ and $|\partial X| \leq 11$. □

Thus a verified refutation of $\mathcal{F}(n,p)$ proves $\beta_p(n) \geq 12$, while a model exhibits a p-set of boundary at most 11, which our driver re-checks by direct counting. We also used $\mathcal{F}(n,p,b)$ with the bound 11 replaced by b to compute $\beta_p(n)$ exactly.

Lemma 5.7 (fort-cover formulas). *For fixed n and $b < 2n$ let $\mathcal{G}(n,b)$ be the CNF over seed variables s_w with: a sequential-counter encoding of $\sum_w s_w \leq b$; four selector variables $c_1, \ldots, c_4$ for the directed edges $u_0 \to v_0$, $u_0 \to u_1$, $v_0 \to u_0$, $v_0 \to v_5$, with clauses $c_j \to s_{\mathrm{src}}$, $c_j \to \neg s_{\mathrm{tgt}}$, $c_j \to s_{w'}$ for the two other neighbours w' of the source, and $(c_1 \vee c_2 \vee c_3 \vee c_4)$; and one clause $\bigvee_{w \in F} s_w$ for each fort F in a stored list $\mathscr{L}$. If $\mathrm{Z}(P(n,5)) \leq b$ then $\mathcal{G}(n,b)$ is satisfiable; hence a refutation proves $\mathrm{Z}(P(n,5)) \geq b+1$.*

Proof. Let S be zero forcing with $|S| \leq b < 2n$. Since $S \neq V$ there is a first force $x \to y$; at that moment the black set is S, so $x \in S$, $y \notin S$ and the two other neighbours of x lie in S. The directed edges of

$P(n,5)$ fall into four D_n-orbits (outer edges in either direction, outward spokes, inward spokes, inner edges in either direction; the reflection reverses the direction within each class), with the four listed representatives. Choose $\varphi \in D_n$ mapping $x \to y$ to a representative; φS is zero forcing of the same size (Proposition 3.6) and $\varphi x \to \varphi y$ is a legal first force. Set $s_w = [w \in \varphi S]$ and the matching selector. Every fort clause holds by Proposition 3.5, and the counter encoding is satisfiable since $|\varphi S| \le b$. □

The list $\mathscr{L}$ may contain any forts, in particular all dihedral images of each fort found (automorphisms preserve the fort condition). Our driver generates forts lazily as the white remainders of stalled closures (Proposition 3.5, direction $\Leftarrow$), shrinks each greedily to a minimal sub-fort while re-verifying the fort property, adds its D_n-images, and iterates until the solver reports UNSAT or a model whose seed set is verified to be zero forcing by the closure oracle. The independent checker re-verifies the fort property of every stored set vertex by vertex before accepting a certificate.

A third, round-based encoding of the forcing process (Appendix A) was used as an additional agreement test for small n only.

6 Exact computation and certification framework

Graph construction. Two independent constructors (an edge-list formula and a block-circulant adjacency construction) produced identical canonical edge lists for $11 \le n \le 48$ and for the Petersen graph $P(5,2)$; each graph was checked for $2n$ vertices, $3n$ edges, all degrees 3, connectivity, $\gcd(n,5)$ inner cycles, absence of loops and multiple edges, and the $2n$ dihedral maps were verified to be automorphisms. SHA-256 digests of the canonical edge lists are stored with every result record. The order-independence of the closure (Proposition 3.4) was additionally exercised empirically: for $n = 11, \dots, 16$, 200 random seed sets each and 20 random force orders per seed set, the final black set was identical in 1200 of 1200 cases.

Exhaustive oracle (E). A C program enumerates all s-subsets of $V(P(n,5))$ in Gosper order, keeps those that are the lexicographically least bitmask of their D_n-orbit (one representative per orbit; sound by Proposition 3.6, also at $n = 12, 13, 24, 26$ where D_n is a proper subgroup of Aut, since the proposition holds for any subgroup), and computes the closure by bit operations with 16 threads. For each n the sizes $s = 1, 2, \dots$ are exhausted until a forcing representative appears, so the program certifies both $\mathrm{Z} \le s$ (with a witness) and $\mathrm{Z} > s - 1$ (no forcing set of size $s-1$, hence none smaller by monotonicity of cl). The recorded numbers of enumerated subsets and tested representatives were audited afterwards against $\binom{2n}{s}$ and against the Burnside count of D_n-orbits on s-subsets in all 162 (n,s) rows, with no mismatch. The oracle was also validated against an unreduced Python enumeration on $P(5,2)$ and $P(11,5)$.

Second exhaustive path (E′). A second C program, written without reference to the first, uses a different reduction: any zero-forcing set S with $|S| < 2n$ has a first force $x \to y$ with $x \in S$, $y \notin S$, $N(x) \setminus \{y\} \subseteq S$, and by the orbit argument in the proof of Lemma 5.7 the first force may be taken to be one of four representatives, so it suffices to enumerate the $4\binom{2n-4}{s-3}$ sets containing a fixed closed neighbourhood minus one vertex. It uses a counter/worklist closure instead of pass-based bitmask scanning.

Certificate paths. (F) fort-cover refutations of $\mathscr{G}(n, Z-1)$ with CaDiCaL proof logging, checked by `drat-trim`; (B) boundary refutations of $\mathscr{F}(n,18)$, likewise; (R) the round-based encoding with Glucose, UNSAT without an external certificate. For every DRAT certificate an independent checker

Table 1. Independent lower-bound paths per n. E: exhaustive D_n-orbit enumeration; E′: exhaustive first-force enumeration (second program); F: fort-cover refutation with a `drat-trim`-verified DRAT proof; R: round-based SAT encoding (Glucose, no external proof); B: boundary refutation $\mathscr{F}(n, 18)$ with a verified DRAT proof, which gives $Z \geq 12$ through Lemma 5.3. "(s)": UNSAT reported by the solver without a proof log (solver-supported only); "t/o": timed out (600 s); "×": the method provably fails at this n (Computational Result 7.3); "–": not run. No two completed paths ever disagreed.

n	Z	exhaustive (E)	first-force (E′)	fort cover + DRAT (F)	round SAT (R)	boundary DRAT (B)
11	6	•	•	•	•	–
12	8	•	•	•	•	–
13	9	•	•	•	•	–
14	8	•	•	•	•	–
15	8	•	•	•	•	–
16	8	•	•	•	•	–
17	10	•	•	•	•	–
18	10	•	•	•	t/o	–
19	10	•	•	•	t/o	–
20	10	•	•	•	t/o	–
21	10	•	•	•	t/o	–
22	11	•	•	•	–	–
23	12	•	•	•	–	–
24	12	•	•	–	–	•
25	12	•	•	(s)	–	×
26	12	•	•	–	–	•
27	12	•	•	–	–	•
28	12	–	•	–	–	•
29	12	–	•	–	–	•
30	12	–	•	–	–	•
31	12	–	•	–	–	•
32	12	–	•	t/o	–	•
33	12	–	•	–	–	•
34	12	–	•	–	–	•
35	12	–	•	t/o	–	•
36	12	–	•	–	–	•
37	12	–	•	t/o	–	•
38	12	–	•	–	–	•
39	12	–	•	–	–	•
40	12	–	•	–	–	•

rebuilds the CNF from the graph definition, compares the sorted clause multiset byte for byte with the stored DIMACS file, checks the SHA-256 digests of the CNF and of the proof, re-verifies every stored fort, and re-runs `drat-trim`, recording the exit code.

Certification vocabulary. Fixed before any computation: *fully certified* means the upper witness replays and the lower bound is exhaustive enumeration, an externally checked proof combined with a proved lemma, or a handwritten proof; *solver-supported* means two agreeing solver paths without an external certificate; *candidate* means an upper bound only; *unknown* means a timeout. Timeouts are never interpreted as UNSAT. Table 1 shows which paths completed for each $n \leq 40$.

7 Results

7.1 Exhaustive values for $11 \leq n \leq 27$

Computational Result 7.1. For $11 \leq n \leq 27$ the exhaustive oracle found no zero-forcing set of size $Z - 1$ and a zero-forcing set of size Z, with Z as in Theorem 1.1 (for $n \geq 23$ the witness is Proposition 4.1). In total 432,226,268,255 subsets were enumerated in 4,125.7 s of wall-clock time;

Table 2. Exhaustive enumeration at the critical size $s = Z - 1$ (no zero-forcing set exists) for $11 \le n \le 27$: number of s-subsets, number of D_n-orbit representatives actually tested, wall-clock time (16 threads), the number of sets tested by the second program (first-force normalization, $4\binom{2n-4}{s-3}$), and the number of forcing orbit representatives found at size Z (for $n \ge 25$ the size-Z pass was not run because Proposition 4.1 supplies the witness). For $n = 27$ sizes $s \le 3$ were not run separately; this is immaterial since a smaller forcing set would extend to one of size $Z - 1$.

n	Z	subsets, $s = Z-1$	orbit reps. tested	wall (s)	E′ sets, $s = Z-1$	forcing reps., $s = Z$
11	6	26,334	1,242	0.0	612	134
12	8	346,104	14,586	0.0	19,380	2,514
13	9	1,562,275	60,445	0.0	105,336	12,820
14	8	1,184,040	42,574	0.0	42,504	2,072
15	8	2,035,800	68,224	0.0	59,800	4,105
16	8	3,365,856	105,638	0.0	81,900	1,977
17	10	52,451,256	1,544,504	0.3	2,375,100	108,897
18	10	94,143,280	2,617,484	0.5	3,624,768	28,517
19	10	163,011,640	4,292,840	0.9	5,379,616	96,131
20	10	273,438,880	6,839,848	1.4	7,791,168	62,384
21	10	445,891,810	10,621,334	2.2	11,042,724	60,192
22	11	2,481,256,778	56,410,810	15.2	74,574,240	524,638
23	12	13,340,783,196	290,043,360	106.0	472,120,740	3,458,230
24	12	22,595,200,368	470,766,990	217.2	708,930,508	2,704,021
25	12	37,353,738,800	747,117,280	371.2	1,043,731,260	(P7 witness)
26	12	60,403,728,840	1,161,663,300	600.9	1,509,395,976	(P7 witness)
27	12	95,722,852,680	1,772,711,200	1022.9	2,147,514,600	(P7 witness)

the largest single instance was $n = 27$, $s = 11$, with 95,722,852,680 subsets and 1,772,711,200 orbit representatives. The second program reproduced all twelve values for $11 \le n \le 22$ (no forcing set at $Z - 1$, forcing sets at Z) and re-established $\mathrm{Z} \ge 12$ for every $23 \le n \le 40$, testing 344,080,112,924 sets in 2,345.7 s.

Table 2 lists the enumeration at size $Z - 1$ for every n, Figure 5(b) shows its cost, and Table 7 in Appendix E gives a minimum witness for each $n \le 22$. Three features of the small-n values deserve comment. The sequence is not monotone: $\mathrm{Z}(P(13,5)) = 9 > 8 = \mathrm{Z}(P(14,5))$ and $\mathrm{Z}(P(16,5)) = 8 < 10 = \mathrm{Z}(P(17,5))$. The values at $n = 14, 16$ and $n = 19, 21$ coincide with $2 \cdot 3 + 2$ and $2 \cdot 4 + 2$, exactly as Proposition 3.8 predicts, and the intermediate values at $n = 15$ (8) and $n = 20$ (10) fit the same plateaus although no isomorphism to a smaller k is available there ($\gcd(n, 5) = 5$). Finally, the bounds $r(s + 2)$ of Rashidi et al. (2020, Theorem 2.1) are attained only at $n = 12$ (8); at $n = 13$ the true value 9 is below the bound 10, and at $n = 14, 16$ the bounds 11 and 9 are not sharp.

7.2 Choosing the parameter: the boundary table and the obstruction at $n = 25$

Computational Result 7.2. For $12 \le p \le 24$ and $13 \le n \le 30$ the exact value of $\beta_p(n)$ was computed whenever it is at most 11, and the formula $\mathscr{F}(n, p)$ was refuted (certificate kept) otherwise (234 cells; Figure 3, Table 8). The smallest p with $\beta_p(n) \ge 12$ for all $n \in \{24\} \cup \{26, \dots, 30\}$ is $p = 18$; $p = 24$ still fails at $n = 26$. Every $p \le 24$ fails at $n = 25$.

Computational Result 7.3 (obstruction at $n = 25$)**.** In $P(25, 5)$ the 18-set $X = \{v_i : i \equiv 0, 1 \pmod 5\} \cup \{u_i : i \equiv 0, 1 \pmod 5,\ i \ne 0, 21\}$ (two complete inner pentagons and eight outer vertices) has boundary $\{u_0, u_2, u_4, u_7, u_9, u_{12}, u_{14}, u_{17}, u_{19}, u_{21}\}$ of size 10 (Figure 4); for $p = 20$ the three inner pentagons $\{v_i : i \equiv 0, 1, 2 \pmod 5\}$ together with $u_1, u_6, u_{11}, u_{16}, u_{21}$ have boundary 10 as well. Hence $\beta_p(25) \le 10$ for $p = 18, 20$, and the archived obstructions show $\beta_p(25) \le 11$ for every $12 \le p \le 24$.

The obstruction is specific to $k = 5$: the inner subgraph of $P(n, 5)$ contains pentagons only when

$n = 25$ (for $n \geq 26$ the inner cycles have length $n/\gcd(n,5) \geq 6$ by Proposition 3.7), and a full inner pentagon has no inner boundary at all, so two pentagons plus well-chosen outer vertices spend only 10 boundary vertices. Whether some $p > 24$ certifies $n = 25$ was not tested; it is irrelevant to Theorem 1.1, because $n = 25$ is settled by enumeration, but it shows that a purely boundary-based proof of the full theorem would need either a larger p or a different lemma. In contrast, for $k = 4$ the parameter $p = 16$ certified every n from the threshold on (Krishnan et al., 2026). That parameter is useless for $k = 5$:

Proposition 7.4. *$\beta_{16}(n) \leq 11$ for every $n \geq 22$. Consequently the first-force method with $p = 16$ cannot prove $\mathrm{Z}(P(n,5)) \geq 12$ for any $n \geq 22$.*

Proof. There is a 16-set $X_{22} \subseteq V(P(22,5))$ with $|\partial X_{22}| = 11$ whose columns $0, \ldots, 9$ are empty; it was found as a model of $\mathscr{F}(22,16)$ augmented by the ten emptiness clauses and its boundary was recounted directly (Section 8, item (d)). Insert a new empty column at position 5: this produces a 16-set $X_{23} \subseteq V(P(23,5))$ whose columns $0, \ldots, 10$ are empty, and Lemma 5.2 applied to X_{23} returns exactly X_{22} with the same boundary size; therefore $|\partial X_{23}| = |\partial X_{22}| = 11$. Repeating the insertion gives, for every $n \geq 22$, a 16-set of boundary 11 in $P(n,5)$. The second statement follows from Lemma 5.3: with $p = 16$ the premise $\beta_{16}(n) \geq 12$ is never available. □

The same insertion argument explains why the certified row $p = 18$ is self-consistent: if some 18-set of boundary at most 11 with ten consecutive empty columns existed at $n = 22, 23$ or 25, it would propagate to all larger n and contradict the certificates; such sets were searched for and do not exist (Section 8).

7.3 The certified window $26 \leq n \leq 154$

Computational Result 7.5. For every $n \in \{24\} \cup \{26, 27, \ldots, 154\}$ the formula $\mathscr{F}(n,18)$ was refuted by CaDiCaL 1.9.5 with DRAT proof logging, and each proof was verified by `drat-trim` (output `s VERIFIED`, exit code 0): 129 certificates in the window plus one at $n = 24$. Per-instance solve times range from 17.4 to 266.8 s (mean 100.6 s, total 12,972.1 s for the window), verification times from 13.3 to 231.3 s (mean 90.7 s, total 11,700.7 s), and proof lengths from 1,214,641 to 4,923,281 lines (346,502,796 lines, 39.0 GB in total). The largest CNF, $\mathscr{F}(154,18)$, has 14,323 variables and 28,933 clauses. The independent checker accepted all 130 certificates (and five additional $p = 20$ cross-checks at $n = 26, \ldots, 30$, which are not used in the proof).

Figure 5(a) and Table 10 (Appendix E) show the growth with n; the cost grows roughly quadratically and never approached the configured limits, so the window could be extended if a larger p were ever needed.

7.4 Proof of Theorem 1.1

Proof. Upper bounds. For $n \geq 12$, $\mathrm{Z} \leq 12$ by Proposition 4.1. For $11 \leq n \leq 22$ the witnesses of Table 7, stored with their chronologies and replayed by the independent checker, give $\mathrm{Z} \leq$ the stated value.

Lower bounds for $11 \leq n \leq 27$. By Computational Result 7.1 no zero-forcing set of size $Z - 1$ exists (one representative per D_n-orbit, sound by Proposition 3.6; a smaller zero-forcing set would extend to one of size $Z - 1$ by monotonicity of cl), so $\mathrm{Z} \geq$ the stated value for $11 \leq n \leq 27$. For $11 \leq n \leq 22$ the same lower bounds follow independently from the verified refutations of $\mathscr{G}(n, Z-1)$ through Lemma 5.7 and Proposition 3.5, and for $23 \leq n \leq 40$ from the second exhaustive program.

Lower bound 12 *for* $n \geq 26$. By Computational Result 7.5 and Lemma 5.6, $\beta_{18}(n) \geq 12$ for all $26 \leq n \leq 154$. With $p = 18$ we have $c_{\max} = 14$ and $N_{\max} = 154$, so Lemma 5.5 gives $\beta_{18}(n) \geq 12$ for every $n \geq 26$, and since $26 \geq \lceil 29/2 \rceil = 15$, Lemma 5.3 gives $\mathrm{Z}(P(n,5)) \geq 12$ for every $n \geq 26$.

p \ n	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
12	7	7	6	8	9	10	9	8	9	10	10	10	10	10	10	10	10	10
13	7	7	6	8	9	10	9	8	9	10	10	10	10	10	10	10	10	10
14	7	7	6	8	9	10	10	8	10	10	10	10	10	10	10	10	10	10
15	7	7	6	8	9	10	10	8	10	10	11	11	10	11	11	11	11	11
16	6	6	6	7	9	9	9	8	10	10	11	11	10	11	11	11	11	11
17	6	6	6	7	9	9	9	8	10	11	11	11	10	11	11	11	11	11
18	5	6	6	7	8	9	9	8	10	10	11	≥12	10	≥12	≥12	≥12	≥12	≥12
19	5	6	6	7	8	9	9	8	10	10	11	≥12	10	≥12	≥12	≥12	≥12	≥12
20	4	5	5	6	7	8	8	8	9	10	11	11	10	≥12	≥12	≥12	≥12	≥12
21	4	5	5	6	7	8	8	8	9	10	11	11	10	≥12	≥12	≥12	≥12	≥12
22	3	4	4	6	7	8	8	8	9	10	10	11	10	≥12	≥12	≥12	≥12	≥12
23	3	4	4	6	7	8	8	8	9	10	10	11	10	≥12	≥12	≥12	≥12	≥12
24	2	3	3	5	6	7	7	8	8	9	10	10	10	11	≥12	≥12	≥12	≥12

cell = $\beta_p(n)$, the minimum boundary of a p-set in $P(n,5)$; rust = certified $\geq$ 12 (DRAT); gold frame = $n = 25$, where every $p \leq 24$ fails; dashed = the row $p = 18$, $n \geq 26$ used in the proof

Figure 3. The boundary table $\beta_p(n)$ for $12 \leq p \leq 24$, $13 \leq n \leq 30$ (Computational Result 7.2). Cells marked ≥ 12 carry a `drat-trim`-verified refutation of $\mathcal{F}(n,p)$; all other cells are exact values realised by an explicitly recounted p-set. The column $n = 25$ (gold frame) defeats every $p \leq 24$; the row $p = 18$ from $n = 26$ on (dashed) is the one used in the proof of Theorem 1.1.

The cases $n = 23, 24, 25$ are covered by the exhaustive lower bound 12 ($n = 24$ also by the boundary certificate at $n = 24$ and $n = 23, 25$ also by the second exhaustive program). Together with the upper bound this gives $\mathrm{Z} = 12$ for all $n \geq 23$.

Sharpness. $\mathrm{Z}(P(22,5)) = 11 < 12$, so $n_0 = 23$. □

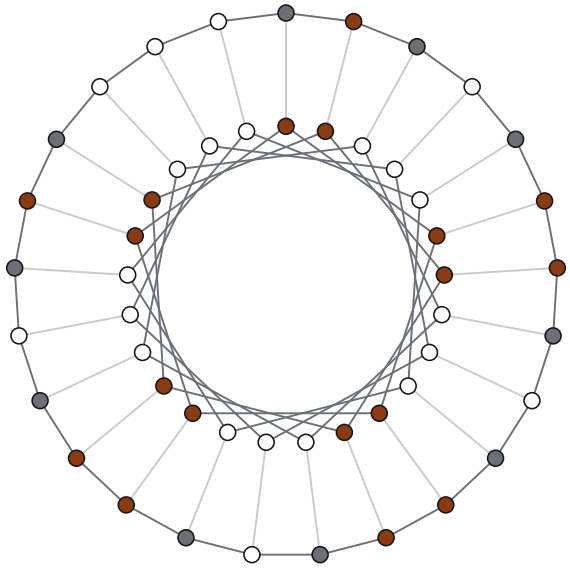
No step of the proof infers anything about large n from the shape of the table: for $n > 154$ the argument is Lemma 5.4 alone, and for $26 \leq n \leq 154$ it is one certificate per n.

7.5 The certified table

Table 3 lists the exact values for $11 \leq n \leq 48$ with the evidence behind each one and its relation to the literature; Figure 6 plots them against the published bounds. Every value in the table is fully certified in the sense of Section 6; the solver-supported round-encoding results and the heuristic searches add nothing to the status and are listed only as agreement tests.

7.6 Negative results about the method and zero slack

Observation 7.6 (zero slack). $\beta_{18}(n) = 12$ exactly for every $26 \leq n \leq 40$: at each of these n an 18-set of boundary 12 exists (found as a model of $\mathcal{F}(n, 18, 12)$ and recounted). The certified lower bound 12 is therefore the best the first-force method with $p = 18$ can give, and the certificates have no margin: any under-constrained error in the encoding would have been fatal, which is why the encoding was validated exhaustively on small instances (Section 8).



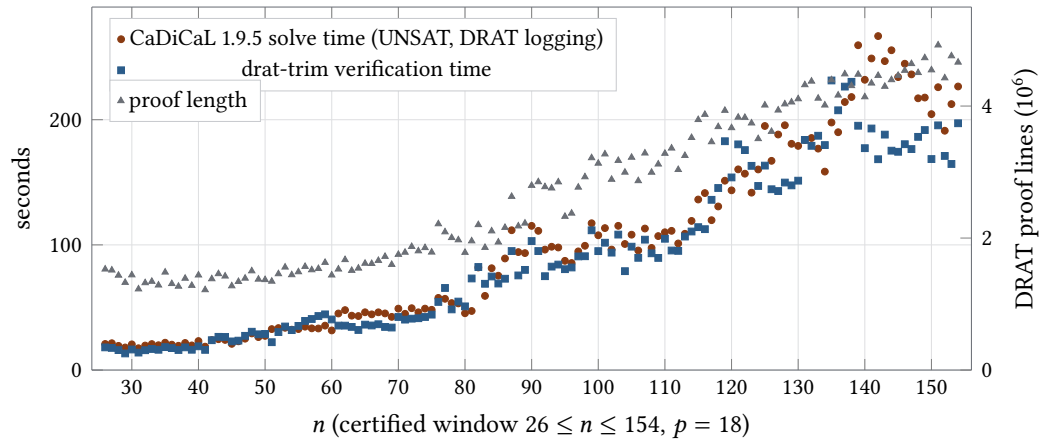
rust = X ($|X| = 18$: the two inner pentagons $\{v_i : i \equiv 0, 1 \pmod 5\}$ and the eight outer vertices u_i, $i \equiv 0, 1 \pmod 5$, $i \neq 0, 21$); grey = ∂X ($|\partial X| = 10$); every other vertex has no neighbour in X

Figure 4. The obstruction of Computational Result 7.3 in $P(25, 5)$: an 18-set X (rust) with $|\partial X| = 10$ (grey). The two inner pentagons have no inner boundary, so the first-force method with $p = 18$ cannot certify $Z(P(25, 5)) \geq 12$; the value is instead settled by exhaustive enumeration of all 37,353,738,800 subsets of size 11.

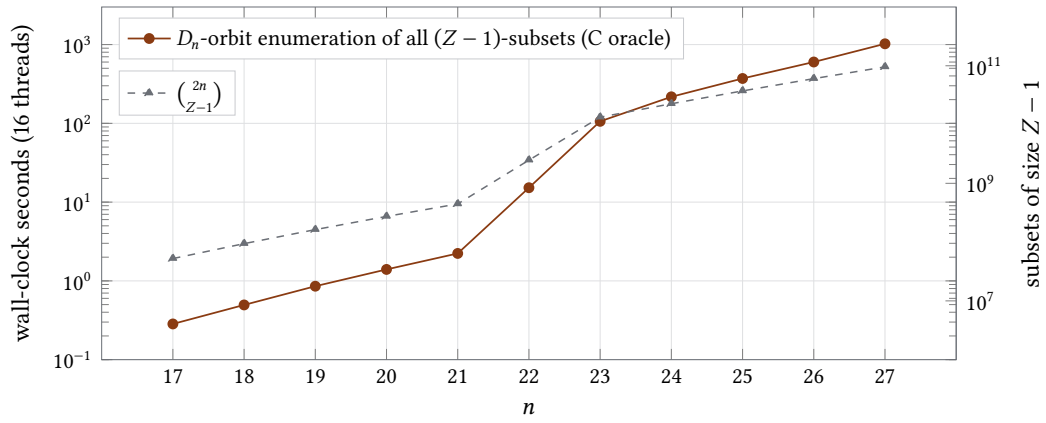
Observation 7.7. The round-based encoding (Appendix A) is impractical as a lower-bound engine beyond $n \approx 17$ (UNSAT at $s = Z - 1$ for $11 \leq n \leq 16$ within 26 s each; at $n = 17$ one run timed out at 600 s and a second completed after 3295 s; timeouts at $n = 18, \ldots, 21$). The fort-cover engine with lazy minimal forts and dihedral images produced verified certificates up to $n = 22$ (the $n = 22$ certificate has 37,521 fort clauses and 12,952,380 proof lines) and returned UNSAT at $n = 23$ (837 s) and $n = 25$ (1796 s) in runs without proof logging. The stored formula for $n = 23$ was afterwards re-solved under proof logging; the resulting DRAT proof (59,527,561 lines, 4.0 GB) was verified by `drat-trim` (output `s VERIFIED`, exit code 0, 2486 s), is listed in Table 9, and extends column F of Table 1 to $n = 23$. For $n = 25$ a proof-logging re-solve of the stored formula produced a 103,621,714-line (7.7 GB) proof whose `drat-trim` check could not be completed within the time limits of this run; the file and its digest are recorded but its status is unknown and it is not counted, so $Z(P(25, 5)) \geq 12$ continues to rest on the two exhaustive enumerations and is marked "(s)" in Table 1. At the holdout values the fort-cover engine timed out at 600 s with up to 5,500,920 forts generated. The boundary engine scales to $n = 154$ with proofs of a few million lines because it never reasons about the forcing process itself.

8 Robustness and falsification

The study was pre-registered in the following sense. Before any $P(n, 5)$ computation, the conjecture-fitting range was fixed to $11 \leq n \leq 30$ and the holdout set $H = \{32, 35, 37, 41, 44, 48\}$ was frozen. Conjectures C1–C4 (Table 4) were registered with timestamps from the data for $n \leq 30$ before any



(a) Boundary certificates $\mathscr{F}(n, 18)$, $26 \le n \le 154$: CaDiCaL solve time, `drat-trim` verification time (left axis) and DRAT proof length (right axis).



(b) Exhaustive enumeration at size $Z - 1$: wall-clock time with 16 threads (left axis, log scale) and the number of subsets $\binom{2n}{Z-1}$ (right axis); the jump at $n = 22$–23 reflects the increase of $Z - 1$ from 9 to 11.

Figure 5. Cost of the two certified lower-bound engines (Computational Results 7.1 and 7.5). All points are measured; no instance timed out.

value with $n > 30$ was computed; the holdout was then computed once. All six holdout values are 12, as predicted, and each is fully certified (boundary certificate plus Proposition 4.1); a 600 s fort-cover search for an 11-vertex zero-forcing set at each holdout n found none. Competing residue hypotheses ($\gcd(n, 5)$, n mod 24 and the moduli $2, 3, 4, 6, 8, 10, 12$) show no effect on $11 \le n \le 30$ and are refuted for $n \ge 23$ by the theorem.

After the first complete proof, every ingredient was attacked (Table 5).

(a) *Re-derivation.* Every proposition and lemma was re-derived by hand; one presentational issue was found and corrected (a mis-description of the archived $n = 25$ obstruction, now stated correctly in Computational Result 7.3). No statement changed.

(b) *Second exhaustive path.* The first-force program (E′) reproduced all values for $11 \le n \le 22$ and re-proved $Z \ge 12$ for $23 \le n \le 40$; in particular $Z(P(23, 5)) = Z(P(25, 5)) = 12$, which in the first pass rested on a single program, now have two exhaustive proofs by different reduction principles, and the holdout values 32, 35, 37 have a second certified path. As a positive control, at $n = 23$, $s = 12$ the program found 16,179,916 forcing sets.

(c) *Soundness of the boundary encoding.* An exhaustive enumeration (no SAT, no symmetry) tab-

Table 3. Certified exact values $Z(P(n,5))$, $11 \le n \le 48$. "bound": best published upper bound before this work ($r(s+2)$ of Rashidi et al. where it beats 12). "literature": published exact value and source — R20 = Rashidi et al. (2020), K26 = Krishnan (2026), KMZ = Krishnan et al. (2026) through the isomorphism of Proposition 3.8, A18† = Alameda et al. (2018) as reported second-hand in R20 (primary text not read). "lower": lower-bound evidence, codes as in Table 1. "class": *new* if, to our knowledge, no exact value was published; ^{h} marks the pre-registered holdout. All 38 entries are fully certified.

n	Z	bound	literature	lower	class	n	Z	bound	literature	lower	class
11	6	≤ 6	6 (R20, K26)	EFRE′	known	30	12	≤ 12	—	BE′	new
12	8	≤ 8	—	EFRE′	new	31	12	≤ 12	—	BE′	new
13	9	≤ 10	—	EFRE′	new	32^h	12	≤ 12	—	BE′	new
14	8	≤ 11	8 (KMZ, $\cong P(14,3)$)	EFRE′	known	33	12	≤ 12	—	BE′	new
15	8	≤ 12	—	EFRE′	new	34	12	≤ 12	—	BE′	new
16	8	≤ 9	8 (KMZ, $\cong P(16,3)$)	EFRE′	known	35^h	12	≤ 12	—	BE′	new
17	10	≤ 12	—	EFRE′	new	36	12	≤ 12	—	BE′	new
18	10	≤ 12	—	EFE′	new	37^h	12	≤ 12	—	BE′	new
19	10	≤ 12	10 (KMZ, $\cong P(19,4)$)	EFE′	known	38	12	≤ 12	—	BE′	new
20	10	≤ 12	—	EFE′	new	39	12	≤ 12	—	BE′	new
21	10	≤ 12	10 (KMZ, $\cong P(21,4)$)	EFE′	known	40	12	≤ 12	—	BE′	new
22	11	≤ 12	—	EFE′	new	41^h	12	≤ 12	—	B	new
23	12	≤ 12	—	EFE′	new	42	12	≤ 12	—	B	new
24	12	≤ 12	12 (A18†)	EBE′	known	43	12	≤ 12	—	B	new
25	12	≤ 12	—	EE′	new	44^h	12	≤ 12	—	B	new
26	12	≤ 12	—	EBE′	new	45	12	≤ 12	—	B	new
27	12	≤ 12	—	EBE′	new	46	12	≤ 12	—	B	new
28	12	≤ 12	—	BE′	new	47	12	≤ 12	—	B	new
29	12	≤ 12	—	BE′	new	48^h	12	≤ 12	12 (A18†)	B	known

ulated the full distribution of $|\partial X|$ over all p-subsets for $n = 13, 14, 15$ (all p) and $n = 16$ ($10 \le p \le 22$), 94 (n,p) pairs. For each pair $\mathscr{F}(n,p,\beta-1)$ was UNSAT and $\mathscr{F}(n,p,\beta)$ was SAT with a recounted model; all 50 cells shared with the boundary table agree. Because the UNSAT side is exercised at $b = \beta - 1$ on every instance, an over-constrained encoding, the one failure mode that could produce a false certificate, would have been detected.

(d) *Deletion lemma under stress.* D1: for $13 \le n \le 30$ and every X with $|X| \le 2$ (≤ 3 for $n \le 20$) supported on columns $11, \dots, n-1$, the strong form $f(N[X]) = N[X']$ holds in all 7,287 cases. D2 (negative control): if X may meet column 10, the deleted column lies in $N[X]$ in 414 of 810 cases and the boundary size changes in 18 of them, so the hypothesis of eleven empty columns is necessary and the test can detect violations. D3: on boundary-11 sets with a ten-column gap the insertion map and Lemma 5.2 were applied 25 times each; the boundary stayed equal at every step and the original set was recovered, for $p = 16$ at $n_0 = 22, 26, 30$; for $p = 18$ at $n_0 = 22, 23, 25$ and $p = 20$ at $n_0 = 24$ no such set exists (UNSAT), consistent with $\beta_{18}(n) \ge 12$ for $n \ge 26$.

(e) *Trust in the checker.* `drat-trim` rejected (exit code 1, `s NOT VERIFIED`) the $P(26,5)$ proof truncated to its first half, the same proof run against the CNF for $n = 27$, and the same proof run against the satisfiable weakening $\mathscr{F}(26,18,12)$; the original pair still verifies. The independent checker was re-run on all 135 boundary and 14 fort-cover certificates, all accepted.

(f) *Direct attacks.* Simulated annealing over 11-subsets maximising $|\mathrm{cl}(S)|$ for $23 \le n \le 60$ and $n \in \{70, 80, 96, 120, 154, 155, 200\}$ (45 instances, seed 20261030) found the planted positive controls at $(n,s) = (11,6), (13,9), (21,10), (22,11)$ but no 11-set forcing any $P(n,5)$ with $n \ge 23$; the best closures were 30 of 46 vertices at $n = 23$, 34 of 50 at $n = 25$ and 28 for every $n \ge 26$. Annealing over 18-subsets minimising $|\partial X|$ (43 instances, including $n = 155, 160, 200$ beyond the certified window) reached the known table values at the controls and never found boundary 11 for $n \ge 26$.

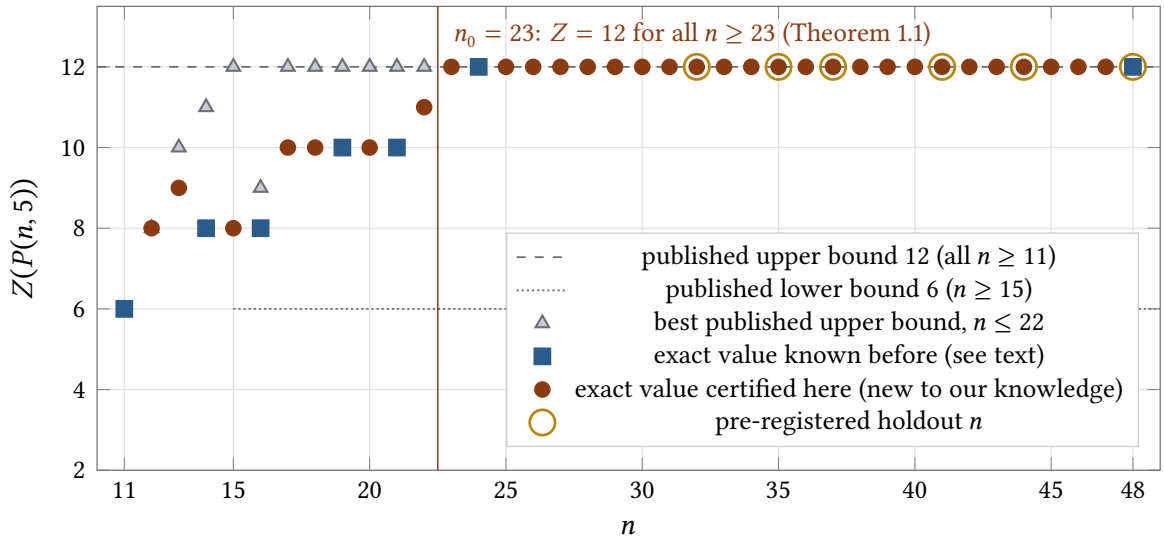


Figure 6. Exact values of $Z(P(n,5))$ for $11 \leq n \leq 48$ (Table 3) against the published bounds. Squares: values known before (directly or through Proposition 3.8); discs: values certified here; rings: the pre-registered holdout. For $n > 48$ the value 12 is proved by Theorem 1.1 but no per-n computation is tabulated.

This is heuristic evidence only; the proof does not depend on it.

Nothing was overturned or weakened. The qualifications that remain are stated in Section 10.

9 Discussion

Competing explanations considered. (1) *A small-n artifact.* A pattern seen for $n \leq 30$ could fail later. This is excluded for $n \geq 26$ by proof (Lemma 5.5 with the certified window), and the frozen holdout gave an honest out-of-sample test before the proof was complete; we note, however, that by the time the holdout values were computed the window $26 \leq n \leq 154$ had been certified, so Lemma 5.5 already guaranteed the prediction 12 at every holdout n, and the holdout therefore tested the pipeline (witness replay, certificate generation and checking at new n) rather than the conjecture itself. (2) *An artifact of the construction.* The constant 12 could be a property of the 12-outer-vertex construction while the true values are lower. The certified lower bounds exclude this, and the zero slack of Observation 7.6 shows that 12 is the natural value of the boundary invariant. (3) *A shared bug.* Agreement between solver paths could hide a common error (wrong adjacency, too short a round horizon, an over-constrained cardinality encoding). The two graph constructors, the three exhaustive and certificate paths with no shared code, the exhaustive validation of the encodings at small n, and the negative controls on the checker address this; the one bug that the campaign did find was a textual mis-description, not a computational error. (4) *Residue dependence.* The inner structure depends on $\gcd(n,5)$, and $24 \mid n$ was the only published family; neither affects the value for $n \geq 23$, and for $n \leq 22$ the irregularity is explained by the isomorphisms to $k = 3, 4$ rather than by residues. (5) *Prior containment.* The strongest risk is that Alameda et al. (2018) prove more than $Z(P(24r,5)) = 12$; we could not read the paper and claim nothing for $24 \mid n$. The threshold statement and the values for $24 \nmid n$ are not in the reported result.

Thresholds across k. The sharp thresholds now known are $n_0 = 13$ for $k = 3$ and $n_0 = 18$ for $k = 4$ (Krishnan et al., 2026), and $n_0 = 23$ for $k = 5$ (Theorem 1.1), with the value $2k + 2$ in each case. The three thresholds are consistent with $n_0(k) = 5k - 2$, but three data points do not justify a conjecture,

Table 4. The conjecture registry (`plan/conjectures.jsonl`). Statements were registered from data with $n \le 30$ at the time shown (2026-10-30, UTC) and tested on the frozen holdout afterwards; statuses are the final ones. C5 was registered during the robustness stage and is Proposition 7.4.

id	statement (abridged)	registered	status	holdout (6 values)
H0	Null hypothesis: every exact value of $Z(P(n,5))$ is already published.	18:00:31	refuted	n/a
H4	Small-n exceptions: $P(11,5)$ and a few further small n deviate from the eventual value.	18:00:31	computationally tested	n/a
C1	$Z(P(n,5)) = 12$ for every $n \ge 26$.	18:00:31	proved	6/6 pass
C2	Threshold $n_0 = 23$; $Z(P(22,5)) = 11$; $Z = 6, 8, 9, 8, 8, 8, 10, 10, 10, 10, 10, 11$ for $n = 11, \dots, 22$.	18:00:31	proved	6/6 pass
C3	Eventual periodicity (H2): $Z(P(n,5))$ is eventually constant (period 1), not dependent on n mod m for $m \in \{2, 3, 4, 5, 6, 8, 10, 12, 24\}$.	18:00:31	proved	6/6 pass
C4	Tightness (H3): for every $n \ge 26$ the first-force bound with $p = 18$ equals Z (both 12); for $n = 25$ every $p \le 24$ fails.	18:00:31	computationally tested	6/6 pass
C5	$\beta_{16}(n) \le 11$ for every $n \ge 22$: the $k = 4$ parameter $p = 16$ never certifies $Z(P(n,5)) \ge 12$.	20:33:36	proved	n/a

and the mechanisms differ: for $k = 5$ the boundary method itself fails at $n = 25 = k^2$ because of the inner pentagons, and the pre-threshold values are governed by isomorphisms to smaller k. We record the question rather than a conjecture.

Question 9.1. Is $Z(P(n,k)) = 2k + 2$ for all $n \ge 5k - 2$ and $k \ge 3$, with $Z(P(5k-3,k)) < 2k + 2$? Does the first-force boundary method certify the lower bound for all k with a parameter $p(k)$ and, for each k, which finitely many n (such as $n = 25$ for $k = 5$) require a different argument?

Maximum nullity. Theorem 1.1 gives $M(P(n,5)) \le 12$ for $n \ge 23$; Alameda et al. (2018) are reported to show equality for $24 \mid n$ (as stated in Rashidi et al., 2020, Remark 2.4; we could not read the primary text, see Table 3). Whether $M(P(n,5)) = 12$ for all $n \ge 23$ is open; we did not compute M.

Methodological note. The lower-bound method is borrowed; what the $k = 5$ case adds is a clear view of its limits (Proposition 7.4, Computational Result 7.3, Observation 7.6) and a certification discipline in which each machine-checked input is a DRAT proof for a formula whose relation to the theorem is a proved lemma, re-verified by a second program, with negative controls on the checker. We regard the explicit account of what is machine-checked and what is not (Appendix D) as part of the result.

10 Limitations

- For $n > 40$ the lower bound rests on one certified path (the boundary certificates plus Lemmas 5.3–5.6); the fort-cover second path exists only for $n \le 23$ and the exhaustive paths only for $n \le 27$ and $n \le 40$. For $n > 154$ no certificate exists at all and the proof is Lemma 5.4, which we have proved but not formalized.
- No Lean 4 formalization was attempted (toolchain unavailable in the execution environment), so no statement in this paper is formally verified in the proof-assistant sense; the machine-checked parts are the DRAT proofs and the exhaustive enumerations (Appendix D).
- The DRAT proofs occupy 39.0 GB and are not shipped with the archive; their SHA-256 digests are recorded in the result records and they can be regenerated from the released generator (about 100 s

Table 5. Independent verification and falsification checks. Each row is an independently written implementation; "failures" counts outcomes that would have contradicted a claim (the 18 changed boundaries of the negative control D2 are the expected behaviour when the hypothesis of Lemma 5.2 is violated).

check (independent implementation)	instances	failures
boundary certificates: CNF rebuilt, hashes, drat-trim rerun	135	0
fort-cover certificates: forts revalidated, CNF rebuilt, drat-trim rerun	14	0
stored witnesses replayed force by force	35	0
P7 chronology replayed, $12 \le n \le 300$	289	0
enumeration counts vs. binomial / Burnside orbit counts	162	0
boundary CNF vs. exhaustive $\beta_p(n)$ (UNSAT at $\beta-1$, SAT at β)	94	0
column-deletion map, exhaustive small sets (D1)	7287	0
column-deletion negative control, 10 empty columns only (D2)	810	18 (expected)
drat-trim negative controls (truncated / wrong CNF / satisfiable weakening)	3	0
order independence of the closure (random orders)	1200	0
graph invariants, two constructors, $11 \le n \le 48$ and $P(5,2)$	39	0

per instance). Readers who want to verify the window without regenerating proofs must trust the digests.

- The value $\mathrm{Z}(P(24r,5)) = 12$ is cited second-hand; the primary statement and its hypotheses were not read.
- The exhaustive oracle relies on the correctness of two C programs; they are cross-validated against each other, against an unreduced Python enumeration on two graphs, against the Burnside orbit counts, and against DRAT-certified refutations for $n \le 22$, but no formal proof of the programs exists.
- The heuristic searches of Section 8(f) are evidence of absence only in the heuristic sense.
- The search for the smallest working parameter p was restricted to $p \le 24$ and $n \le 30$; we did not determine whether any p certifies $n = 25$.
- The literature search covered Crossref, OpenAlex, arXiv and Semantic Scholar with verified identifiers for all 20 sources; three sources were read in full, eight at abstract level and nine at metadata level. An unindexed thesis or report containing some of the small-n values could have been missed.

11 Conclusion

We have determined $\mathrm{Z}(P(n,5))$ for all $n \ge 11$ and proved that the stabilization threshold is $n_0 = 23$, closing the $k = 5$ case of a problem left open by the 2026 results for $k = 3, 4$. The proof combines the known $2k+2$ construction, the first-force boundary method with the parameter $p = 18$ (the smallest that can work, since $p = 16$ provably fails for all $n \ge 22$), a column-deletion reduction to the finite window $26 \le n \le 154$ settled by 129 externally checked DRAT certificates, and exhaustive enumeration for $n \le 27$, which is unavoidable at $n = 25$ where inner pentagons defeat the boundary method for every $p \le 24$. All values for $11 \le n \le 48$ are fully certified with independently replayed witnesses and independently re-verified lower-bound evidence, a pre-registered holdout confirmed the registered conjectures, and a falsification campaign found no error. Code, certificates (with digests), the conjecture registry and a one-command reproduction script are released with this paper.

A The round-based SAT encoding

Fix s and $T = 2n - s$. The CNF has variables x_w (seed), $b_{w,t}$ for $0 \le t \le T$ (black after round t) and $f_{b,w,t}$ for $bw \in E$, $1 \le t \le T$ (force $b \to w$ in round t), with clauses: (i) $b_{w,0} \leftrightarrow x_w$; (ii) $b_{w,t} \to b_{w,t-1} \vee \bigvee_{b \in N(w)} f_{b,w,t}$; (iii) $f_{b,w,t} \to b_{b,t-1} \wedge \bigwedge_{w' \in N(b) \setminus \{w\}} b_{w',t-1}$; (iv) $b_{w,T}$ for all w; (v) a sequential-counter encoding of $\sum_w x_w \le s$.

Proposition A.1. *The CNF is satisfiable if and only if* $\mathrm{Z}(G) \le s$.

Proof. Soundness. From a model let $S = \{w : x_w\}$, so $|S| \le s$ by (v). We show by induction on t that $\{w : b_{w,t}\} \subseteq \mathrm{cl}(S)$. For $t = 0$ this is (i). If $b_{w,t}$ holds then by (ii) either $b_{w,t-1}$ holds (induction) or some $f_{b,w,t}$ holds; by (iii) b and all of $N(b) \setminus \{w\}$ are black at round $t-1$, hence in $\mathrm{cl}(S)$ by induction. If $w \notin \mathrm{cl}(S)$ then $b \in \mathrm{cl}(S)$ would have $N(b) \setminus \mathrm{cl}(S) = \{w\}$ and could force at the fixed point $\mathrm{cl}(S)$, which is impossible; so $w \in \mathrm{cl}(S)$. By (iv), $V \subseteq \mathrm{cl}(S)$. *Completeness.* Let S be zero forcing with $|S| \le s$ and run the synchronous process $B_0 = S$, $B_t = \Phi(B_{t-1})$ with Φ as in the proof of Proposition 3.4. Set $b_{w,t} = [w \in B_t]$ and $f_{b,w,t} = [b \in B_{t-1},\ N(b) \setminus B_{t-1} = \{w\}]$; (i)–(iii) hold by construction and (v) by $|S| \le s$. Each B_t is reachable by a chronology (perform the available forces one at a time; by Step 1 of Proposition 3.4 each remains legal), so $B_t \subseteq \mathrm{cl}(S) = V$, and while $B_{t-1} \ne V$ some force is available (otherwise B_{t-1} would be a fixed point $\ne V$ above S), so $|B_t| \ge |B_{t-1}| + 1$. Hence $B_t = V$ for $t \ge 2n - |S|$, and $2n - |S| \le 2n - s = T$ gives (iv). □

The horizon $T = 2n - s$ is exact in the sense that every forcing process from a set of size s terminates within T rounds; the propagation time of Hogben et al. (2012) is typically far smaller, but the exact horizon keeps the encoding complete without further argument.

B Computational protocol

Hardware and software. All computations ran on a single MacLustr worker node (identifier `M4M64c`; Apple Silicon, 16 CPU cores, macOS 26.4). Python 3.14.7; PySAT 1.9.dev17 with CaDiCaL 1.9.5 (boundary and fort-cover formulas, DRAT logging) and Glucose 4.1 (round encoding); `drat-trim` built from the GitHub source fetched on 2026-10-30 (source SHA-256 `d834b649...a0c2ee`, full digest in the result records); SymPy 1.14.0 for the symbolic identities; C programs compiled with `gcc -O2`. No GPU, no cluster scheduler and no external service were used; all reported timings are wall-clock times measured on this node, and CPU times where stated.

Job structure and limits. One process per graph order n with a bounded pool of workers (6 for the boundary table, 5–12 for the certified window, 16 threads inside each C enumeration). Limits: exhaustive oracle 3,600 s per n; fort-cover 600 s (holdout searches) to 1,200 s (exact runs); round encoding 600 s (one rerun at $n = 17$ without limit). A timed-out solver is recorded as *unknown*. No boundary-window instance approached its limit (maximum 266.8 s).

Randomness. The only random inputs are the seed sets and force orders of the order-independence test (seed $20261030 \cdot 1000 + n$), the random instances of the lemma checks, and the simulated-annealing attacks (seed 20261030); all are inputs to deterministic checks and are recorded. Nothing in this study is simulated or synthetic data in the statistical sense: all objects are generated from Definition 3.1.

Records. Every solver run is a JSON record with UTC start time, n, k, graph digest, method, bound, solver name and version, wall time, timeout flag, witness or certificate paths and their SHA-256 digests (`results/solver_runs.jsonl`, `results/boundary_certificates.jsonl`, `results/`

lower_bound_certificates.jsonl, results/brute_force_runs.jsonl); the independent checker's verdicts are in results/lower_bound_verification.json and results/upper_bound_verification.json.

C Reproducibility and artifact layout

The released directory contains code/ (graph constructors, closure oracle, both C enumerators, the SAT drivers, the independent checkers, the figure and table generators, reproduce.sh and pinned requirements.txt), code/proofs/handwritten/ (the step-numbered proofs), code/proofs/certificates/ (DIMACS files, fort lists and, where shipped, DRAT proofs), results/ (all tables, records, the counterexample archive, derivations.md, robustness.md and metrics.json), and plan/ (the pre-registered plan with the frozen holdout, the prior-art audit and conjectures.jsonl). Running shcode/reproduce.sh from the root makes a working copy of the code and of the small result files under results/reproduction/work/, rebuilds the tools there, re-runs the invariant and lemma checks, replays every witness, re-verifies shipped certificates whose proof files are present, regenerates a bounded part of the computation with fresh proofs, and finally compares the working copy with the shipped files (results/reproduction/summary.json); the shipped artifacts are never modified. Every step is logged with its command, start and end time and exit code. The three modes regenerate different parts of the tables, and we state them precisely because the DRAT proofs of the boundary window (39.0 GB) and the two large fort-cover proofs (4.0 GB at $n = 23$, 7.7 GB at $n = 25$) may be absent from distributed copies of the archive (their digests are in the result records and in Tables 9 and 10; a proof that is not on disk is reported as *skipped*, never as verified):

- QUICK=1 (about ten minutes): steps 0–3, 9 and 10. It re-verifies with drat-trim the shipped fort-cover proofs for $n \le 21$ and a fixed sample of the boundary proofs ($n = 24, 26, \dots, 40, 60$) when they are present, and re-runs the robustness checks that need no enumeration.
- default (about 25 minutes on 16 cores; bounded so that it stays well inside the platform's one-hour limit): adds the exhaustive oracle for $11 \le n \le 22$ (rows 11–22 of Table 2), the boundary table for $16 \le p \le 20$, the certified window for $26 \le n \le 40$ regenerated with fresh DRAT proofs (each checked; the regenerated CNF must have the digest of the shipped one), the fort-cover certificates for $n \le 21$ (the round encoding only for $n \le 17$, with a 60 s limit), and a re-check of the regenerated certificates.
- FULL=1 (about six further hours): adds the exhaustive oracle at $23 \le n \le 27$ (the last five rows of Table 2), the complete boundary table, the complete window $26 \le n \le 154$ (about 35 minutes on 12 cores), the fort-cover certificate at $n = 22$ (about eleven minutes), a drat-trim re-check of *every* shipped proof present on disk including the two large fort-cover proofs (about three hours), the proof-logging re-solves of the stored fort-cover formulas for $n = 23$ and $n = 25$ (column F of Table 1 at these n; two to three hours), the round encoding for $18 \le n \le 22$ with its original 600 s limit, the second enumerator E′ through $n = 40$ (column E′ of Tables 1 and 2 for $n \ge 23$) and the exhaustive boundary-encoding checks of Section 8(c).

An interrupted step leaves the corresponding result rows unchanged; a killed or timed-out solver is never recorded as UNSAT. The outcome of our own runs of the default and QUICK modes, and of the platform's earlier clean re-execution, is described in manuscript/reproducibility.md.

D Scope of machine checking

Table 6 states, for each component of the proof, what is proved by hand for all n, what is a machine-checked finite input and with which checker, and whether any formal (proof-assistant) verification

exists. There is none; the claim "verified" in this paper always means "verified by `drat-trim` and re-verified by an independent re-implementation", never "formally verified".

Table 6. Human proof, machine-checked input and formal verification for each component.

statement	human proof	machine-checked input	Lean 4
P1–P5 (closure, forts, symmetry, SAT encoding, graph facts)	complete, all n	empirical companions only	no
P6–P9 (isomorphism, 12-vertex bound, first-force lemma, column deletion)	complete, all n	symbolic identity (sympy); replays $12 \le n \le 300$	no
Lemmas 3.1–3.5 (boundary machinery, CNF correctness)	complete, all n, p	integer bookkeeping $12 \le p \le 30$	no
$Z \ge 12$ for $26 \le n \le 154$	via Lemma 3.4	130 DRAT proofs, drat-trim exit 0	no
$Z \ge 12$ for $n > 154$	Lemmas 3.2–3.3 (no computation)	—	no
$Z(P(n,5))$, $11 \le n \le 27$	P3 (orbit reduction)	exhaustive enumeration, two independent programs	no
$Z(P(n,5)) \ge$ table, $11 \le n \le 22$	Lemma 3.5, P2	12 fort-cover DRAT proofs	no
upper bounds, all n	P7	replayed chronologies	no

E Additional tables

Table 7. One minimum zero-forcing set for each $11 \le n \le 22$ (the first D_n-orbit representative found by the exhaustive oracle) and the number of forcing orbit representatives at size Z. Each set is stored with a complete chronology and was replayed by the independent checker. Only $n = 15$ needs inner vertices in the listed representative; for $n \ge 17$ ten or eleven consecutive outer vertices (with one gap for $n = 17, 18$) suffice, and for $n = 11$ two triples of consecutive outer vertices do.

n	Z	a minimum zero-forcing set (first orbit representative found)	forcing reps. at size Z
11	6	$\{u_0, u_1, u_2, u_5, u_6, u_7\}$	134
12	8	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$	2,514
13	9	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$	12,820
14	8	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$	2,072
15	8	$\{u_0, u_{10}, u_{11}, u_{12}, u_{13}, v_0, v_1, v_3\}$	4,105
16	8	$\{u_0, u_1, u_2, u_3, u_4, u_7, u_8, u_9\}$	1,977
17	10	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_7, u_8, u_9, u_{10}\}$	108,897
18	10	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_8, u_9, u_{10}\}$	28,517
19	10	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9\}$	96,131
20	10	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9\}$	62,384
21	10	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9\}$	60,192
22	11	$\{u_0, u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, u_{10}\}$	524,638

F Note on the autonomous research process

This working paper was produced by an autonomous research pipeline under a fixed brief. The scientific content (definitions, proofs, encodings, programs and the interpretation of results) was generated and checked within that pipeline; the pre-registration of the fitting range and holdout, the three-valued certification vocabulary, the independent re-implementation of every checker, and the requirement that every reported number be traceable to a stored artifact were imposed as process rules before computation began. The literature audit verified the identifiers of all 20 cited sources and recorded the reading depth of each. The MacLustr compute infrastructure is credited for computation only. Nothing in this appendix is a mathematical claim.

Table 8. The boundary table $\beta_p(n)$ for $20 \le n \le 30$ (the full range $13 \le n \le 30$ is in Figure 3 and `results/tables/boundary_table.csv`). Bold entries are certified refutations of $\mathcal{F}(n, p)$; all other entries are exact values with a recounted witness.

$p \backslash n$	20	21	22	23	24	25	26	27	28	29	30
12	8	9	10	10	10	10	10	10	10	10	10
13	8	9	10	10	10	10	10	10	10	10	10
14	8	10	10	10	10	10	10	10	10	10	10
15	8	10	10	11	11	10	11	11	11	11	11
16	8	10	10	11	11	10	11	11	11	11	11
17	8	10	11	11	11	10	11	11	11	11	11
18	8	10	10	11	**≥ 12**	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
19	8	10	10	11	**≥ 12**	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
20	8	9	10	11	11	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
21	8	9	10	11	11	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
22	8	9	10	10	11	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
23	8	9	10	10	11	10	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**
24	8	8	9	10	10	10	11	**≥ 12**	**≥ 12**	**≥ 12**	**≥ 12**

References

Joseph S. Alameda, Emelie Curl, Armando Grez, Leslie Hogben, O'Neill Kingston, Alex Schulte, Derek Young, and Michael Young. Families of graphs with maximum nullity equal to zero forcing number. *Special Matrices*, 6(1):56–67, 2018. doi: 10.1515/spma-2018-0006. URL https://doi.org/10.1515/spma-2018-0006.

Francesco Barioli, Wayne Barrett, Shaun M. Fallat, H. Tracy Hall, Leslie Hogben, Bryan Shader, P. van den Driessche, and Hein van der Holst. Zero forcing parameters and minimum rank problems. *Linear Algebra and its Applications*, 433(2):401–411, 2010. doi: 10.1016/j.laa.2010.03.008. URL https://doi.org/10.1016/j.laa.2010.03.008.

Beth Bjorkman, Lei Cao, Franklin Kenter, Ryan Moruzzi, Carolyn Reinhart, and Violeta Vasilevska. Leaky forcing: Extending zero forcing results to a fault-tolerant setting. *arXiv preprint arXiv:2508.02564*, 2025. URL https://arxiv.org/abs/2508.02564.

Boris Brimkov, Caleb C. Fast, and Illya V. Hicks. Computational approaches for zero forcing and related problems. *European Journal of Operational Research*, 273(3):889–903, 2019. doi: 10.1016/j.ejor.2018.09.030. URL https://doi.org/10.1016/j.ejor.2018.09.030.

Boris Brimkov, Derek Mikesell, and Illya V. Hicks. Improved computational approaches and heuristics for zero forcing. *INFORMS Journal on Computing*, 2021. doi: 10.1287/ijoc.2020.1032. URL https://doi.org/10.1287/ijoc.2020.1032.

Thomas R. Cameron, Leslie Hogben, Franklin H. J. Kenter, Seyed Ahmad Mojallal, and Houston Schuerger. Forts, (fractional) zero forcing, and Cartesian products of graphs. *arXiv preprint arXiv:2310.17904*, 2023. URL https://arxiv.org/abs/2310.17904.

Randy Davila and Franklin Kenter. Bounds for the zero forcing number of graphs with large girth. *Theory and Applications of Graphs*, 2(2), 2015. doi: 10.20429/tag.2015.020201. URL https://doi.org/10.20429/tag.2015.020201.

Caleb C. Fast and Illya V. Hicks. Effects of vertex degrees on the zero-forcing number and propagation time of a graph. *Discrete Applied Mathematics*, 250:215–226, 2018. doi: 10.1016/j.dam.2018.05.002. URL https://doi.org/10.1016/j.dam.2018.05.002.

Table 9. Fort-cover certificates $\mathcal{G}(n, Z-1)$ for $11 \le n \le 23$ (path F): size of the fort list at termination, CNF size, DRAT proof length and CaDiCaL time of the proof-logging re-solve. Every certificate was verified by `drat-trim`; those for $n \le 22$ were also re-verified by the independent checker, which re-validates every listed fort (the $n = 23$ certificate was checked by `drat-trim` through a plain subprocess, Observation 7.7). *n/r: the proof-logging re-solve at $n = 23$ ran as a separate process that ended before its record was written, so no solve time is recorded.

n	bound $b = Z-1$	forts in $\mathcal{L}$	variables	clauses	proof lines	CaDiCaL (s)
11	5	187	111	386	401	0.00
12	7	650	147	915	12,310	0.05
13	8	1,404	174	1,719	38,256	0.17
14	7	1,344	179	1,669	25,506	0.13
15	7	930	195	1,285	9,359	0.05
16	7	2,372	211	2,757	45,055	0.26
17	9	7,684	263	8,167	508,267	3.86
18	9	8,832	283	9,353	808,609	9.43
19	9	11,571	303	12,130	1,168,686	16.68
20	9	16,290	323	16,887	924,560	11.39
21	9	22,001	343	22,636	2,229,150	21.48
22	10	37,521	388	38,242	12,952,380	206.70
23	11	87,239	435	88,050	59,527,561	n/r*

Roberto Frucht, Jack E. Graver, and Mark E. Watkins. The groups of the generalized Petersen graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 70(2):211–218, 1971. doi: 10.1017/s0305004100049811. URL https://doi.org/10.1017/s0305004100049811.

AIM Minimum Rank – Special Graphs Work Group. Zero forcing sets and the minimum rank of graphs. *Linear Algebra and its Applications*, 428(7):1628–1648, 2008. doi: 10.1016/j.laa.2007.10.009. URL https://doi.org/10.1016/j.laa.2007.10.009.

Leslie Hogben, My Huynh, Nicole Kingsley, Sarah Meyer, Shanise Walker, and Michael Young. Propagation time for zero forcing on a graph. *Discrete Applied Mathematics*, 160(13-14):1994–2005, 2012. doi: 10.1016/j.dam.2012.04.003. URL https://doi.org/10.1016/j.dam.2012.04.003.

Leslie Hogben, Jephian Lin, and Bryan Shader. Inverse problems and zero forcing for graphs. *Mathematical Surveys and Monographs*, 2022. doi: 10.1090/surv/270. URL https://doi.org/10.1090/surv/270.

Alexey Ignatiev, Antonio Morgado, and Joao Marques-Silva. PySAT: A Python toolkit for prototyping with SAT oracles. In *Lecture Notes in Computer Science*, pages 428–437. Springer International Publishing, 2018. doi: 10.1007/978-3-319-94144-8_26. URL https://doi.org/10.1007/978-3-319-94144-8_26.

Arnav Krishnan. A correction to the zero forcing number of the generalized Petersen graphs $P(n, 3)$. *arXiv preprint arXiv:2607.19412*, 2026. URL https://arxiv.org/abs/2607.19412.

Arnav Krishnan, Haobo Ma, and Wenlin Zhang. Sharp zero forcing numbers for the generalized Petersen graphs $P(n, 3)$ and $P(n, 4)$. *arXiv preprint arXiv:2610.24059*, 2026. URL https://arxiv.org/abs/2610.24059.

Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Lecture Notes in Computer Science*, pages 625–635. Springer International Publishing, 2021. doi: 10.1007/978-3-030-79876-5_37. URL https://doi.org/10.1007/978-3-030-79876-5_37.

Table 10. Boundary certificates $\mathcal{F}(n, 18)$ at selected n in the certified window (all 129 rows: `results/tables/boundary_window_p18.csv`). Digests are the first 16 hex characters of the SHA-256 of the DIMACS file and of the DRAT proof.

n	variables	clauses	CaDiCaL (s)	proof lines	drat-trim (s)	CNF SHA-256	proof SHA-256
26	1,779	3,589	20.8	1,528,402	18.1	8b2b39c3769f94b2	3d1a3b3652f8e732
30	2,171	4,381	20.5	1,440,157	16.4	c526d49854f86414	e2656fd1b6e21c20
40	3,151	6,361	23.2	1,438,038	19.0	108bc245d1983d35	5f3a371fc8fb20c3
50	4,131	8,341	27.3	1,367,546	28.9	925b74a12023475f	cf91a7b47240bf41
60	5,111	10,321	31.6	1,437,601	40.4	7f5d6d4cc0c34a73	a2ff4fe669cd7e2b
70	6,091	12,301	49.1	1,751,245	42.4	41cbe28199c05b53	01aebd7b76c4e379
80	7,071	14,281	45.4	1,780,669	50.9	6ef97ab92c165054	a52be3cb21ad3a59
90	8,051	16,261	115.2	2,796,452	103.0	217a9f570f340a2d	1ac406abdd73beee
100	9,031	18,241	107.7	3,132,080	95.0	9716f6fe9658d889	defbfdb12994c8dc
110	10,011	20,221	110.1	3,281,673	104.9	90cc4aceefceca54	85b0a09a20712e6e
120	10,991	22,201	143.6	3,669,833	153.9	0a5576422e6c22f1	4ef95714eacd70ae
130	11,971	24,181	179.1	4,107,598	151.3	705fda01008e9e47	3259fd182fabf86c
140	12,951	26,161	231.9	4,137,876	177.3	8d53e0c5734cb054	5058e939dc808016
150	13,931	28,141	204.5	4,545,857	168.6	c44db207409a5796	4d9c5a910260bbe2
154	14,323	28,933	226.5	4,662,907	197.2	fa15c4c63db1c686	41ea487bdc464bf0

Saeedeh Rashidi, Nosratollah Shajareh Poursalavati, and Maryam Tavakkoli. Computing the zero forcing number for generalized Petersen graphs. *Journal of Algebra Combinatorics Discrete Structures and Applications*, 7(2):183–193, 2020. doi: 10.13069/jacodesmath.729465. URL https://doi.org/10.13069/jacodesmath.729465.

Carsten Sinz. Towards an optimal CNF encoding of Boolean cardinality constraints. In *Lecture Notes in Computer Science*, pages 827–831. Springer Berlin Heidelberg, 2005. doi: 10.1007/11564751_73. URL https://doi.org/10.1007/11564751_73.

Mark E. Watkins. A theorem on Tait colorings with an application to the generalized Petersen graphs. *Journal of Combinatorial Theory*, 6(2):152–164, 1969. doi: 10.1016/s0021-9800(69)80116-x. URL https://doi.org/10.1016/s0021-9800(69)80116-x.

Nathan Wetzler, Marijn J. H. Heule, and Warren A. Hunt. DRAT-trim: Efficient checking and trimming using expressive clausal proofs. In *Lecture Notes in Computer Science*, pages 422–429. Springer International Publishing, 2014. doi: 10.1007/978-3-319-09284-3_31. URL https://doi.org/10.1007/978-3-319-09284-3_31.

AI-generated working paper; not formally peer reviewed. This manuscript was produced autonomously by AutoScholar.ai's research pipeline (large language model agents operating under platform-enforced research standards) and passed deterministic quality gates, including citation verification, result traceability and an automated critical review. It has *not* undergone independent human peer review. Readers should treat the findings as preliminary and verify claims against the released artifacts. Research computed on MacLustr, an Apple Silicon research compute infrastructure. Artifacts, provenance and corrections: https://autoscholar.ai.