FROM LOCAL REWARDS TO RECURSIVE GAINS: HYBRID AI SEARCH FOR C_4-FREE HYPERCUBES

Anonymous authors
Paper under double-blind review

ABSTRACT

Erdős Problem #86 asks how many edges of the n-dimensional hypercube can be retained without forming a 4-cycle. AI-assisted search for recursive constructions must decide both where to search in this exponentially large space of coupled constraints and what to value, since immediate edge count can misrank seeds by their recursive potential. We present a hybrid system with two modes over compact symmetry-quotient states. A frozen language model proposes typed search actions that constraint programming executes and an independent verifier certifies; a low-dimensional mode saturates within one dimension and hands its verified seed pool to a high-dimensional mode, which a jump cycle reaches on already-computed recursive values. Both modes share the proposer, the verifier and a fitted head; the high-dimensional mode adds an analytic, exactly computed recursive advantage under the classical Brass–Harborth–Nienborg (BHN) lift to the score, and these values feed back into repair objectives and retention. The archive shows that a recursive state does not fix graph structure and that seeds with fewer edges can produce stronger descendants: a searched Q_{15} seed starts 1,024 edges below the classical seed yet exceeds its same-template rollout by 12,288 edges after one lift and stays ahead at the next two depths. Combining complementary seed lineages yields a construction family that provably matches or exceeds the refined BHN benchmark for every $n \geq 7$. Across $9 \leq n \leq 16{,}384$, six independently verified seeds realize the archive envelope, with strict improvements in 13,805 dimensions (84.3%). These results connect structural exploration to high-dimensional construction guarantees through template-specific recursive value.

1 INTRODUCTION

Hybrid AI systems for mathematical construction must coordinate learned search decisions with reliable evaluation. When a candidate is a seed for a recursive construction, immediate quality need not predict the quality of its descendants. We study this problem in Erdős Problem #86: finding C_4-free subgraphs of the hypercube Q_n with many edges (Erdős Problems, n.d.). Our goal is to discover verifiable seeds whose recursive descendants improve the classical Brass–Harborth–Nienborg (BHN) benchmark (Brass et al., 1995).

We build on AI-assisted discovery with executable evaluators (Romera-Paredes et al., 2024; Lange et al., 2026) and learned search decisions (Gasse et al., 2019). Our focus is template-specific recursive value: the same mathematical recurrence constructs descendants and evaluates the seeds from which they grow.

Two challenges organize the system. *Where should it search?* Overlapping C_4 constraints couple individual edge decisions. Symmetry quotients expose edge orbits as decision units, with exact constraints within a chosen symmetry class; controlled symmetry release enlarges the represented space. *What should it value?* Under the BHN lift, descendant edge counts depend on the directional profile as well as the current count. A searched Q_{15} seed starts $1{,}024$ edges below the classical reference yet produces $12{,}288$ more edges after one identical lift. Retaining only the seed with more current edges would prefer the weaker recursive starting point.

We present a *hybrid AI system* connecting these decisions through feedback (Figure 1). A frozen language model proposes parameters for typed search actions over quotient states. Constraint program-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

ming executes admissible proposals, and an independent verifier checks their outputs. The archive records outcomes and recursive states (n, E, P): dimension, edge count, and sorted directional profile. Archived outcomes support fitted value estimates; the BHN recurrence separately evaluates fixed-template continuation without further solver calls. Recursive values inform candidate retention and repair objectives, connecting exploration, verified execution, and analytic evaluation.

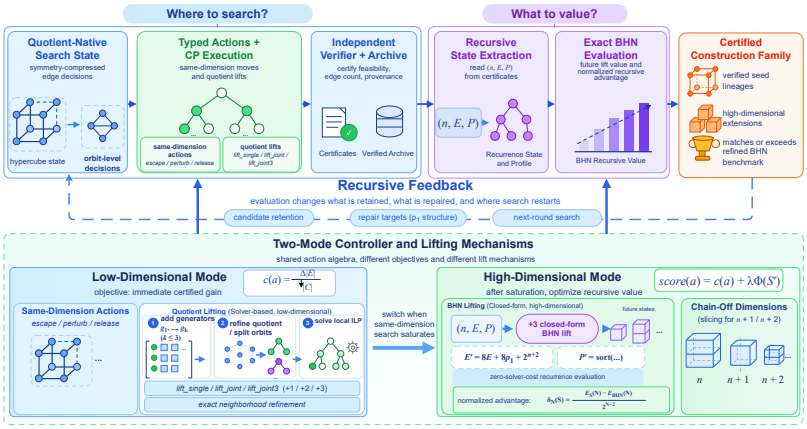


Figure 1: A frozen language model proposes typed search actions; constraint programming and independent verification produce certified seeds. BHN recursive evaluation feeds back into candidate retention and repair objectives. Bounded protocols test selected feedback choices.

Our contributions are:

- **A representation for verifiable search.** Quotient states, typed operators, and controlled symmetry release provide an interface for AI-assisted exploration within explicitly defined feasible search spaces (Sections 2–3).
- **Recursive evaluation as search feedback.** We show that distinct graph realizations can share a recursive state and that immediate edge count can misrank seeds. These findings motivate profile-based evaluation for retention and repair (Sections 4–6).
- **A certified construction family.** We prove that the family matches or exceeds refined BHN for every $n \geq 7$. Across $9 \leq n \leq 16{,}384$, six independently verified roots realize an archive envelope with 13,805 strict improvements (84.3%), evaluated by recurrence and coordinate-restriction bounds (Section 7).

The construction guarantees are independent of controller efficiency: they hold for the certified family whichever controller produced the seed.

2 The Representation Problem: States an AI Can Operate On

Exponential size and coupled constraints. The hypercube has 2^n vertices and $n2^{n-1}$ edges, giving $2^{245,760}$ candidate edge subsets already at $n = 15$. Every C_4 is a coordinate square, and each edge belongs to $n - 1$ such squares, so feasibility couples local edits across the graph. We expose symmetry orbits as the decision units of AI-assisted search.

Quotient states. Choose a full-row-rank binary matrix $H \in \mathbb{F}_2^{r\times n}$, with kernel $C = \ker H$ and columns $g_i = He_i$, distinct and nonzero. Translation by C partitions vertices into 2^r syndrome classes; an *edge orbit* in direction i is the unordered pair $\{A, A + g_i\}$ with its direction label. A *quotient state* $z = (H, F)$ selects a set F of these labeled orbits, and its lift is

$$\mathcal{E}(\Phi_H(F)) = \big\{\{x, x + e_i\} : (\{Hx, Hx + g_i\}, i) \in F\big\}. \tag{1}$$

Every orbit contains $|C|$ edges, so $E(\Phi_H(F)) = |C|\,|F|$ and the state carries only $n2^{r-1}$ binary decisions: one decision controls a whole orbit of $|C|$ edges instead of one edge. After duplicate

square constraints are removed, both the decision and the constraint counts are reduced by exactly $|C|$ (Figure 2; Appendix A).

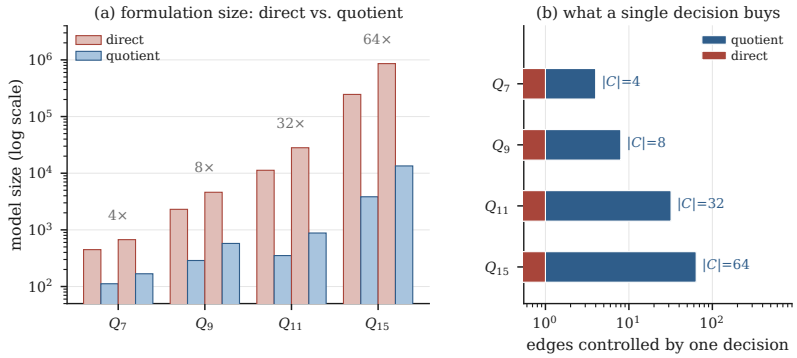


Figure 2: Quotient compression. **(a)** One binary decision controls an entire edge orbit of $|C|$ edges **(b)**, reducing decision and square-constraint counts by $|C|$. Formulation size, not solver speed.

Exact constraints, exact edge counts. Let $u_{A,i} = u_{A+g_i,i} \in \{0,1\}$ denote the single decision for an unordered edge orbit. For every syndrome A and pair $i < j$, impose

$$u_{A,i} + u_{A+g_j,i} + u_{A,j} + u_{A+g_i,j} \leq 3. \tag{2}$$

For a cube vertex x with $Hx = A$, these are exactly the four edges of its (i,j)-square, so the constraints are necessary and sufficient for a C-invariant lift to be C_4-free: within the chosen symmetry class the encoding preserves feasibility and edge count *exactly*.

Compression versus expressivity. A quotient state is not an isomorphism quotient: it covers exactly the C-invariant graphs of one chosen symmetry class — not all graphs in Q_n nor all isomorphic copies — and requires only $C \leq \mathrm{Stab}(E)$. To recover freedom, the controller can *release* symmetry: replacing C by an index-two subgroup splits each orbit into two independently selectable orbits while keeping the incumbent representable. It can therefore search within a compressed class or release symmetry to enlarge it.

3 The Hybrid AI System over Quotient States

The controller of the hybrid system is the tuple

$$\Pi = (\theta,\ \hat{V},\ \mathcal{G}) \tag{3}$$

— a frozen language-model proposer θ, a fitted value head $\hat{V}$, and an exact verifier gate $\mathcal{G}$ — plus the typed operator algebra $\mathcal{O}$ and two mutable artefacts, a mechanism library $\mathcal{L}_t$ and an outcome archive $\mathcal{A}_t$. This section presents the system's *lower-dimensional warm-up*: within a fixed dimension, θ proposes action parameters, CP-SAT executes admissible proposals, and $\mathcal{G}$ certifies every output before it enters the archive that the high-dimensional mode of §6 searches from. The proposer weights remain fixed; the head is fitted to executed outcomes (Appendix L).

One round: propose, execute, verify, select. The round dispatches into the action families of Table 1. Given a certified state, its representation, the archive, and the remaining budget, θ emits a batch of typed actions. The prompt requests coverage of the operator kinds; the model chooses their parameters, including directions for lifting and perturbation choices for repair. A pre-solve check rejects unknown kinds, malformed parameter blocks, and inadmissible values. Every admissible proposal is executed under its recorded solver limit. Only FEASIBLE or OPTIMAL outputs provide candidates for independent verification; only the latter certifies a model optimum. Accepted outputs retain their parent basename and hash, parameters, status, cost, and recursive fingerprint (n, E, P). Unresolved attempts retain status records without contributing graphs. Successor selection uses these verified outcomes; the archive also retains complementary and intermediate states.

Table 1: Action families in the hybrid system. Full parameter schemas are in Appendix B.

Family	Operation	Representation effect
Repair	Perturb or exclude orbit choices	Fixed n, H
Symmetry release	Split orbits and re-optimize	Fixed n, higher rank
Quotient recursion	Append $k \in \{1, 2, 3\}$ directions	Higher n, fixed rank

An exact recursive neighborhood. Appending k columns preserves 2^r quotient vertices and multiplies the kernel size by 2^k. Binary r_e deletes a selected old orbit and y_p selects a new-direction orbit; maximize $\sum_p y_p - \sum_{e \in F} r_e$.

Proposition 1 (Exactness within the recursive-lift neighborhood)**.** *Let (H, F) be certified, with nonzero, pairwise distinct augmented columns. Allow deletions from F and additions only in new directions. Impose $y_p + y_q \leq 1 + r_a + r_b$ on every old–new square whose old orbits a, b lie in F, and $y_1 + y_2 + y_3 + y_4 \leq 3$ on every new–new square. The feasible solutions are exactly the C_4-free states in this neighborhood, with optimum $\max_{F'}(|F'| - |F|)$.*

Appendix C gives the proof. Exactness of this formulation does not certify optimality or improvement from a time-limited feasible solve.

Two modes, one learned search interface. For same-dimension actions, immediate improvement in parent-orbit units is

$$c(a) = \frac{|E(S')| - |E(S)|}{|C|}, \qquad n(S') = n(S). \tag{4}$$

The immediate-only setting ranks verified outcomes by c; a bootstrap variant adds $\gamma \hat{V}(S')$. Both use the same proposer, operator library, verifier, and archive. The high-dimensional mode additionally evaluates $\Phi(S')$ under legal BHN continuation (Section 6); the fitted head estimates search return, whereas Φ is computed analytically. Thus learning guides search choices while exact mathematics evaluates the constructions they produce (Figure 3).

Failure to improve within the tested actions and limits triggers an operational stop, not a proof of representation exhaustion. A dimension jump selects among already verified successors in $n+1, n+2, n+3$ by their common-target recursive value. Selection needs no additional solve; generating the candidates still costs solver time. Promotion criteria and implementation details are in Appendix R.

Outputs of the search loop. Table 13 reports certified outputs of the two head settings alongside the archived baseline. The $\gamma = 0.5$ run yields 371 distinct recursive states versus 338 at $\gamma = 0$; neither exceeds an archived dimension record. These states are the interface to the next stages: Discovery I studies what their fingerprints determine, Discovery II shows why edge count can misrank them, and Section 6 uses their recursive value for feedback.

4 Discovery I: A Recursive State, Many Realizations, Many Routes

The BHN recurrence does not read graphs; it reads a numerical fingerprint. Three certified examples delimit what this fingerprint fixes and what it does not.

A searched graph with the classical state. The searched Q_7 graph has $E = 304$ and sorted direction profile $P = [48^{\times 3}, 40^{\times 4}]$, matching the classical reference. Yet it is non-isomorphic to that reference and lies outside the tested standard $Q_4 \to Q_7$ lift image, which the classical graph passes as a positive control (Appendix E). Thus the archive distinguishes a graph's recursive state from its construction provenance.

Fix a classical reference chain at dimensions $n = 4 + 3t$, with states $(E_n^{\mathrm{ref}}, P_n^{\mathrm{ref}})$. Matching this state is a purely numerical condition — the BHN *state family* (Definition 4, Appendix F) — weaker than membership in the *construction family*, the graphs reached by the iterated classical template from Q_4. The Q_7 example separates the two. Does sharing (E, P) fix what comes next?

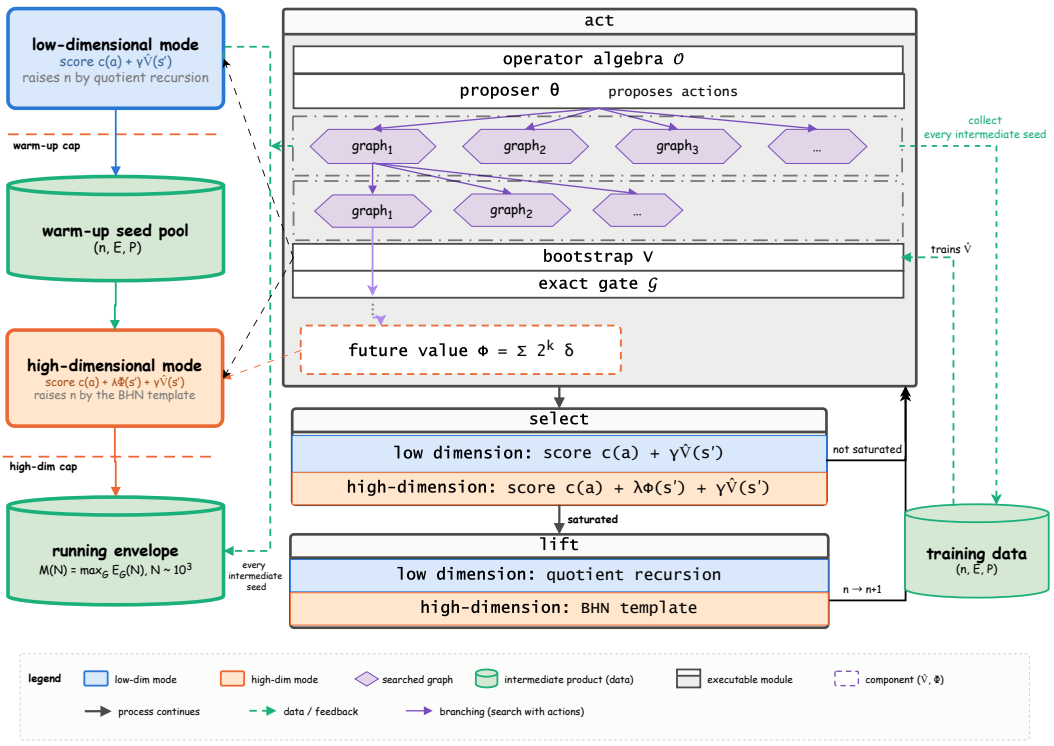


Figure 3: The controller. *Left*: the modes run in series, and the warm-up seed pool passes between them; each has its own cap, past which only lifting continues. *Right*: the mechanism both modes share — the proposer opens a batch of candidate graphs, the value head is bootstrapped from the archived rows, the exact gate certifies each graph, and the recursive advantage Φ returns into the score. The modules ACT, SELECT and LIFT form the control loop; the two-coloured variants are the parts that switch with the mode.

The recurrence reads only the state. It does. The maximum-direction BHN lift consumes only the *recursive state* $S(G) := (n, E(G), P(G))$, with P the sorted direction profile: choosing a maximum-load direction p_1,

$$\begin{aligned} E' &= 8E + 8p_1 + 2^{n+2}, \\ P' &= \operatorname{sort}\big(8p_2, \ldots, 8p_n, (2^n + 4p_1)^{\times 4}\big). \end{aligned} \tag{5}$$

Graph-level dynamics can branch—the three maximum directions at Q_7 give children that are not all cube-equivalent—yet the state dynamics is deterministic: every branch shares (E, P) (Appendix E).

Proposition 2 (State determinism and closure)**.** *If* $(E(G), P(G)) = (E_n^{\mathrm{ref}}, P_n^{\mathrm{ref}})$, *then* $(E(T_d(G)), P(T_d(G))) = (E_{n+3}^{\mathrm{ref}}, P_{n+3}^{\mathrm{ref}})$ *for every maximum-load direction* d *of* G*: Equation (5) reads only* (n, E, P)*, ties among maximum directions do not change the sorted output multiset, and the template preserves* C_4*-freeness (Brass et al., 1995).*

Entry through an external lineage. Closure describes trajectories that begin inside; entry need not. A joint quotient extension of the searched Q_7 seed yields a certified Q_9 root off the reference chain entirely: $E = 1{,}472$, 32 edges below the refined benchmark. Under strict BHN lifts and one coordinate slice, its trajectory stays below the benchmark at every lift and lands exactly on it:

$$\underbrace{1{,}472}_{Q_9} \xrightarrow{\text{lift}} \underbrace{15{,}360}_{Q_{12}} \xrightarrow{\text{lift}} \underbrace{151{,}552}_{Q_{15}} \xrightarrow{\text{slice}} \underbrace{33{,}792}_{Q_{13}}, \qquad \begin{array}{c} 1{,}472 < 1{,}504 \\ 15{,}360 < 15{,}616 \\ 151{,}552 < 153{,}600 \\ 33{,}792 = \text{ref}. \end{array} \tag{6}$$

The slice matches the classical Q_{13} reference state, while structural invariants separate it from the archived classical realization (Appendix G). Membership is thus attainable from outside: the BHN

lift with coordinate slicing is itself a route into the family, and subsequent lifts stay in it by Proposition 2.

What the archive must retain. Initial state-family membership is thus sufficient for closure, but is not required for eventual entry when slicing is also allowed. For the AI archive, (n, E, P) provides an exact interface for BHN evaluation without identifying graph structure or provenance. The remaining question is whether E alone is enough to rank states for extension.

5 Discovery II: Edge Count Misranks Recursive Potential

Discovery I establishes that (n, E, P) determines the BHN trajectory. For the AI controller, does the simpler score E suffice to decide which seeds to retain and extend? Two certified Q_{15} seeds show that it does not: better immediate quality can imply worse value under the same recursive template.

A controlled ranking reversal. The searched seed G_{s} has $E = 152{,}576$ and maximum direction load $p_1 = 12{,}800$; the classical reference G_{c} has $E = 153{,}600$ and $p_1 = 10{,}240$. Current edge count therefore prefers G_{c} by $1{,}024$ edges. Under the same maximum-direction BHN lift, Equation equation 5 gives, at $n = 15$,

$$J_1(G) = 8\big(E(G) + p_1(G)\big) + 2^{17}. \tag{7}$$

The searched seed's $2{,}560$-edge advantage in p_1 outweighs its $1{,}024$-edge deficit in E, yielding $J_1(G_{\mathrm{s}}) - J_1(G_{\mathrm{c}}) = 12{,}288$ at Q_{18}. It also exceeds the distinct refined BHN benchmark by $4{,}096$ edges. The reversal persists at the next two evaluated lift depths (Appendix H).

From a counterexample to a search signal. An edge-count-only retention rule would prefer the weaker recursive seed in this comparison. For one lift at fixed dimension, $E + p_1$ gives the exact ranking. More generally, let $J_h(G) = E(T^h(G))$, where T applies the maximum-direction BHN lift. Iterating Equation equation 5 evaluates J_h from the archived profile without another solver call; deeper horizons can depend on entries beyond p_1 (the two-step form is derived in Appendix N). These values describe fixed-template continuation, not the best outcome of further search.

This supplies an analytic value signal for the AI controller: the language model proposes search actions, while the recurrence evaluates the extension potential of verified outcomes. Section 6 incorporates this signal into candidate retention and repair objectives. The reversal establishes a limitation of the immediate objective and motivates the recursive signal that the controller scores with.

6 The High-Dimensional Mode: Recursive Value as Search Feedback

The hybrid system's *high-dimensional mode* evaluates what the lower-dimensional warm-up of §3 collects: the BHN recurrence propagates each retained seed's recursive state to dimensions where explicit edge search is impractical, and the resulting values feed back into retention, repair objectives, and successor selection (Discovery II supplies the evaluation target).

A value defined by legal extension. To reach a target N outside a seed's $+3$ chain, use the next lift node $G \subseteq Q_m$, where $m - N = k \in \{0, 1, 2\}$. Deleting the k lightest directions D leaves 2^k coordinate slices, with

$$\max_b E(G|_{x_D=b}) \geq \frac{E(G) - \sum_{i\in D} p_i}{2^k}. \tag{8}$$

Off a chain node the ceiling of this average is a guaranteed bound rather than the exact largest slice count; on a chain node the value is exact. The evaluator needs profile arithmetic only, never another solver call; the scale-free implementation is given in Appendix R. The recursive advantage is

$$\delta_N(S) = \frac{E_S(N) - E_{\mathrm{BHN}}(N)}{2^{N-2}}, \tag{9}$$

with $E_S(N)$ the value the recurrence assigns to S at N. Here, for $4^a \le N < 4^{a+1}$, the refined benchmark is

$$E_{\mathrm{BHN}}(N) = N2^{N-2} + \left\lfloor \frac{2N + 4^{a+1}}{3} \right\rfloor 2^{N-3-a}. \tag{10}$$

Normalization puts gains at different target dimensions on a comparable exponential scale; subtracting the common benchmark centers them without changing rankings at a fixed target. Target dimensions are taken *congruent to* 0 *modulo* 3, the recursion's own stride, as a common evaluation grid.

The objective. The candidate score mixes the immediate gain $c(a)$, the recursive advantage, and a bootstrap of the value still reachable from the successor,

$$\mathrm{score}(a) = c(a) + \lambda\, \Phi(S') + \gamma\, \hat{V}(S'), \tag{11}$$

$$\Phi(S) = \sum_{k=1}^{m} r^k\, \delta_{N_k}(S), \qquad N_k = N_1 + 3(k-1), \tag{12}$$

with $N_1 = 1002$, $m = 12$ and $r = 2$. The nodes are *fixed absolute dimensions*, the same twelve for every candidate, not offsets from a candidate's own dimension: a given N_k must mean the same thing at $n = 11$ and at $n = 17$ or the scores are not comparable, and indexing by lift depth would make "N_5" denote a different dimension per seed. A target N_k is a chain node for a seed S only when $N_k \equiv n(S) \pmod 3$; at other targets equation 8 applies. The ranking power comes from the recurrence itself: δ_{N_k} depends on the full direction profile, not only on the current edge count shown uninformative in §5 and §3. The last term is the fitted head of §3: $\hat{V}$ estimates what a successor still leaves accessible, so a move is scored by what it is, what it becomes under the lift, and what it leaves reachable.

The weights r^k increase with k — a reverse exponential, the opposite of a conventional discount — emphasizing later targets as a design choice.

The loop, and where the feedback enters. The two modes are joined by a saturation-and-jump cycle: execute the batch θ proposes, score by equation 11, saturate, then jump to the successor among $n+1, n+2, n+3$ with the largest Φ (lift outcomes already computed, no additional solve); past the warm-up cap only lifting continues. The evaluation returns into the search at three points, closing the loop of Figure 1: the direction structure the recurrence reveals (the leading load p_1) enters repair objectives; seeds with an advantage at some target dimension are retained instead of discarded; recursive gains and costs inform the next round's states, actions, and budget (the executed protocols are in Appendix I). *This loop is the concrete meaning of the hybrid system*: quotient search explores and reshapes structure, the BHN mechanism evaluates and extends these structures, and the evaluation returns to steer the search.

What the mode yields. Run on the pool the low-dimensional mode hands over, this mode produces verified certificates that dominate the earlier quotient chain at every dimension from 15 to 20 and exceed the refined BHN benchmark at Q_{14}, Q_{18}, Q_{20}, Q_{21} and Q_{22}:

Improvement over the chain is distinct from improvement over the stronger archived envelope: the loop reaches that envelope but does not raise it (Appendix M), and the envelope column above shows the two agreeing to within one edge at every dimension. Since different seeds are strong at different target dimensions, the system retains *complementary lineages*, and their running maximum M forms the construction family of Section 7.

What the components contribute. Each component is removed in turn at equal budget. The value head's screen is scored against the downstream advantage Φ: at $\gamma = 0$ the screen is c itself, so regret against c is zero by construction. Against Φ the screen improves monotonically — normalised top-1 regret $0.594 \to 0.537 \to 0.431$, top-1 hit rate $0.203 \to 0.354$, rank correlation $0.053 \to 0.149$ (Figure 7) — a transferred alignment, since the head is fitted on realised return, never on Φ; $\gamma = 1.0$ is best against downstream value and worst against immediate gain. Proposer and action-library arms and the head's composition are in Appendices M and L.

Table 2: **Truncated**: $14 \le n \le 22$; rows 6–8 are settled and omitted. The full table is Table 11.

n	refined BHN	quotient chain	this work	gain over chain
14	71,680	—	**72,192**	—
15	153,600	152,576	**153,600**	+1,024
16	327,680	321,792	**327,680**	+5,888
17	688,128	676,864	**688,128**	+11,264
18	1,449,984	1,418,240	**1,454,080**	+35,840
19	3,047,424	2,967,552	**3,047,424**	+79,872
20	6,356,992	6,193,152	**6,389,760**	+196,608
21	13,303,808	—	**13,352,960**	—
22	27,787,264	—	**27,852,800**	—

7 Discovery III: Broad and Growing Gains from Hybrid Search

The hybrid system combines retained seed profiles with verified recursive extension into a dimension-wise construction envelope. **Over the audited range $9 \le n \le 16{,}384$, the envelope of the 318 archived states beats the refined BHN benchmark of Eq. equation 10 in $13{,}805$ of the $16{,}376$ dimensions (84.3%), ties it in the remaining $2{,}571$, and never falls below it; by Theorem 1 the same family matches or exceeds the benchmark for every $n \ge 7$.** The gains follow a repeated block structure whose maxima keep growing after the exponential edge scale is removed.

From retained profiles to a construction family. Using the construction bounds $B_n(S)$ of Section 6, let $L_{\mathcal{A}}(n) = \max_{S \in \mathcal{A}} B_n(S)$, over the archive states S that can reach n, be the archive envelope; it equals the running maximum $M(n)$ at the end of the search. Its normalized advantage is $g_{\mathcal{A}}(n) = \big(L_{\mathcal{A}}(n) - E_{\mathrm{BHN}}(n)\big)/2^{n-2}$, the quantity already used as $\delta_N(S)$ in Section 6, so $g_{\mathcal{A}}(n) = 0$ means equality with the benchmark. Six independently verified roots attain every envelope value over $9 \le n \le 16{,}384$ under a fixed tie-breaking order (Table 12); the searched Q_7 branch supplies a uniform benchmark floor, and complementary recursive profiles supply the strict gains. Every envelope value is a construction bound derived from a verified root by legal lifts and coordinate restriction, so no graph has to be materialized at each target dimension; stored explicit certificates are a separate evidence layer and reach Q_{30} (Appendix Q).

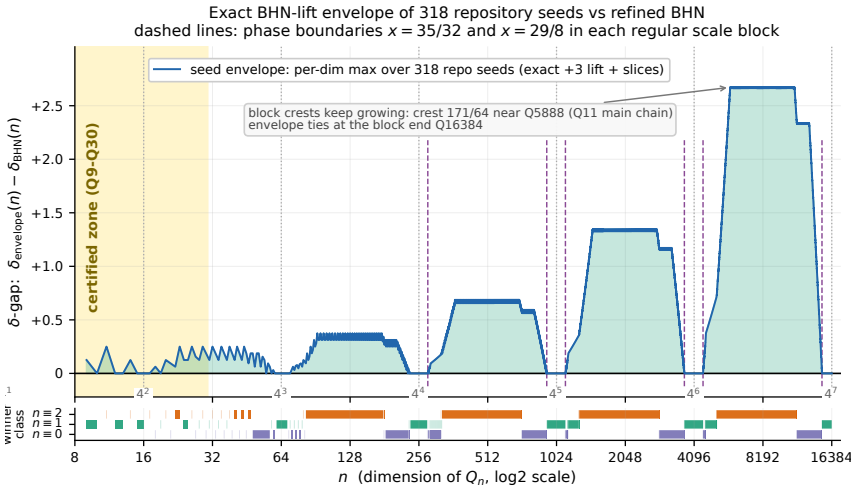


Figure 4: The archive envelope relative to the refined BHN benchmark. The curve is the normalized advantage $g_{\mathcal{A}}(n)$; the zero line is equality with the benchmark. Dashed lines mark the phase boundaries $x = 35/32$ and $x = 29/8$ of the regular scale blocks, and the lower strip records the residue class $n \bmod 3$ of the winning root's dimension. Shading marks the range covered by explicit certificates; larger dimensions use the recurrence and slice bounds from the six verified roots. The normalized block maximum first reaches $171/64$ at $Q_{5,888}$.

Theorem 1 (An all-dimension construction guarantee)**.** *Let $\mathcal{A}$ contain the searched Q_7 seed, whose recursive state is the classical one. Maximum-direction BHN lifts and coordinate restriction of that seed give a C_4-free graph with at least $E_{\mathrm{BHN}}(n)$ edges, and every bound $B_n(S)$ is met or exceeded by some coordinate slice of S; hence*

$$\mathrm{ex}(Q_n, C_4) \;\geq\; L_{\mathcal{A}}(n) \;\geq\; E_{\mathrm{BHN}}(n), \qquad n \geq 7.$$

The proof is in Appendix D. The guarantee holds for all $n \geq 7$, beyond the plotted range, and it does not rely on extrapolating the observed phase pattern.

The strict gains occupy a scale-locked phase window. Partition dimensions into scale blocks $[4^a, 4^{a+1})$, writing $x = n/4^a \in [1, 4)$ for the within-block phase. In every fully audited regular block, strict gains occupy the window $\frac{35}{32} < x < \frac{29}{8}$ and tie the two flanking windows, up to four boundary exceptions per block (Figure 4); the audited recurrence continuation through $Q_{4,194,304}$ observes the same pattern (Appendix Q).

The 84.3% is a structural, not incidental, quantity. The gains make up $f(4^7) = 13{,}805/16{,}376 = 0.843002$ of the evaluated dimensions—not an isolated percentage, but the $R = 7$ block-boundary value of a regular sequence: for the audited boundaries $4 \leq R \leq 11$ (exact rational arithmetic through $R = 7$, validated continuation beyond), the cumulative gain fraction satisfies the closed form

$$f(4^R) = \frac{27}{32} - \frac{2R - \frac{7}{4}}{4^R - 8}. \tag{13}$$

If this block self-similarity persists at all scales, $f(4^R) \to \frac{27}{32}$, while the full cumulative fraction does not converge, oscillating between $\frac{27}{35}$ and $\frac{27}{29}$ within every block (Appendix Q).

Growing gains beyond the exponential edge scale. Let $g_{\mathrm{crest}}(a)$ be the largest normalized advantage $g_{\mathcal{A}}$ in block $[4^a, 4^{a+1})$, first attained at $n_{\mathrm{crest}}(a)$ (lift mechanics in Appendix Q). In the exactly evaluated blocks $2 \leq a \leq 6$,

$$\begin{aligned} n_{\mathrm{crest}}(a) &= 23 \cdot 4^{\,a-2}, \\ g_{\mathrm{crest}}(a) &= \frac{2^a}{24} + \frac{1}{3 \cdot 2^a}. \end{aligned} \tag{14}$$

The block maximum therefore rises from $1/4$ at Q_{23} to $171/64$ at $Q_{5,888}$, a 10.6875-fold increase in a quantity that already divides out $2^{\,n-2}$: the advantage over the benchmark grows beyond the exponential growth of raw edge counts. The maximum is a plateau rather than an isolated point, attained on an arithmetic progression of $7 \cdot 4^{\,a-2} + 1$ dimensions (1,793 of them in the block ending at $Q_{16,384}$). If the observed crest formula persists at all scales, the crest advantage keeps growing,

$$g_{\mathrm{crest}}(a) = \frac{2^a}{24} + \frac{1}{3 \cdot 2^a} \;\longrightarrow\; +\infty, \tag{15}$$

so the normalized gap at the crest is unbounded. This divergence concerns the peak of the gain window only: the benchmark floor of Theorem 1 holds at every dimension unconditionally. The finite-range verification and the conditional asymptotics are separated in Appendix Q. Together, these results make explicit how retention acts in the hybrid system: an immediate edge-count objective can discard a recursive seed whose descendants are stronger (Discovery II), and the envelope shows that different retained states pay off at different target dimensions; keeping complementary profiles, rather than the single best current one, is what turns the archive into a construction family.

8 Discussion and Limitations

The hybrid system connects structural search to recursive value through a compact state, a legal extension rule, and independent verification. The evaluator describes continuation under the BHN template, so it measures the recursive potential of a fixed template rather than an unconstrained optimum.

The implemented protocols include human review. Bounded repair studies show profile steering and a retention–diversity trade-off, and offline ranking uses already-solved candidates. Root verification, recurrence-derived bounds, and explicit high-dimensional artifacts are distinct evidence levels (Appendices I–Q).

Ethics Statement

This work studies an extremal combinatorics problem; it involves no human subjects or personal data, and its outputs are mathematical constructions and search procedures.

Reproducibility Statement

The accompanying repository `https://anonymous.4open.science/r/AI-Assisted-Search-for-C4-Free-Hypercube-Constructions-with-Recursive-Feedback-F305` ships the 318-key seed library, the bundled certificates with SHA-256 manifests, the exact envelope series, and two standalone verifiers; its `README.md` gives the layout and entry points. The search-loop code and run data are not part of this release. The artifacts separate three evidence layers: root-certificate checks, exact recurrence bounds, and archived solver outcomes. The independent checker imports no generation module and verifies cube feasibility, code–kernel consistency, stabilizer inclusion, and quotient–lift correspondence; the numerical bound audit rechecks root graphs and direction profiles separately. Reproduction of the expanded archive requires its state manifest, dimension-wise witnesses, figure values, and source hashes alongside the evaluator. Appendices give metric definitions, solver budgets, termination statuses, and the all-dimension proof; long-range counts require the recurrence and valid roots, not explicit edge lists at every dimension.

AI Use Statement

Generative AI assisted search-experiment design and manuscript editing. All construction claims rest on certificate checks or deterministic recurrences from valid roots, and all empirical summaries come from archived outcomes. The authors reviewed the AI-assisted work and take responsibility for the final manuscript and artifacts.

References

Peter Brass, Heiko Harborth, and Hauke Nienborg. On the maximum number of edges in a C_4-free subgraph of Q_n. *Journal of Graph Theory*, 19(1):17–23, 1995. doi: 10.1002/jgt.3190190104.

Erdős Problems. Problem 86: Four-cycle-free subgraphs of the hypercube, n.d. URL `https://www.erdosproblems.com/86`.

Maxime Gasse, Didier Chételat, Nicola Ferroni, Laurent Charlin, and Andrea Lodi. Exact combinatorial optimization with graph convolutional neural networks. In *Advances in Neural Information Processing Systems*, 2019. URL `https://arxiv.org/abs/1906.01629`.

Robert Lange, Yuki Imajuku, and Edoardo Cetin. ShinkaEvolve: Towards open-ended and sample-efficient program evolution. In *International Conference on Learning Representations*, 2026. URL `https://proceedings.iclr.cc/paper_files/paper/2026/hash/7886b9bafe76c52fd568db10ff9772df-Abstract-Conference.html`.

Minamo Minamoto. C_4-free subgraphs of the hypercubes Q_6, Q_7, and Q_8: Odd squares, fully frustrated models, and computational structure. *arXiv preprint arXiv:2603.29127v5*, 2026. URL `https://arxiv.org/abs/2603.29127v5`.

Bernardino Romera-Paredes, Mohammadamin Barekatain, Alexander Novikov, Matej Balog, M. Pawan Kumar, Emilien Dupont, Francisco J. R. Ruiz, Jordan S. Ellenberg, Pengming Wang, Omar Fawzi, Pushmeet Kohli, and Alhussein Fawzi. Mathematical discoveries from program search with large language models. *Nature*, 625(7995):468–475, 2024. doi: 10.1038/s41586-023-06924-6.

Rafał Wrona. Explicit C_4-free subgraphs of hypercubes Q_9 through Q_{15}. Repository with explicit edge-list certificates and an independent verifier, 2026. No exactness claim is made for $n \geq 7$.

A Exactness of the Quotient Encoding

We record the counting and exactness facts behind Section 2. Throughout, H has full row rank and its columns $g_i = He_i$ are distinct and nonzero.

Orbit counts. Translation by C acts freely on edges: a nontrivial $c \in C$ fixing the edge $\{x, x+e_i\}$ would require $c = e_i$, i.e. $g_i = 0$. Hence every edge orbit has exactly $|C| = 2^{n-r}$ edges, giving $E(\Phi_H(F)) = |C|\,|F|$; and direction i forms the perfect matching $\{\{A, A+g_i\} : A \in \mathbb{F}_2^r\}$ of size 2^{r-1}, so the state carries $n2^{r-1}$ binary decisions. The same freeness holds on squares: a square in directions (i, j) fixed by a nontrivial $c \in C$ would need $c \in \{e_i, e_j, e_i + e_j\}$, forcing $g_i = 0$, $g_j = 0$, or $g_i = g_j$. The constraint of Equation (2) is therefore identical at the four syndromes $A, A+g_i, A+g_j, A+g_i+g_j$ and is imposed once per square orbit, so both decision and constraint counts are reduced by exactly $|C|$.

Proposition 3 (Exactness of the quotient encoding)**.** *A C-invariant lift $\Phi_H(F)$ is C_4-free if and only if F satisfies Equation (2) on every square orbit; the encoding then preserves the edge count, $E(\Phi_H(F)) = |C|\,|F|$.*

Proof. Every C_4 in Q_n is a coordinate square, and by C-invariance squares appear in orbits. The four edges of the square at x with $Hx = A$ in directions $i < j$ are exactly the four variables of Equation (2), so that constraint is violated exactly when all four edges of a square are selected. The edge count is the orbit count above. □

Release safety. Let $C' \le C$ be an index-two subgroup, so each C-orbit splits into two C'-orbits. Selecting both copies of every incumbent orbit re-represents the incumbent edge set in the finer quotient: symmetry release enlarges the represented space without losing any certified state.

Verification scope. The independent checkers in the accompanying repository (`verifiers/verify_all_new_certs.py` and `verifiers/verify_certs_independent.py`, importing no generation-side module) verify, for each certificate: ... ... and provenance (parent basename and SHA-256). The repository also ships the deduplicated 318-state seed library with SHA-256 manifests (`seedlib_318/`), whose `tools/verify_package.py` rechecks the envelope from the archived profiles without the generation pipeline.

B The Instantiated Operator Algebra

Table 3 instantiates each kind of §3 as an executable procedure; the parameter blocks are what the proposer fills in.

C Proof of Proposition 1

Proof of Proposition 1. Every 4-cycle of Q_n is a coordinate square. In the quotient each square is an orbit of four edge classes $\{a, b, p, q\}$ with a, b old-direction and p, q new-direction. Old–old squares stay satisfied because deleting edges cannot create a cycle. Each old–new square is covered by $y_p + y_q \le 1 + r_a + r_b$, which is violated exactly when all four members are selected. Each new–new square is covered by $y_1 + y_2 + y_3 + y_4 \le 3$. Conversely, any $0/1$ point satisfying all these constraints selects a C_4-free F': the feasible region of the local program coincides with the set of C_4-free states reachable by the move, so the program's optimum equals the maximum over the neighbourhood. □

D Why the State-Equivalent Branch Preserves BHN

Here we prove the averaging equality used in Theorem 1. For the classical main chain $m \equiv 1 \pmod 3$, the sorted profile has two possible values. Write $4^a \le m \le 4^{a+1}$ and define

$$\begin{aligned} p_H &= 2^{m-2} + 2^{m-a-2}, \quad N_H = (4^{a+1} - m)/3, \\ p_L &= 2^{m-2} + 2^{m-a-3}, \quad N_L = (4m - 4^{a+1})/3. \end{aligned} \tag{16}$$

Table 3: The instantiated operator algebra $\mathcal{O}$. Every kind is an executable procedure taking a certified quotient state to a certified quotient state; `params` is the machine-readable block consumed by the frozen executor, and its typing is what makes a proposal checkable before it is run (Section 3). `escape_multiround` is a composite that repeats exclusion-based repair and retains the best verified output. The final block is the dimension-lifting family: the lifts add generator columns and re-solve the exact local program of Proposition 1, while `full_refine` refines the representation without changing n.

Kind	Effect	`params`
`escape`	force a different orbit choice	`k_force`
`escape_structured`	force change in one direction	`block_dir`
`escape_multi`	exclude incumbents, re-solve	`n_exclusions`
`escape_structured_excl.`	both of the above	`n_excl.`, `block_dir`
`perturb`	pin a core, free the rest	`protect_frac`
`perturb_exclusion`	protected escape + exclusion	`n_excl.`, `protect_frac`
`release`	refine C to an index-2 subgroup	`h`
`escape_multiround`	k exclusion rounds, keep best	`k_rounds`, `n_excl.`, `best_of`
`lift_single`	add one column, exact re-solve	`g`
`lift_joint`	add two columns, joint re-solve	`g1`, `g2`
`lift_joint3`	add three columns, joint re-solve	`g1`, `g2`, `g3`
`full_refine`	complete refinement at fixed n	—

The profile is $[p_H^{\times N_H}, p_L^{\times N_L}]$. At $m = 4$ it is $[6^{\times 4}]$; after one lift it is the Q_7 profile in Table 4. Within a scale interval, Equation (5) replaces one heavy direction by four light directions, after multiplication by eight. Specifically, an inherited heavy direction becomes the heavy value at $m+3$, an inherited light direction becomes the light value there, and $2^m + 4p_H$ equals that light value. The multiplicities change by $N_H \mapsto N_H - 1$ and $N_L \mapsto N_L + 4$, precisely as in Equation (16).

At a scale boundary $m = 4^{a+1}$, $N_H = 0$ and all directions have the same value. Re-expressing that value as the heavy value for the next interval starts the same induction. Since the searched Q_7 seed has the identical initial profile, Proposition 2 transfers this entire profile evolution to its lineage.

Now fix $n \geq 7$, let $4^a \leq n < 4^{a+1}$, and choose the smallest $m \geq n$ with $m \equiv 1 \pmod 3$. Then $j = m - n \in \{0, 1, 2\}$ and $m \leq 4^{a+1}$. If $j > 0$, the node has at least four light directions, so the j deleted directions can all have load p_L. Summing the profile and applying Equation (8) gives

$$\begin{aligned} \frac{E_m - jp_L}{2^j} &= n2^{n-2} + \frac{2n + 4^{a+1} - j}{3}\, 2^{n-3-a} \\ &= n2^{n-2} + \left\lfloor \frac{2n + 4^{a+1}}{3} \right\rfloor 2^{n-3-a} = E_{\mathrm{BHN}}(n). \end{aligned} \tag{17}$$

For the floor equality, $j \equiv 1-n \pmod 3$ and $4^{a+1} \equiv 1 \pmod 3$, so j is the remainder of $2n+4^{a+1}$ modulo three. The same calculation holds for $j = 0$. A maximum slice has at least this many edges. Neither equal slice counts nor finite computational extrapolation is needed. Removing edges, if necessary, gives a graph with exactly the benchmark count.

E Structural Checks for the Q_7 Discovery

The two root edge lists are checked for valid endpoints, hypercube adjacency, duplicate edges, and absence of a coordinate square. Direction profiles are recomputed from edge differences. Stable WL refinement, initialized uniformly and iterated to stability, produces four classes of size 32 for the searched graph and two classes of sizes 32 and 96 for the classical graph. Different stable class counts certify graph non-isomorphism. Translation stabilizers are obtained by testing whether each translation in $\mathbb{F}_2^7$ preserves the edge set; their orders are four and two.

Membership testing concerns the standard BHN parity-projection $Q_4 \to Q_7$ lift under $\mathrm{Aut}(Q_7)$, not arbitrary variants of the construction. The common profile forces the three inherited directions and the four gadget directions into separate blocks. There are $3!4!2^7 = 18{,}432$ alignments from block

Table 4: Same recursive state, different graph structure in Q_7. The lift-image test concerns the standard BHN $Q_4 \to Q_7$ template under cube automorphisms (Appendix E).

Property	Classical BHN seed	Search-generated seed
Edges E	304	304
Sorted direction profile P	$[48^{\times 3}, 40^{\times 4}]$	$[48^{\times 3}, 40^{\times 4}]$
Stable WL vertex classes	2	4
Translation stabilizer order	2	4
Standard lift-image consistency test	Pass	Fail

permutations and translations.The archived membership audit finds zero witnesses for the searched graph: all $18{,}432$ alignments fail a necessary inherited-edge fiber-consistency condition. The classical positive control passes under all $18{,}432$ tested alignments, so the test is non-vacuous; the first witness reconstructs the independently stored Q_4 parent edge-for-edge. A separate checker tests all 35 choices of four gadget coordinates: none passes the necessary condition for the searched graph, while the classical gadget block does. A failure of this necessary condition excludes membership in the specified image. A VF2 check certifies that the two Q_7 graphs are not abstractly isomorphic.

Branching one lift step later. The Q_7 profile has three maximum directions, so the template is set-valued at the graph level while remaining single-valued at the state level. For the classical Q_7 graph, the children T_0 and T_1 are cube-equivalent while T_2 is not, yet all three share $(E, P) = (3328, [384, 384, 320^{\times 8}])$. For the cross test, all 3×3 pairs $T_e(G_7^{\text{ours}})$ versus $T_d(G_7^{\text{ref}})$ are non-cube-equivalent: each pair is decided by exhaustive search over the $80{,}640$ profile-compatible coordinate permutations and $1{,}024$ translations ($82{,}575{,}360$ candidates, no witness). The conclusion is scoped to the tested branches; completeness over all cube-equivalent parent copies is unproved.

Both root certificates ship in the supplementary package (`seedlib_318/certs/`), and the invariants in Table 4 are recomputed directly from the stored edge lists. The archived checks also construct and verify a Q_{10} descendant with $3{,}328$ edges. These structural tests establish a non-classical realization with a known recursive state, distinct from the archived classical realization.

F The BHN State Family

Definition 4 (BHN state family)**.** *With $(E_n^{\text{ref}}, P_n^{\text{ref}})$ the classical reference state at dimension $n = 4 + 3t$, $\mathcal{S}_n^{\text{BHN}} = \{G \subseteq Q_n : G$ is C_4-free, $(E(G), P(G)) = (E_n^{\text{ref}}, P_n^{\text{ref}})\}$. The* construction family *consists of graphs reached by the specified iterated BHN template from Q_4.*

State-family membership is preserved under the maximum-direction lift: the classical profile evolution and the induction are in Appendix D, and the searched Q_7 seed inherits the same evolution by Proposition 2.

G Cross-Lineage State Capture at Q_{13}

Lineage and certificate. The lineage originates outside the classical BHN lift chain: the searched Q_7 quotient seed is extended by a joint $Q_7 \to Q_9$ step (not the $+3$ template) to a Q_9 root with $E = 1{,}472$ and profile $[192^{\times 3}, 160^{\times 4}, 128^{\times 2}]$ (the classical Q_9 value is $1{,}504$, and Q_9 is not a chain dimension). From there the chain uses only template operations: a strict BHN lift to a Q_{12} node with $E = 15{,}360$ (classical: $15{,}616$; `q12_altchainC_15360.json`, operation `BHN_extend`), a second strict BHN lift to a Q_{15} node with $E = 151{,}552$ and profile $[12288, 10240^{\times 12}, 8192^{\times 2}]$ (classical: $153{,}600$; `q15_altchainC_151552.json`, operation `BHN_extend`), and finally a slice deleting the two light directions (load $8{,}192$ each). The four slices have equal edge counts $33{,}792$, computed directly from the stored Q_{15} parent certificate; one component is materialized and certified as H_{13} (`q13_altchainCsl_33792.json`). The certificate is independently certified: endpoints, adjacency, all $159{,}744$ coordinate squares, C_4-freeness, and parent-hash consistency all pass.

State equality and graph separation. Table 5 lists the invariants. The recursive state coincides with the classical reference exactly, so $H_{13} \in \mathcal{S}_{13}^{\mathrm{BHN}}$; the degree sequence does not separate the graphs, while the trace of A^6 and the union WL histogram do. The conclusion is that an external construction lineage reaches the BHN state family through a graph structurally distinct from the archived classical reference realization. We do not claim separation from the entire branch family $\mathcal{C}_{13}^{\mathrm{BHN}}$, whose full orbit structure is unenumerated; sibling certificates at Q_{16} and Q_{19} exclude cube automorphisms only. All certificates named above are bundled in the supplementary package and anchored in `seedlib_318/cert_manifest.csv`.

Table 5: State capture at Q_{13}: identical recursive state, different graph. The union WL histogram refines the disjoint union of the two graphs with shared color labels.

Property	H_{13} (external lineage)	G_{13}^{ref} (classical)
Edges E	33,792	33,792
Sorted direction profile P	$[3072, 2560^{\times 12}]$	$[3072, 2560^{\times 12}]$
Degree sequence	{8:6144, 9:2048}	{8:6144, 9:2048}
$\mathrm{tr}(A^k)$, $k \le 5$	$0, 67584, 0, 1050624, 0$	same
$\mathrm{tr}(A^6)$	23,531,520	23,396,352
Union 1-WL color classes	8	28

H Exact Counts Behind the Ranking Reversal

Table 6 separates the classical seed rolled forward by the same maximum-direction template from the refined dimension-wise BHN benchmark. These are distinct references at later dimensions. All counts are integer recurrence evaluations from the root profiles; no high-dimensional optimization is needed. Extending the same recurrence from the two root profiles (in the accompanying repository: `seedlib_318/seeds_318.jsonl` and `seedlib_318/bhn_reference.json`) over every reachable dimension $13 \le n \le 16{,}384$, the searched lineage leads the classical rollout at 16,358 of the 16,372 dimensions; the classical seed falls below the refined benchmark from Q_{18} onward, and the searched lineage's non-leading dimensions all sit in the troughs at powers of four, where its margin against the benchmark changes sign. In units of 2^{n-1}, the classical rollout's margin is $-35/32$ at Q_{1000} and the searched lineage's is $-1709/512$ at Q_{16384}.

Table 6: Exact values underlying the reversal of §5 (Figure 5, first lift steps). G_{s} is the searched Q_{15} seed with profile $[12800, 9984^{\times 14}]$ (`seedlib_318/seeds_318.jsonl`); G_{c} is the classical Q_{15} seed (`seedlib_318/bhn_reference.json`).

h	n	$J_h(G_{\mathrm{s}})$	$J_h(G_{\mathrm{c}})$	$E_{\mathrm{BHN}}(n)$	Δ vs. G_{c}	Δ vs. BHN
0	15	152,576	153,600	153,600	$-1,024$	$-1,024$
1	18	1,454,080	1,441,792	1,449,984	$+12,288$	$+4,096$
2	21	13,352,960	13,238,272	13,303,808	$+114,688$	$+49,152$
3	24	120,586,240	119,537,664	120,061,952	$+1,048,576$	$+524,288$

The two plotted normalized margins, in units of 2^{n-1}, are $(-1/16, 3/32, 7/64, 1/8)$ against the classical seed rollout and $(-1/16, 1/32, 3/64, 1/16)$ against the BHN benchmark. The result is a ranking reversal at the displayed depths, rather than a claim that a single stronger lineage dominates at all dimensions or horizons.

I Feedback Protocols and Outcome Definitions

Mode-level certificates. The certificates of Table 2 are stored and independently re-verified. The envelope of §7 maximises over a larger root set and dominates these single-chain values where they differ; the table's point is that the mode's own objective already yields benchmark-exceeding certificates. Establishing the mode's central selection claim — that δ-guided selection outperforms edge-guided selection when the pool is too large to enumerate — requires a pool in which exhaustive

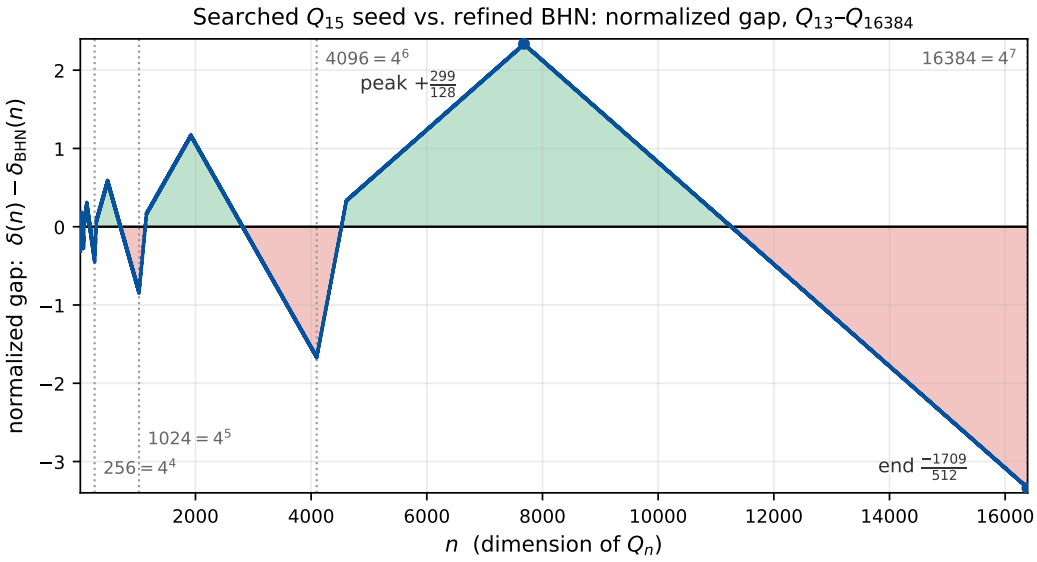


Figure 5: The extended comparison of the two rollouts of §5, on the same recurrence as Table 6, out to Q_{16384}. The searched lineage leads at 16,358 of 16,372 dimensions; the classical seed's rollout falls below the refined benchmark from Q_{18} onward. This is a comparison of two recurrence rollouts from fixed roots, not a claim about the maximum.

evaluation is infeasible, and the pre-registered comparison is a held-out pairwise ranking and top-k regret against the $|F|$ baseline, split by root family rather than by row.

Overview and main outcomes. We attach $V(G) = (J_1, J_2, J_3)$ to each certified state and compare candidates at common dimensions and lift depths. Figure 6 summarises the two main outcomes.

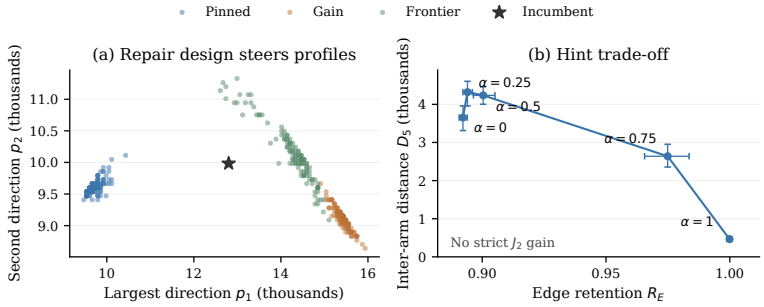


Figure 6: Repair designs steer profiles; hints trade retention for differentiation. **(a)** Three arms on the same archived deep-release states produce distinct profiles; the star marks the incumbent. **(b)** Increasing the hinted fraction α changes edge retention in the pinned arm and mean inter-arm distance in the five largest direction loads. Intervals are 95% child-bootstrap intervals, conditional on feasible outputs. Neither protocol produces a strict incumbent J_2 gain.

In the deep-release outputs, complete profile signatures do not overlap between arms (Figure 6a). All 552 deep and shallow outputs are FEASIBLE; they establish construction values, not model optima. Deep-release differentiation costs edges, whereas shallow repairs retain more of the incumbent. Hints suggest incumbent orbit values without fixing variables, so they change search behavior while preserving the feasible set. At $\alpha = 0.75$ the pinned arm retains approximately 97.5% of incumbent edges; full hints raise retention and reduce mean inter-arm profile distance (Figure 6b). Metrics exclude unknown outcomes and are conditional on feasibility. Neither protocol adds a new future-value Pareto point relative to the reference archive, and no hint setting strictly improves incumbent

J_2. These observations establish that hint strength and no-good exclusion steer structural exploration while the incumbent value is preserved.

Repair designs and decision timing. The process experiments use three fixed repair arms inspired by recursive evaluation. Let d_1, d_2 be the two largest direction loads of the parent state. The profile-pinned arm maximizes edge count while preserving the parent load in d_1. The gain-oriented arm uses direction weights two on d_1, four on d_2, and one elsewhere, with an edge floor. The frontier-oriented arm uses weights two, three, and one, respectively, and requires the output load in d_1 to differ from the parent load. Thus the arms differ in objective and associated repair constraints; they are not an ablation of objective coefficients alone. Their names describe intended search directions, not guarantees of future-value improvement.

Deep-release runs use an edge floor of four quotient orbits below the parent in the gain-oriented arm. Shallow-release and hint-sweep runs use 96% of the parent's orbit count as that floor. Hints give selected parent orbit variables value one without fixing them. For the sweep, the hinted fraction $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$ is chosen by the protocol; h remains reserved for recursive lift depth. Subsets are ordered using the archived seed-dependent hash rule. Solver random seeds vary by shard and task. These protocols do not invoke the offline selector before choosing an arm.

Validity and value improvement. Only FEASIBLE or OPTIMAL outputs provide candidates. The records include C_4 checks and direction profiles; the manuscript audit recomputes J_1, J_2, J_3 from each valid profile. A structural signature is either the complete sorted profile, when testing inter-arm overlap, or its three largest entries, when reporting the diversity counts in Table 7. Structural novelty is separate from a new future-value Pareto point. A vector dominates another when it is no smaller at all three horizons and strictly larger at one. The reference archive used by these protocols contains $(1{,}454{,}080, 13{,}352{,}960, 120{,}586{,}240)$ and two additional archived vectors. No output in Table 7 adds a new point relative to that reference archive.

Table 7: Bounded repair outcomes. All 552 recorded outputs are FEASIBLE, hence establish construction values rather than model optima. The deep and shallow protocols use different states and budgets; their rows are not an equal-budget representation ablation.

Protocol	Arm	Outputs	Edge-count range	Top-three signatures	Maximum J_2
Deep	Pinned	120	134,656–143,040	70	12,529,664
Deep	Gain	120	131,648–135,616	90	12,492,800
Deep	Frontier	120	132,288–140,736	117	12,648,448
Shallow	Pinned	64	152,448–152,576	6	13,352,960
Shallow	Gain	64	152,064–152,576	5	13,352,960
Shallow	Frontier	64	152,064–152,576	7	13,352,960

Budgets and replication. The deep protocol comprises six shards, each with 20 archived child states and three arms, for 360 solves. Each has a 90-second limit and four CP-SAT workers, a nominal total of nine solver-hours. The shallow protocol records 64 children and three arms, for 192 solves, at 120 seconds and four workers (6.4 nominal solver-hours). All outputs in both snapshots are FEASIBLE. These allocations are wall-time limits per solve, not speedup measurements or total infrastructure cost. The shard seeds and state-wise observations describe the sampled variability of the arms compared.

Hint strength and metric definitions. The complete archived sweep has 64 children, five hint fractions, and three arms: 960 attempts at 90 seconds with four workers (24 nominal solver-hours). It returns 851 FEASIBLE outputs and 109 UNKNOWN outcomes. Unknown outcomes have no edge count or profile and are excluded from candidate metrics. Retention is $R_E = E/152{,}576$. For each child and fraction, the differentiation statistic uses feasible arm pairs:

$$D_5(P^{(a)}, P^{(b)}) = \sum_{i=1}^{5} |p_i^{(a)} - p_i^{(b)}|. \tag{18}$$

Table 8 averages this distance over available pairs. Figure 6(b) uses retention from the profile-pinned arm on its horizontal axis, not retention averaged across all arms. Its intervals resample child states

2,000 times, preserving within-child pair dependence, using bootstrap seed 20260918. Unequal feasible counts, especially in the gain arm, make these conditional-on-success comparisons; the graph is not a universal optimum for hint strength.

Table 8: Complete hint-sweep snapshot. Parenthesized counts are feasible outputs out of 64 per arm. Each mean D_5 is over the reported number of available arm pairs. No fraction produces a strict incumbent J_2 gain.

α	R_E, pinned (n)	R_E, gain (n)	R_E, frontier (n)	Mean D_5	Pairs
0.00	0.892 (64)	0.959 (11)	0.889 (64)	3,652	86
0.25	0.894 (64)	0.959 (20)	0.890 (64)	4,320	104
0.50	0.900 (64)	0.983 (52)	0.895 (64)	4,235	168
0.75	0.975 (64)	0.997 (64)	0.936 (64)	2,639	192
1.00	1.000 (64)	1.000 (64)	0.999 (64)	465	192

No-good repairs under full hints. A related 256-solve protocol uses 64 shallow-release children, full hints, and 120-second limits with four workers. The pinned control has six distinct top-three profiles and mean $E = 152{,}560$. Adding a no-good exclusion to that same arm gives six profiles and mean $E = 152{,}474$. Gain and frontier arms with no-good exclusions give eight and six profiles, with mean counts 152,386 and 152,392. Across the three no-good arms, the union contains 11 profiles, including five not seen in the pinned control. That union is a comparison of three repair arms with one control arm; it does not isolate the exclusion's diversity effect at fixed objective. The results show that full hints are compatible with alternative profiles, but do not prove that hints inherently suppress freedom or that a no-good alone increases diversity. All four arms attain at most the incumbent J_2.

J Offline Candidate Selection

Table 9 reports the outcome.

The archived selector study uses 1,176 solved-candidate rows grouped into 283 child states from three representation families. It reserves ten children per family for an evaluation set of 30 children, and trains on the remaining children. Candidate features include family and generating-arm indicators, E, p_1, p_2, p_3, and profile spread. A binary head predicts membership on the child's (J_1, J_2, J_3) Pareto front; a regression head predicts J_2. The implementation selects the row maximizing $\widehat{p}_{\text{front}} + 10^{-8}\widehat{J}_2$. This is ranking of already solved candidates, not choice of an unexecuted search action from parent-only features.

For child c with candidate pool $\mathcal{P}_c$, the reported selection statistic is

$$\text{Imp}_{J_2}(c) = J_2(G_{\text{selected}}) - \frac{1}{|\mathcal{P}_c|}\sum_{G\in\mathcal{P}_c} J_2(G). \tag{19}$$

The report averages this raw count across the evaluation children. It is a within-pool selection margin, not improvement over the incumbent or BHN. Pools include both Q_{14} and Q_{15} seeds, so their unnormalized magnitudes should not be interpreted as a dimension-independent performance gain.

The gain reference maximizes $(J_2 - J_{2,\text{ref}})/(|E - E_{\text{ref}}| + 1)$ within the solved pool; the frontier reference selects using its true Pareto label. Both are scored on complete candidate profiles from the archived solved pool.

K Controller Calibration Measurements

Proposal executability. In the controlled probe of §3, 186 proposals from the frozen model were 98.4% well-formed and 55.9% mechanically correct in prose, yet only 31.2% executable; every loss was decidable before execution, from free-text parameters or out-of-algebra kinds.

Table 9: Archived offline selection results; no new solves. The frontier baseline and oracle use candidate ground truth. The learned selector ties both in this snapshot, rather than exceeding the best fixed reference.

Candidate-selection rule	Mean Imp_{J_2}	On child Pareto front
Learned two-head score	$+539,989$	30/30
Gain reference (true-value gain/cost ratio)	$+413,013$	27/30
Frontier reference (true Pareto label)	$+539,989$	30/30
Random candidate	$-357,035$	12/30
Oracle (true maximum J_2)	$+539,989$	30/30

Fitted heads. Estimated on the archive and evaluated leave-one-state-out, the cost head reaches $R^2 = 0.31$ from search-time features while the value head reaches $R^2 = 0.11$: which mechanism will gain is largely not predictable before paying for it. Allocating solve budget by predicted cost against predicted value does not beat uniform allocation at equal budget: -7.3% at a tight budget and at most $+18.8\%$ over 30 seeds, the latter within two standard deviations of the uniform baseline, while an oracle with true costs is consistently better. The headroom is real, but a cost model of this accuracy does not capture it; with whole-batch execution the allocation question does not arise.

Saturation-pool measurements. The low-dimensional loop of §3 was run on a pool of 37 low-dimensional states chosen to be *below* their dimension's record (the records themselves have no headroom and are useless for this test). The loop pushes 5 states to their dimension's record, improves 4 more, and produces a seed bank of 774 distinct constructions. The action statistics make the saturation criterion concrete: across the pool, same-dimension actions attain a best gain of 0 and a negative mean, whereas lifting actions attain $+11{,}968$ to $+239{,}872$ depending on the parent. The representation saturates almost immediately at these dimensions, and nearly all progress comes from jumps — the situation the objective switch of §6 is designed for.

Jump-horizon study. The "at most three dimensions per jump" rule is an empirical finding, not conservatism. On 24 parents, a one-step value objective beat a max-edge objective $16/2/6$ (win/tie/loss), while an exact two-step objective lost $4/2/18$: a longer horizon does not pay here, and the candidate successors are lift outcomes already computed during the loop, so the jump itself costs no additional solve.

Candidate ranking within one pool. Fixing a parent state and varying only the release functional ℓ gives candidates with identical starting edge counts, so ranking by current edges is exactly random. Over 255 candidates with a unique optimum, the best of eleven cheap static features places the true winner at rank 61 (permutation $p \approx 0.23$); the archive's default sparsity-first ordering places it at rank 246. In a second parent the winner is at rank 65 ($p \approx 0.25$). Ranking by current edges and ranking by downstream value select disjoint candidate sets (Spearman -0.775).

Offline policy ablation. The reward-based policy of §3 is trained on 7,790 archived solves over 79 states and reduces top-1 regret from 485 (random) to 246. A model with only action features reproduces the full result bit-for-bit, while a model with only state features degenerates to the constant baseline. Across the escape family the strict-gain *rate* is nearly constant (18.0%–19.6%) while the mean gain ranges from $+168$ to -178 edges: the family separates on magnitude, not on success probability.

L The Value Head: Composition and Online Fit

What it is. Figure 8 fixes the composition of $\hat{V}$ — its inputs, its target and its update rule. Three choices carry the method, and each is a decision that a naive regressor would have got wrong.

How well it fits, and when its class changes. Figure 9 tracks one run of the high-dimensional mode against the number of training rows accumulated so far.

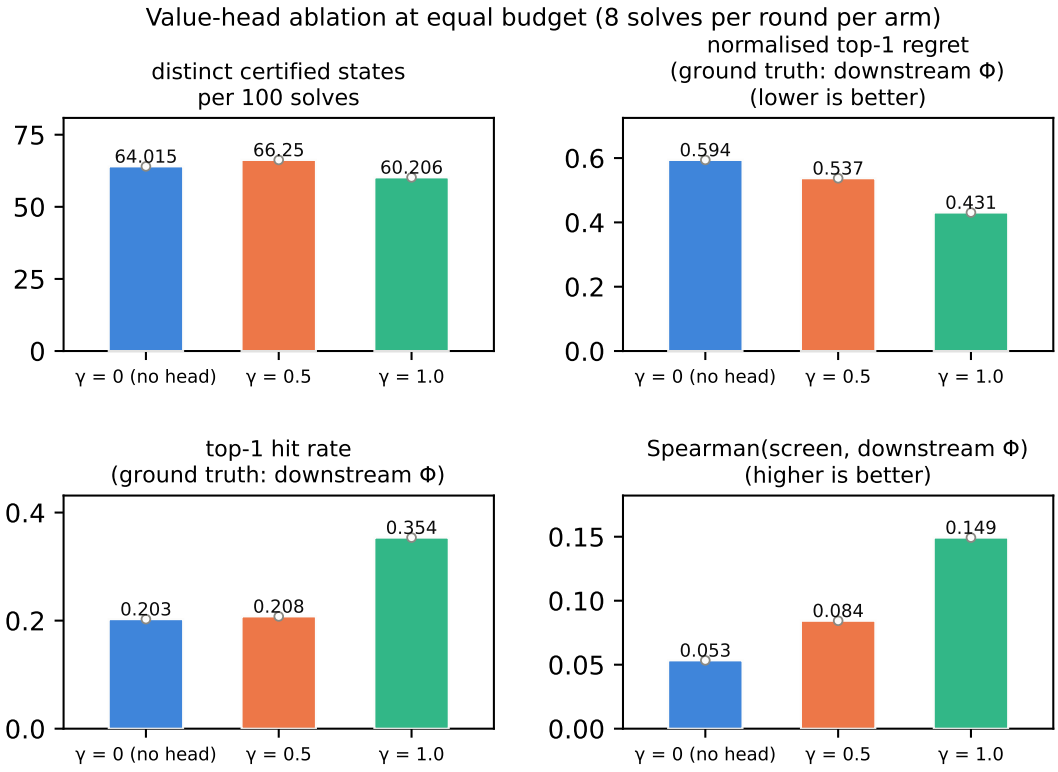


Figure 7: The value-head ablation, $\gamma = 0$ against a sweep, at equal budget. No arm produced a state above its dimension's record: the comparison measures the screen, not the envelope.

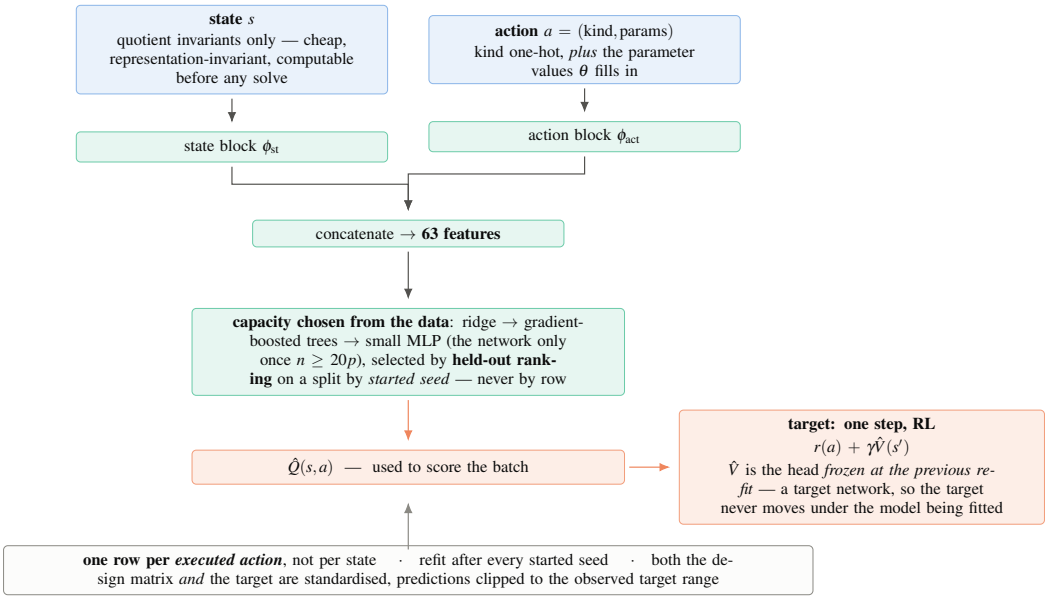


Figure 8: The composition of the value head. Three choices carry the method: the action block carries the *parameter values* the proposer filled in, not only the kind (the prompt requires every kind each round, so the parameters are where its freedom lies, and a kind-only input would hide its contribution); the target is a one-step Q-learning target $r(a) + \gamma \hat{V}(s')$ with $\hat{V}$ frozen at the previous refit, so the target does not move under the model being fitted; and there is one row per *executed action*, not per state.

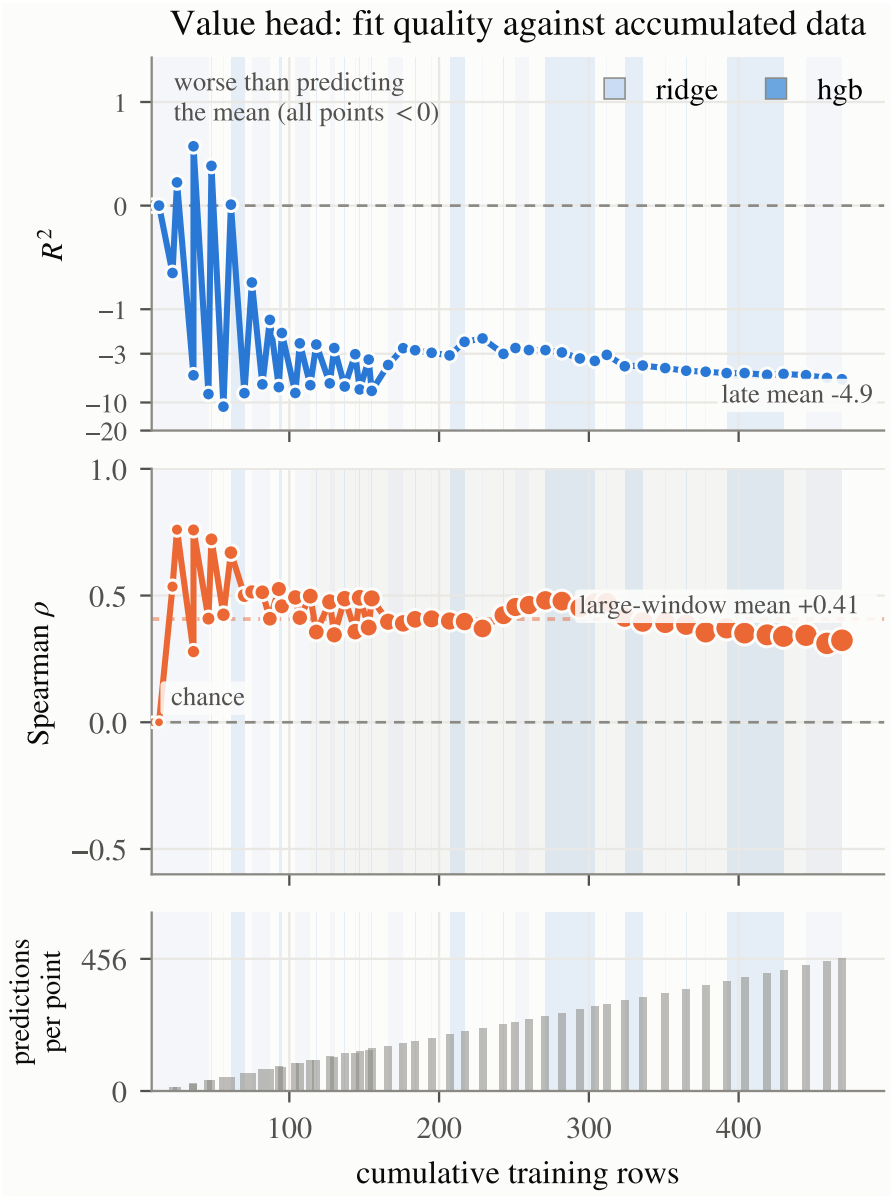


Figure 9: The value head against accumulated data, over one run of the high-dimensional mode (54 refits, 27 class changes). Top: R^2 is negative at every refit and settles near -4.9 — numerically the head never beats the mean, and on a target whose labels span orders of magnitude R^2 is the wrong instrument. Middle: online forward Spearman on unseen states (chance dashed, marker size set by the predictions behind the point) holds at $+0.41$ once the window exceeds a hundred predictions — which is why the loop ranks rather than regresses. Bottom: predictions per point. Shading marks the model class in force, chosen by held-out ranking on a split by started seed.

M Proposer and Action-Library Ablations

The loop's own output against the no-feedback envelope. The certificates of Table 2 dominate the earlier chain, but they do not raise the envelope of this line of work. Lifting the loop's own outputs with the *same* recurrence as the archived baseline gives two envelopes that coincide at every target: over 5,089 loop outputs against 79 archived states, 0 strictly above, 11 tied, 0 below.

Figure 10 draws that comparison.

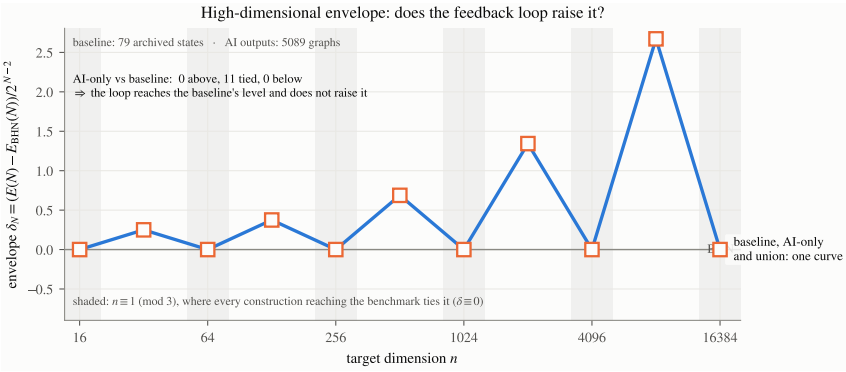


Figure 10: The loop's output against the no-feedback baseline, on one axis, both lifted by the same recurrence so that the only difference is the seed sets. Circle and square markers are the two envelopes drawn on top of one another. On this run the loop *reaches* the baseline's level and does not raise it. Shaded strips mark $n \equiv 1 \pmod 3$, where every construction reaching the benchmark ties it.

Both arms below run at exactly equal budget: the same number of solves per round, the same starting seeds, and a separate training store per arm so that no arm inherits another's data. No arm in either ablation produced a state above its dimension's record — including the whole-library arm — so the comparisons measure the screen, not a change in the envelope.

Proposer: θ against a random draw. Figure 11 gives the two arms at equal budget. θ produces a finer batch (66.3 distinct states per hundred solves against 57.4; median within-batch spread 1890 edges against 781) and exercises more kinds (12 against 9). Its top-1 regret against the *immediate* gain is half the random arm's (0.064 against 0.131); against the downstream Φ the ordering reverses (0.537 against 0.413, top-1 hit rate 0.208 against 0.317, Spearman 0.084 against 0.221). Choosing actions thus helps on the quantity the proposer was asked about and not on the one the loop ultimately wants — the disagreement of §5 reappearing inside the controller. A third arm, an enumerating fixed grid, is not plotted: its median within-batch spread is one edge, so it has almost nothing to choose between and its zero regret measures the batch rather than the screen.

Action library: whole library against one family. Figure 12 varies *what* can be proposed, holding the proposer fixed so that the only thing changing is the operator set. No single family substitutes for the mix: the lift-only arm produces the most distinct states per solve (43.7 against 41.9 for the whole library), the same-dimension arm the fewest (35.7), and the single-family arm reaches the highest ratio of all (54.7) only because it exhausts its four kinds and terminates early — on 172 executions against 1621. The same-dimension arm's batch is degenerate (median gain spread of one edge), so its zero regret measures the batch, not the screen. Least comfortably, the lift-only screen anti-correlates with downstream value (**Spearman** -0.209): a loop restricted to lifting raises the dimension while its screen points away from what the dimension is worth.

N Further Analysis of the Recursive Value

The recurrence also gives an explicit two-step objective. At $n = 15$, let $K_1 = E + p_1$. After the first lift, the largest direction has load $\max(8p_2, 2^{15} + 4p_1)$, so substituting into the second lift gives

$$J_2 = 64K_1 + 8\max(8p_2, 2^{15} + 4p_1) + 2^{21}. \tag{20}$$

This identity follows algebraically from Equation (5), independently of solver status. Below $p_2^* = 4{,}096 + p_1/2$, J_2 is independent of p_2 when E and p_1 are held fixed. Above that line,

$$J_2 = 64K_2 + 2^{21}, \qquad K_2 = E + p_1 + p_2. \tag{21}$$

For the searched seed, $p_2^* = 10{,}496$ and $J_2 = 13{,}352{,}960$. In the activated regime, strictly exceeding that value requires $K_2 > 175{,}872$. The seed's own $K_2 = 175{,}360$ is below this threshold because

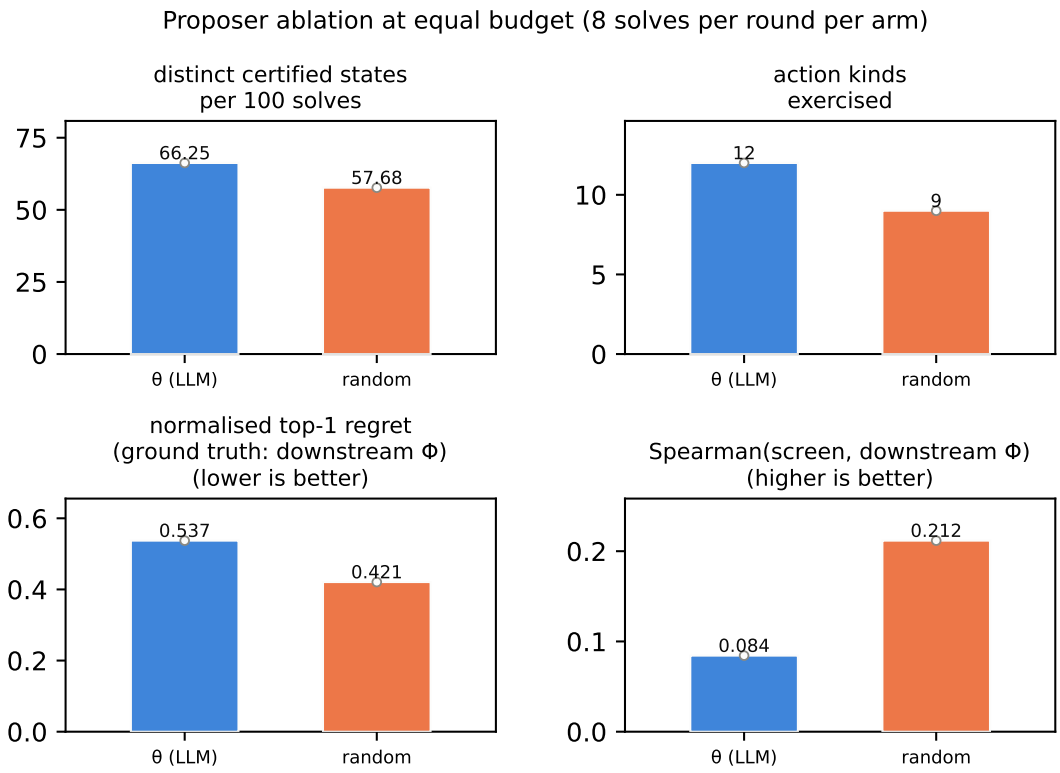


Figure 11: The proposer ablation, two arms at exactly equal budget. Panels: distinct certified states produced per hundred solves; action kinds exercised; the median spread of realised gain *within one batch*; and the normalised top-1 regret of the screen against the downstream Φ.

its second direction remains in the other regime. Raising p_2 without preserving E need not improve recursive value. This distinction explains why profile differentiation in Figure 6 is not itself a new value frontier.

The archived multi-step analysis evaluates 1,110 profiles and reports a Pearson correlation of 0.986 between depth-two and depth-five counts. High correlation in this pool does not rule out ranking reversals elsewhere, establish that longer horizons add no signal, or prove sufficiency of three horizons. Similarly, a reported infeasibility check for improving K_1 applies to the specified `rel255` codimension-seven model, not all Q_{15} graphs or all quotient representations. Local blocker costs and representation scans are diagnostic evidence within their tested neighborhoods. They are not needed for the construction guarantee or the ranking-reversal claim in the main text.

O Construction Data and Evidence Levels

Table 10 gives selected dimension-wise counts.

Root artifacts. The root edge lists are checked independently of the generation code for endpoints, adjacency, duplicate edges, and all coordinate squares by the verifier in the accompanying repository (`verifiers/verify_all_new_certs.py`), and the complete direction profiles are recomputed from the stored edge lists. The inputs include the searched and classical Q_7 certificates, the separate Q_9 root, the Q_{11} root, and the searched Q_{15} root; certificates with $n > 15$ are hash-anchored in `seedlib_318/cert_manifest.csv`. The non-isomorphism and lift-image findings use the structural audit in Appendix E.

Finite and long-range bounds. The exact lift profiles and slicing averages behind the original three-lineage family $L(n)$ are recomputed from the archived profiles by `seedlib_318/tools/`

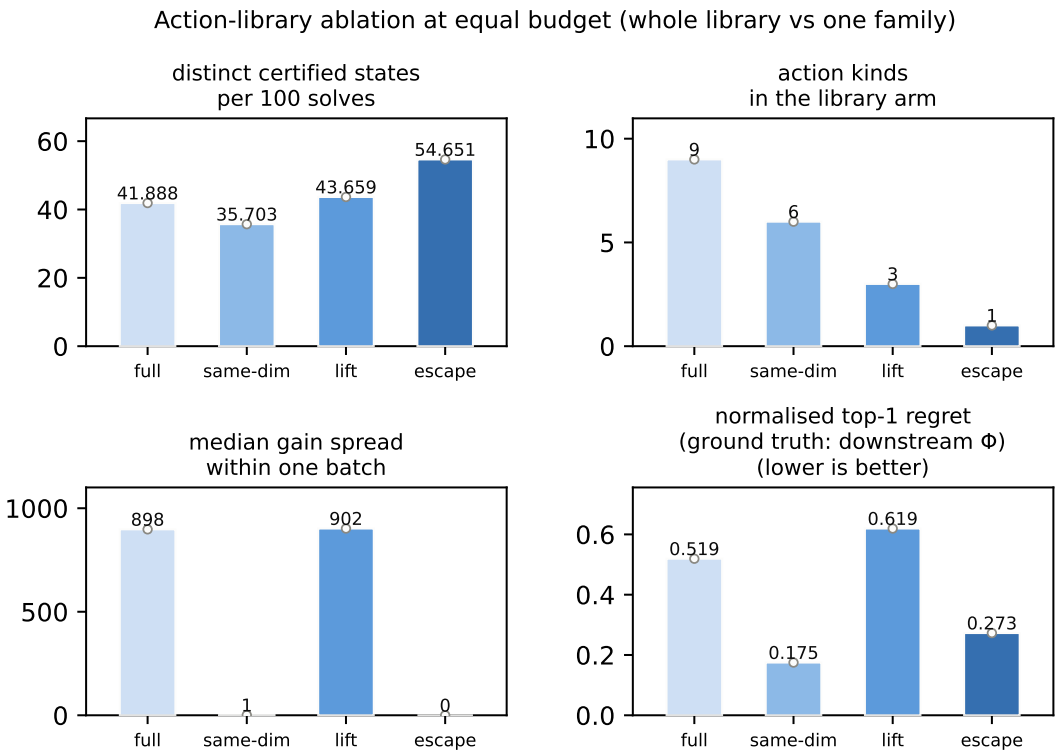


Figure 12: The action-library ablation at equal budget. Four arms, ordered by how much of the operator algebra they retain: the whole library, same-dimension repairs only, recursive lifting only, and a single repair family. The ordinal ramp encodes that order — these are nested degrees of one quantity, not four independent categories.

`verify_package.py`, using exact rational arithmetic. In that range, these averages are integers, so upward rounding is immaterial. Excluding the Q_7 branch gives 15 strict gains, three ties, and four deficits over Q_9–Q_{30}; including it gives 15 gains, seven ties, and no deficits. Over Q_9–Q_{1000}, the combined lower bounds give 720 gains, 272 ties, and no deficits. These finite arithmetic checks concern $L(n)$ and agree with, but do not substitute for, Theorem 1. The larger archive envelope $L_{\mathcal{A}}(n)$ in Figure 4 is audited separately in Appendix Q.

Table 10: Selected dimension-wise counts illustrating lineage complementarity. These are construction lower bounds from valid roots, exact lifting, and slicing averages.

n	Refined BHN	Combined $L(n)$	Margin	Branch supplying the shown gain
9	1,504	1,520	$+16$	Separate Q_9 root
11	7,168	7,296	$+128$	Q_{11}
15	153,600	153,600	0	Q_7 benchmark branch
18	1,449,984	1,454,080	$+4,096$	Q_{15}
20	6,356,992	6,389,760	$+32,768$	Q_{11}
21	13,303,808	13,352,960	$+49,152$	Q_{15}
30	9,428,795,392	9,479,127,040	$+50,331,648$	Q_{15}

Levels of verification. Root-certificate validity, theorem-derived bounds, and archived high-dimensional verification records carry different levels of evidence. The Q_{30} count follows from the valid Q_{15} root and the legal lift recurrence; an archived server record agrees with it. The recurrence-derived lower bound does not depend on which explicit binary artifact is materialized. The original comparison through Q_{1000} and the enlarged archive plot through Q_{16384} use recurrence-and-slicing bounds; they do not report explicit graph generation or square-by-square checking at those dimensions.

Figure regeneration. The figures are regenerated from archived records without running a solver or training a model. The enlarged archive figure uses the audited pipeline of Appendix Q. JSON figure records retain exact counts, sampled-protocol summaries, and input hashes for the feedback and family data.

P Provenance of the Comparison Table

Every column of Table 2, and what it is not.

Refined BHN. The closed form $N2^{N-2} + \lfloor (2N + 4^{a+1})/3 \rfloor 2^{N-3-a}$ for $4^a \le N < 4^{a+1}$. This is a benchmark, not a certificate.

Table 11: The full comparison table: every certified lower bound, by source, for $6 \le n \le 22$. See Appendix P for what each column is and is not.

n	refined BHN	external	quotient chain	archived envelope	this work	vs. BHN	vs. chain	vs. env.
6	132	132	—	—	**132**	0	—	—
7	304	304	—	304	**304**	0	—	0
8	672	680	—	680	**680**	+8	—	0
9	1,504	1,505	—	1,520	**1,520**	+16	—	0
10	3,328	3,304	—	3,329	**3,328**	0	—	−1
11	7,168	7,164	—	7,296	**7,296**	+128	—	0
12	15,616	15,372	—	15,617	**15,616**	0	—	−1
13	33,792	32,856	—	33,793	**33,664**	−128	—	−129
14	71,680	69,909	—	72,193	**72,192**	+512	—	−1
15	153,600	148,126	152,576	153,600	**153,600**	0	+1024	0
16	327,680	—	321,792	327,681	**327,680**	0	+5888	−1
17	688,128	—	676,864	688,129	**688,128**	0	+11264	−1
18	1,449,984	—	1,418,240	1,454,081	**1,454,080**	+4096	+35840	−1
19	3,047,424	—	2,967,552	3,047,425	**3,047,424**	0	+79872	−1
20	6,356,992	—	6,193,152	6,389,761	**6,389,760**	+32768	+196608	−1
21	13,303,808	—	—	13,352,961	**13,352,960**	+49152	—	−1
22	27,787,264	—	—	27,852,801	**27,852,800**	+65536	—	−1

External. The best explicit certificate published by others at that dimension. **One column carries two works because they cover disjoint ranges**: Minamoto (2026) reports Q_6–Q_8 (with Q_6 reproduced from Harborth–Nienborg) and Wrona (2026) reports Q_9–Q_{15}.

Quotient chain. Not an external method and not a published baseline. It is the quotient chain of the *non-AI* pipeline used earlier within this line of work, included as the reference point a within-paper ablation would provide. It contains no language model, no fitted head and no recursive objective. It is reported only for Q_{15}–Q_{20}, which is why the column is empty elsewhere.

Archived envelope. The guaranteed construction bound of the archived seed pool under the same lift and slice rules. It is computed over the same pool that Appendix Q describes, and it coincides with this work's certificate set to within a single edge at every dimension, because a bound and the construction attaining it differ only by the integrality of the recurrence.

This work , the best archived certificate: the maximum over the retained pool — not the same quantity as the "this work" column of Table 2 in the body, if the two differ.

No entry asserts $\mathrm{ex}(Q_n, C_4)$ equals the listed value; every non-BHN, non-envelope entry is a construction lower bound.

Q The 318-State Archive Envelope

Pool and comparison scope. The expanded archive contains 318 distinct recursive states (n, E, P). A state is a numerical fingerprint, not an isomorphism class or an independent search

trial. The envelope is computed over this archived pool under the specified BHN lift and coordinate-restriction rules, not over all C_4-free hypercube graphs. The searched Q_7 branch supplies the refined-BHN floor; complementary profiles supply strict improvements.

Independent exact reconstruction. The audit script `seedlib_318/tools/verify_package.py` (supplementary package) reads the source profiles and maintains an exact multiset of normalized direction loads $q_i = p_i/2^{n-2}$, using rational arithmetic. A lift removes a largest load q and adds four copies of $(1+q)/2$; the normalized excess $\delta = E/2^{n-2} - n$ increases by $q-1$. Deleting one or two light directions gives $\delta + 1 - q_{\min}$ or $\delta + 2 - q_{\min,1} - q_{\min,2}$. The audit maximizes these values separately at every dimension from 9 through 16,384, using lift nodes through 16,386 so that the last two target dimensions receive all slice candidates. The resulting average edge counts in the evaluated range are integers.

Figure 4 displays the expanded archive through Q_{16384}, where the final dimension ties refined BHN. Its ordinate is $g_{\mathcal{A}}(n) = (L_{\mathcal{A}}(n) - E_{\mathrm{BHN}}(n))/2^{n-2}$. The six root certificates and profile arithmetic support the envelope; the plotted explicit-artifact range is not a statement that every descendant in that range was independently rechecked in this audit. Figure 4 displays the expanded archive through Q_{16384}, where the final dimension ties refined BHN. Its ordinate is $g_{\mathcal{A}}(n) = (L_{\mathcal{A}}(n) - E_{\mathrm{BHN}}(n))/2^{n-2}$. The six root certificates and profile arithmetic support the envelope; the plotted explicit-artifact range is not a statement that every descendant in that range was independently rechecked in this audit.

Six verified roots suffice in the evaluated range. The selected witnesses use only six root states (Table 12). For each, the audit independently reads an explicit edge list, checks endpoint ranges, adjacency, duplicates and every coordinate square, and recomputes the complete direction profile. Every dimension's maximizing value is attained by a descendant or slice of one of these checked roots. This supports the plotted construction envelope without asserting that all 198 root files, all quotient metadata, or explicit high-dimensional descendants were reverified. Six roots are sufficient for this interval; no minimality claim is made.

Table 12: Verified root profiles realizing the archive envelope over Q9–Q16384. Witness counts use a fixed archive order to resolve ties and do not measure unique necessity or statistical frequency.

Root n	Edges	Sorted direction profile	Selected dimensions
7	304	$[48^{\times 3}, 40^{\times 4}]$	2,551
10	3,328	$[352^{\times 5}, 320^{\times 4}, 288]$	9
11	7,296	$[704^{\times 4}, 640^{\times 7}]$	8,495
15	152,576	$[12800, 9984^{\times 14}]$	32
15	152,576	$[12800, 10240^{\times 7}, 9728^{\times 7}]$	4,585
16	322,048	$[25600, 19968^{\times 14}, 16896]$	704

Why the full recursive profile matters. The two Q_{15} roots in Table 12 have identical $E = 152{,}576$ and $p_1 = 12{,}800$, but different lower direction loads. Both occur among selected witnesses; these counts depend on tie-breaking and do not prove that both are indispensable. Their different profiles illustrate the need to retain the full recurrence input rather than only the immediate edge count and its leading direction.

Phase law in the audited blocks. Write $x = n/4^a$ for the phase of dimension n in its scale block $[4^a, 4^{a+1})$, and call a dimension a gain when $L_{\mathcal{A}}(n) > E_{\mathrm{BHN}}(n)$ and a tie when equality holds. In every fully audited regular block $4 \le a \le 6$ (exact rational arithmetic) and in the continued blocks $7 \le a \le 10$ (lattice-snapped double precision, validated against the exact series through Q_{16384}), the audit observes the same pattern: all dimensions with $\frac{35}{32} < x < \frac{29}{8}$ are gains except those at offsets $h, h+3, t-3$ (dimensions $\equiv 1 \pmod 3$), and all dimensions with $x < \frac{35}{32}$ or $x > \frac{29}{8}$ are ties except the dimension at offset $t+1$ ($\equiv 2 \pmod 3$), where $h = \frac{3}{32}4^a$ and $t = \frac{21}{8}4^a$ mark the window boundaries. Blocks below $a = 4$ deviate by single dimensions: at $a = 2$ the boundary $\frac{35}{32}4^2 = 17.5$ is not an integer dimension, and at $a = 3$ the offset $h + 3$ (dimension 73) is a gain rather than a tie. The crest of the gain window is first attained at $n = 23 \cdot 4^{a-2}$ in every block

$2 \le a \le 10$, with value equal to the closed form $\frac{2^a}{24} + \frac{1}{3 \cdot 2^a}$ (exact for $a \le 6$, validated continuation for $7 \le a \le 10$), whose successive ratios $1.50, 1.83, 1.955, 1.988, \ldots$ approach two.

Exact block-boundary gain rate. Let $N(R) = 4^R - 8$ count the evaluated dimensions up to 4^R and $W(R)$ the strict gains among them. For $a \ge 4$ both window boundaries are integer dimensions, and block a splits into a head tie window of $\frac{3}{32}4^a$ dimensions, a gain zone of $\frac{81}{32}4^a$ dimensions containing exactly three ties at offsets $h, h+3, t-3$, and a tail tie window of $\frac{3}{8}4^a$ dimensions containing exactly one gain at offset $t+1$. Hence $W_a = \left(\frac{81}{32}4^a - 3\right) + 1 = \frac{81}{32}4^a - 2$: the constant -2 records the four exception dimensions of each block. The irregular blocks below $a = 4$ contribute 203 gains over the 247 dimensions below 4^4, so

$$W(R) = 203 + \sum_{a=4}^{R-1} \left(\tfrac{81}{32}4^a - 2 \right) = \tfrac{27}{32}4^R - 2R - 5, \qquad f(4^R) = \frac{W(R)}{4^R - 8} = \frac{27}{32} - \frac{2R - \frac{7}{4}}{4^R - 8}.$$

The identity holds for every $4 \le R \le 11$: for $R \le 7$ both sides come from exact rational arithmetic, and for $R \le 11$ the required per-block counts are verified for $a \le 10$ in the lattice-snapped continuation. The deficit decays as $\approx 2R \cdot 4^{-R}$, so the value 0.843002 at $R = 7$ is accounted for rather than merely measured; the continuation computed from the counted blocks is $0.843533, 0.843688, 0.843733, 0.843745$ at $R = 8, 9, 10, 11$, and only $f(4^{12}) = 0.8437487$ lies beyond the audited range.

Conditional asymptotic consequences. Suppose the observed pattern persists for all a (*block self-similarity*): each block contributes $\frac{81}{32}4^a - 2$ gains and no deficits. Then the cumulative gain fraction $f(N)$ satisfies: along block boundaries, $f(4^R) \to \frac{27}{32} = 0.843750$; the full sequence does not converge, since adding a gain strictly raises f and adding a tie strictly lowers it, so per-block extrema occur at the window boundaries, giving $\liminf f = \frac{27}{35} \approx 0.771429$ (at $x = \frac{35}{32}$) and $\limsup f = \frac{27}{29} \approx 0.931034$ (at $x = \frac{29}{8}$); in particular the gain set has no natural density, while its logarithmic density exists and equals $\ln \frac{116}{35} \big/ \ln 4 \approx 0.864349$, a different averaging notion. The block-boundary limit and the oscillation bounds follow from geometric-series bookkeeping applied to the per-block counts; the monotonicity observation is unconditional. These consequences describe the observed recurrence law, not a proved property of all future blocks; the matching guarantee of Theorem 1 is independent of them.

Finite-range growth and attribution. Successive observed crest values are $1/4, 3/8, 11/16, 43/32, 171/64$, first attained at $n = 23, 92, 368, 1472, 5888$. These are exact finite-range values, continued numerically by $\frac{683}{128}, \frac{2731}{256}, \frac{10923}{512}, \frac{43691}{1024}$ at $Q_{23552}, Q_{94208}, Q_{376832}, Q_{1507328}$. They do not by themselves prove an asymptotic growth law or a strict gain at every future dimension. The universal non-inferiority statement instead follows from Theorem 1. The archive comparison concerns the refined BHN benchmark, not an exhaustive survey of current best-known constructions. The states' existence and recursive consequences also do not establish an efficiency advantage for the AI controller.

Data and reproduction. The source is the archived-profile dataset shipped in the supplementary package (`seedlib_318/`): `seeds_318.jsonl` records one row per state key, `six_roots.json` the six envelope roots, and `envelope_q16384_series.json` the dimension-wise values, with SHA-256 hashes in `cert_manifest.csv` and `SHA256SUMS.txt`. Running `seedlib_318/tools/verify_package.py` regenerates the envelope check from these profiles using exact rational arithmetic, without the generation pipeline.

R Hybrid-System Implementation and Recursive Evaluation

Controller interfaces and learning. The controller tuple $\Pi = (\theta, \hat{V}, \mathcal{G})$ of Equation equation 3 is part of the hybrid system, which also includes the executable operator algebra, constraint solver, mechanism library, outcome archive, and analytic recursive evaluator. The frozen proposer supplies parameterized actions. Each executed action contributes an outcome record with its status and cost; only verified feasible outputs contribute construction data and recursive fingerprints. The head's features, bootstrap target, and refit protocol are described in Appendix L.

Mechanism admission and representation guards. The same-dimension promotion criteria require certified trials, at least three evaluations, a strict gain or no losses, comparison with a degenerate parameterization, and reproduction on a held-out state. These criteria concern the tested cases; they do not guarantee improvement on an unseen state. Recursive operators have an exact neighborhood encoding by Proposition 1, but feasible outcomes must still be verified and cannot be called optimal without an OPTIMAL solver status. Implementations preserving minimum code distance four also reject weight-three kernel words in the augmented column set. This is stronger than the nonzero, distinct-column condition required by the square encoding.

An earlier cost-allocation design. An earlier design ranked batches by predicted value relative to predicted cost. The reported leave-one-state-out R^2 values are 0.31 for cost and 0.11 for value. The archived allocation comparisons report -7.3% at a tight budget and at most $+18.8\%$ over 30 seeds, with the latter within two standard deviations of uniform allocation. The implemented execute-all batch protocol does not use that allocation policy. These historical measurements are separate from the value-head settings in Table 13.

Scale-free recursive evaluation. For a lift node in Q_m, put $\varepsilon = E/2^m$ and $u_i = p_i/2^m$, with u_1 a maximum load. Equation equation 5 gives

$$\varepsilon' = \varepsilon + u_1 + \tfrac{1}{2}, \qquad u' = \{u_2, \dots, u_m\} \cup \{\tfrac{1}{8} + \tfrac{u_1}{2}\}^{\times 4}. \tag{22}$$

For example, $E'/2^{m+3} = E/2^m + p_1/2^m + 1/2$; dividing the inherited and four new direction counts by 2^{m+3} gives the multiset update. Maintaining extrema with a suitable ordered data structure permits logarithmic-time updates in the number of stored loads, excluding the bit cost of exact arithmetic. Reaching a target dimension still requires the intervening lift steps; a wall-time claim additionally needs an implementation, arithmetic mode, hardware, and measured timing protocol.

Absolute targets and coordinate restriction. The score uses $N_j = 1002 + 3(j-1)$ for $1 \le j \le 12$ for every candidate. A seed in dimension n has direct lift nodes $n + 3h$. If N_j is outside this residue class, use the next node and restrict one or two coordinates. For instance, a Q_{11} seed reaches Q_{1004} on its chain and uses two coordinate restrictions for target Q_{1002}. The ceiling in Equation equation 8 gives a construction lower bound, not necessarily the exact largest slice size. Targets congruent to zero modulo three are a common evaluation grid, not a general proof that other residue classes cannot distinguish candidates. Increasing weights emphasize later targets across the twelve-node grid.

Same-dimension scores and dimension jumps. The term $c(a)$ in Equation equation 4 applies to same-dimension moves. For fixed-rank extension by k columns, write $|F'| = |F| + \Delta$. Then

$$\frac{E' - E}{|C|} = (2^k - 1)|F| + 2^k\Delta, \qquad \frac{E' - 2^k E}{2^k |C|} = \Delta.$$

The first expression includes replicated parent edges; the second is net orbit gain. Dimension jumps in the stated protocol compare already verified candidates through common-target Φ. No extra solve is required for selection, but candidate-generation time remains part of the search cost. Historical scores and the configured stopping and warm-up rules must be reported as actually used, rather than inferred from the resulting envelope.

Certificate evidence for the mode-level table. Table 2 reports stored certificates rechecked independently by a checker importing no generation code. The checks enumerate coordinate two-faces, test adjacency and duplicate edges, and find no C_4. For Q_{17}, the full cube has $\binom{17}{2}2^{15} = 4{,}456{,}448$ such faces. These checks certify the tabulated lower bounds. The comparison with the archived no-feedback envelope is separately reported in Appendix M.

S A CONDITIONAL COUNTING BOUND FOR WIRING FAMILIES

A restricted construction family can have an analytic ceiling, but a stalled search does not establish one. The following counting argument states the extra condition needed for such a ceiling. It is separate from the quotient constraints and from the all-dimension lower bound.

Table 13: Low-dimensional search outputs. The two head settings start from the same 43 seeds; the archived baseline uses 61. Protocol and budget records accompany the component studies in Appendices L and M.

Quantity	Archived baseline	$\gamma = 0$	$\gamma = 0.5$
Seeds started	61	43	43
Certified executions	1300	528	560
Distinct states (n, E, P)	—	338	371
Action kinds exercised	—	12	12
States reaching their record	5	1	2
States above their record	0	0	0
Mean solver seconds per state	—	45.3	46.0

Explicit assumptions. Let $m \geq 2$, and assign each edge f of the gadget Q_m a subset S_f of a common universe U of size $N = 2^{k-1}$. Suppose $|S_f| \leq \alpha$ for every edge. In a wiring model this bound may come from independence in a specified conflict graph. Impose the additional *disjoint-pair condition*: on every coordinate two-face, at least two of the four assigned sets are disjoint. This is a restriction defining the family considered here. It does not follow from the four sets having empty common intersection: the four sets $S_i = \{1, 2, 3, 4\} \setminus \{i\}$ have empty common intersection and intersect pairwise.

Proposition 5 (Counting bound under the disjoint-pair condition)**.** *Under these assumptions, the total wiring size obeys*

$$\Sigma := \sum_{f \in E(Q_m)} |S_f| \leq m 2^{m-2} \big(\alpha + N/2\big). \tag{23}$$

Proof. On each two-face, choose a disjoint pair. Its two sets have total size at most N; the two remaining sets contribute at most 2α. Sum this four-edge inequality over the $\binom{m}{2} 2^{m-2}$ two-faces. Every gadget edge belongs to exactly $m - 1$ such faces, so

$$(m-1)\Sigma \leq \binom{m}{2} 2^{m-2}(N + 2\alpha).$$

Dividing by $m - 1$ proves the bound. No assumption on how often an edge is chosen as a witness is needed. □

Conditional numerical consequence. For $m = 4$, $k = 7$, and a certified cardinality bound $\alpha \leq 48$, the proposition gives $\Sigma \leq 16(48 + 32) = 1280$. If the remaining parent contribution is $8(304 - 48) = 2048$, the resulting construction satisfies

$$E' \leq 2048 + 1280 = 3328. \tag{24}$$

This matches the refined BHN count at Q_{10} as an upper bound for the stated restricted family. Applying it to a concrete wiring model requires establishing both its cardinality bound and its disjoint-pair condition. Equality in the count alone neither proves attainability in that model nor identifies a unique extremal wiring.

Implication for search scope. Once those premises are certified, exceeding the bound requires leaving that restricted family. Without them, this argument does not prove exhaustion of the general BHN gadget family, the quotient representation, or all C_4-free subgraphs. In particular, changing a wiring template and releasing quotient symmetry are different changes of search space. The unconditional construction guarantee in Theorem 1 uses the legal BHN lift and slicing, and does not depend on this conditional upper bound.