

A New Upper Bound for the First Autocorrelation Inequality and 306 Circle-Packing Records via Active-Set Escape from Equioscillating and KKT Optima

Lindell S. Cumes

QWNTL Labs

Project page: <https://qwntl.com/paper-new-constructions>

Abstract

As of October 30, 2026, the Einstein Arena’s leader on the first autocorrelation constant C_1 stands at 1.5027434867, and MILO’s evolved agent harnesses report 1.5027435984. Under a two-hour, four-thread budget, Claude Opus 5.5 surpasses none of the Endless Exam’s 30 published frontiers. In the published optima of these minimax problems, the maximum is attained, to within the venue’s minimum improvement, at thousands of points, or the circles sit at a KKT point. Any improvement larger than that minimum must lower the value at every one of these points at once. Active-set escape, which we introduce, is a search in which each move changes the near-active set and a global re-melt and a second-order solve finish the new basin. We gave Claude Opus 5.5 agents hours of autonomous work and an exact evaluator, and they wrote these moves themselves. Three agents sharing an archive of verified attempts, with no weight update, certify $C_1 \leq 1.50266484 \dots$ after 40.75 working hours. The bound lies 7,865 times the Arena’s minimum improvement below its leader. A single agent with no outer loop also passes the leader, by 7,200 times the minimum improvement. The same moves lower the signed autocorrelation constant below MILO’s value and set the best sum of radii for unequal circles in a square at 306 Packomania sizes up to $N = 10,000$. A preregistered classical optimizer started from the same constructions stalls after gains about three thousand times smaller. Witnesses and exact verifiers ship with the paper.

Release. <https://qwntl.com/paper-new-constructions>. The arXiv ancillary files hold every witness (the C_1 , C_3^+ , and Erdős step functions and all 306 packings), the exact checkers, and a driver that rechecks every value in under a minute on one CPU core.

1 Introduction

For N circles in a square, every feasible list of centers and radii gives a lower bound on the largest sum of radii. For the first autocorrelation constant C_1 , every nonnegative step function on $[-\frac{1}{4}, \frac{1}{4}]$ gives an upper bound. Exact arithmetic can check such a bound from its example alone, without trusting the search that produced it. Machine-discovery systems search for such examples: FunSearch built cap-set constructions [32], and AlphaEvolve built autocorrelation and circle-packing constructions [17, 30]. The public Einstein Arena [4, 13] and Packomania’s tables [37] list the current records.

On October 30, 2026, the Einstein Arena’s leader on C_1 , entry #2615, stands at 1.5027434867. Among published systems, two learn a harness. MILO evolves complete agent harnesses and reaches $C_1 \leq 1.5027435984$ [21]. HarnessTTT trains a harness proposer at test time and reaches 1.503778 [27]. Two others change model weights: TTT-Discover updates them during the search and reaches 1.50286 [43], and ThetaEvolve, built for test-time reinforcement

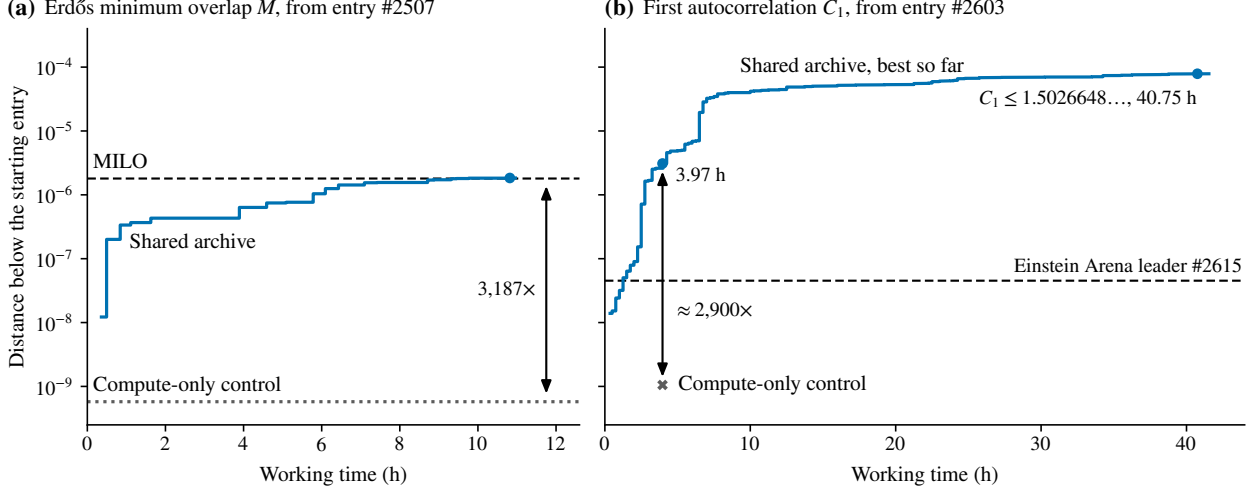


Figure 1: **From the same stationary entries, active-set escape goes about 3,000 times further than classical descent.** Each panel plots the distance below the Einstein Arena entry the search started from, on a log scale, against working time, the matched axis of a preregistered study on one machine type. (a) Erdős’s M from entry #2507. The shared archive’s 19-attempt lineage (three Claude Opus 5.5 agents) ends at the certified construction (filled circle, 10.82 h, 1.836×10^{-6} below the entry). The dotted line is the compute-only control from the same entry, whose best stayed 5.76×10^{-10} below it. The dashed line is MILO’s printed value. (b) C_1 from entry #2603: the same run’s best so far. Filled circles mark the certificates at 3.97 h, 3.10×10^{-6} below the entry, and at 40.75 h ($C_1 \leq 1.5026648\dots$). At the first, the control (cross) stood 1.05×10^{-9} below the entry. The dashed line marks #2615, the Arena’s leader on October 30, 2026, 7.86×10^{-5} above the second certificate. Each arrow’s label is the ratio of the certified distance to the control’s. Curves are the protected evaluator’s conservative floating-point scores; the certificates are exact.

learning, reports 1.503133 from a run started at AlphaEvolve’s construction [41]. Test-time memory evolution reaches 1.502745 [25]. AgentDiscover, in which a coding agent (Claude Opus 5 on these problems) plans its own search over a database of its attempts, posted the entry #2603 from which our C_1 searches started [15]. Under a tight budget Claude Opus 5.5 does not pass published values: the Endless Exam gives it code and web access and a two-hour limit on four CPU threads, and on that budget it matches 23 of the benchmark’s 30 published frontiers and surpasses none, on problem families other than these [44].

The C_1 entry #2603 attains its maximum, to within 10^{-8} , at 24,397 of its 131,071 knots, and any improvement on it larger than that tolerance must lower all of them at once. The best packings at large N are KKT points of their active constraints, and local polishing gains less than 10^{-9} from them. We ran a preregistered classical optimizer from the Erdős and C_1 entries, both stationary (Definition 1). It lowered them by 5.8×10^{-10} and 1.05×10^{-9} and then stalled (Figure 1).

Active-set escape leaves such points with moves that change the active set, meaning the knots that attain the maximum or, for a packing, the contacts and wall touches that hold with equality (Figure 2). We gave Claude Opus 5.5 agents in Claude Code hours of autonomous work on a pool of machines and an evaluator that scores in exact rational arithmetic, while keeping the model weights fixed and using no learned harness, and the agents wrote the moves themselves, as programs that call no language model. In the shared-archive Erdős lineage they broke symmetry along the unstable modes of a saddle and then crossed independent lineages. Other searches changed the discretization that defines the active set: they moved functions to finer grids, by refinement before a re-melt or by rebuilding a published coarse function at full resolution, and they re-gridded populations onto grids nested in neither parent’s. The packing agent re-seated the circles around a defect before relaxing the whole packing. The packings and the C_3^+ bound came from single agents with no outer loop, and a second C_1 certificate came from another.

On C_1 , three agents sharing an archive of verified attempts certified $C_1 \leq 1.5027404359\dots$ after 3.97 working hours and $C_1 \leq 1.50266484\dots$ after 40.75. The second bound is 7.86×10^{-5} below the Einstein Arena’s leader, 7,865 times its minimum improvement, and below each published value as of October 30, 2026 (Table 1). A single agent escaped the KKT points of the packing table with defect moves. Its packings exceed the best published sum of radii at 306 sizes

Table 1: **Our bounds on C_1 and C_3^+ and our packings at 306 sizes, against the published values as of October 30, 2026.** Lower is better for the constants and higher for packings. Our bounds are exact rationals, shown truncated; other values are as printed. Parentheses give the margin to our best value, computed exactly from the full printed decimal. Values from other sources are below the rule. “—” means no reported value.

Source (date)	$C_1 \downarrow$	$C_3^+ \downarrow$	Sum of radii, $N = 137\text{--}10,000 \uparrow$
This paper (exact)	1.5026648407... (40.75 h); 1.5027404359... (3.97 h)	1.4488670313... (97,000 cells)	above every published value at each of the 306 sizes
Einstein Arena leaders (Oct. 30)	1.5027434867 #2615 ($+7.86 \times 10^{-5}$)	1.4491097893 #2617 ($+2.43 \times 10^{-4}$)	—
MILO [21] (Sep. 29)	1.5027435984 ($+7.88 \times 10^{-5}$)	1.4488860143 ($+1.90 \times 10^{-5}$)	—
TTME [25] (Oct. 23)	1.502745 ($+8.02 \times 10^{-5}$)	—	—
TTT-Discover [43]	1.50286 ($+1.95 \times 10^{-4}$)	—	$N = 26, 32$ only
ThetaEvolve [41]	1.503133 ($+4.68 \times 10^{-4}$)	—	$N = 26$ only
AlphaEvolve [17]	1.5032 ($+5.35 \times 10^{-4}$)	1.4557 ($+6.83 \times 10^{-3}$)	$N = 26, 32$ only
HarnessTTT [27] (Oct. 12)	1.503778 ($+1.11 \times 10^{-3}$)	—	$N = 26$ only
The Station [9]	1.504473 ($+1.81 \times 10^{-3}$)	—	—
Liang [24] (head Oct. 5)	—	—	previous best at each of the 306 sizes; ours higher by 3.7×10^{-7} to 5.1×10^{-2}
Packomania csqv [37] (Sep. 30)	—	—	ours higher at each of the 306 sizes, by 3.5×10^{-4} to 9.9×10^{-2}

between $N = 137$ and 10,000 on the same date (Figure 3). They include 58 of the 60 largest sizes Packomania lists. For the signed autocorrelation constant, another single agent reached $C_3^+ \leq 1.4488670313 \dots$ on 97,000 cells, below MILO and below each Einstein Arena entry listed on October 30, 2026.

On Erdős’s minimum overlap problem we can follow the mechanism step by step, against a control from the same start. The mirror-symmetric optimum is a saddle with 75 to 85 unstable antisymmetric directions. The agents’ 19-step descent broke symmetry along them and then crossed independent lineages. It went 3,187 times as far below Einstein Arena entry #2507 as that control did. The resulting certificate, timestamped October 8, 2026 in the archive, is the earliest construction below MILO’s value in the dated record; the first public one followed on October 16 (Table 12). A re-grid of MILO’s published construction then certified $M \leq 0.38085654787558 \dots$, below the value reported by each paper compared here as of October 30, 2026.

We certify each value in exact arithmetic (Section 4) and trace it to an escape move from a stationary published optimum (Sections 2 and 3). The preregistered control (Section 5) applies the classical toolbox, perturbation restarts included, to the same entries. In Section 6 we compare the agents with that control to separate the effect of the moves from the effect of compute and polishing.

2 Problems, certificates, and the stationarity barrier

2.1 Problem statement

We write each task as a triple $\tau = (d, s_0, V)$: a problem description d , a public starting construction s_0 , and an exact verifier V that maps a construction file to a rational score or rejects it. An agent is a pair $A = \langle M, H \rangle$ of a model M , here Claude Opus 5.5, and a harness H , here Claude Code with a machine pool, the protected evaluator, and, on two problems, a shared archive. We measure the budget B in working hours. The objective is discovery:

$$\text{minimize} \quad \min_{x \in \mathcal{X}_B(A, \tau)} V(x), \quad (1)$$

where $\mathcal{X}_B$ is the set of constructions the agent submits within B , and we count a result only if $V(x)$ is strictly better than every published value on the same instance. For packings the objective is a maximum.

2.2 The four problems

First autocorrelation constant. C_1 is the largest constant with $\max_t (f * f)(t) \geq C_1 (\int f)^2$ for every nonnegative f supported on $[-\frac{1}{4}, \frac{1}{4}]$ [10, 28]. Each such f gives $C_1 \leq \|f * f\|_\infty / \|f\|_1^2$. Schinzel and Schmidt conjectured the

value $\pi/2$, attained by $f_0(x) = (2x + \frac{1}{2})^{-1/2}$, which is singular at the left endpoint. Matolcsi and Vinuesa refuted the conjecture with a step function at 1.50972. They found no natural candidate for the extremal function [28]. The best lower bound is $C_1 \geq 1.2937$ [22].

Signed autocorrelation constant. C_3^+ is the same constant for real-valued f with $\int f \neq 0$, still measured by $\max_t (f * f)(t)$. Negative values lower it: Matolcsi and Vinuesa found a signed step function at 1.45810. Since $\int f * f = (\int f)^2$ and $f * f$ lives on an interval of length one, $C_3^+ \geq 1$. This is the constant of the corrected AlphaEvolve notebook, Einstein Arena problem 4, and MILO.

Unequal circles in a square. For Packomania’s sum-of-radii problem (csqv, its table name), place N circles with centers (x_i, y_i) and radii $r_i > 0$ in the unit square and maximize $\sum_i r_i$ subject to

$$r_i \leq x_i, y_i \leq 1 - r_i, \quad (x_i - x_j)^2 + (y_i - y_j)^2 \geq (r_i + r_j)^2.$$

Packomania lists every $N \leq 500$ except 493. It also lists 84 larger sizes up to 10,000, of which 78 are perfect squares.

Erdős’s minimum overlap problem. Split $\{1, \dots, 2n\}$ into sets A and B of size n , let M_k count pairs $(a, b) \in A \times B$ with $a - b = k$, put $M(n) = \min_{A,B} \max_k M_k$, and let $M = \lim_n M(n)/n$ [14]. Swinnerton-Dyer proved, in a result published by Haugland, that the limit exists and that

$$M \leq \sup_t O_h(t), \quad O_h(t) = \int_{\mathbb{R}} h(x)(\mathbf{1}_{[0,2]} - h)(x+t) dx, \quad (2)$$

for every step function $h : [0, 2] \rightarrow [0, 1]$ with $\int h = 1$ [19, 20]. A step function on n cells of width $w = 2/n$ is admissible when its heights lie in $[0, 1]$ and sum to $n/2$. The best lower bound published on arXiv is $M > 0.380552$, with a Lean-verified proof [9]. Higher lower bounds, up to 0.3806274, have been claimed on the Einstein Arena’s public discussion threads without refereed proofs [13].

Knot reduction. A step function turns each supremum into a finite maximum, and a certificate then becomes a computation on integers.

Lemma 1 (Knot reduction). *Let h be a step function on n cells of width w . Then $O_h(t)$ is linear between consecutive multiples of w , so $\sup_t O_h(t) = \max_k C_k$ with $C_k = w \sum_i h_i(1 - h_{i+k})$, the sum over $1 \leq i, i+k \leq n$. Likewise, for f with heights a_i on n cells of $[-\frac{1}{4}, \frac{1}{4}]$, $(f * f)(t)$ is linear between the knots $t_k = -\frac{1}{2} + (k+1)/(2n)$, where it equals $c_k/(2n)$ with $c_k = \sum_{i+j=k} a_i a_j$. With $S = \sum_i a_i \neq 0$,*

$$\frac{\max_t (f * f)(t)}{(\int f)^2} = \frac{2n \max_k c_k}{S^2}.$$

Proof. For t between consecutive knots, each pair of cells overlaps on an interval whose length is affine in t , and the integrands are constant on cells. Both functions are therefore affine on each such interval, and a piecewise-linear function attains its supremum at a breakpoint. The identity $\int f = S/(2n)$ gives the ratio. $\square$

2.3 Why local descent stops at the published optima

After knot reduction each analytic problem minimizes $F(h) = \max_k C_k(h)$ over a polyhedron P of admissible heights, with each C_k a quadratic. For C_1 , whose ratio is scale-invariant, P fixes the sum of the heights. We bound descent on P with two lemmas: an exact first-order condition, which is local, and a bound that holds for steps of any size and covers improvements as large as the venue’s minimum.

Lemma 2 (First-order condition for a finite maximum). *Let $F = \max_k C_k$ with each C_k continuously differentiable, and let $\mathcal{A}(h) = \{k : C_k(h) = F(h)\}$ be the active set at $h \in P$. For every feasible direction d at h ,*

$$F'(h; d) = \max_{k \in \mathcal{A}(h)} \langle \nabla C_k(h), d \rangle.$$

A feasible direction therefore lowers F to first order exactly when it lowers every active C_k at once. If no feasible direction does, there are weights $\lambda_k \geq 0$ on $\mathcal{A}(h)$ with $\sum_k \lambda_k = 1$ such that $-\sum_k \lambda_k \nabla C_k(h)$ lies in the normal cone of P at h .

Proof. The formula for $F'(h; d)$ is Danskin’s theorem for a maximum over a finite index set [12]. The directions that lower F form the solution set of the strict linear system $\langle \nabla C_k(h), d \rangle < 0$ for $k \in \mathcal{A}(h)$ with d in the polyhedral tangent cone of P . By linear-programming duality (Gordan’s theorem of the alternative relative to a polyhedral cone) the system is infeasible exactly when the stated weights exist. $\square$

Lemma 3 (Near-active knots bind every improvement). *For $\varepsilon \geq 0$ let $\mathcal{A}_\varepsilon(h) = \{k : C_k(h) \geq F(h) - \varepsilon\}$. If $F(h') \leq F(h) - \delta$ with $\delta > \varepsilon$, then $C_k(h') \leq C_k(h) - (\delta - \varepsilon)$ for every $k \in \mathcal{A}_\varepsilon(h)$.*

Proof. For $k \in \mathcal{A}_\varepsilon(h)$, $C_k(h') \leq F(h') \leq F(h) - \delta \leq C_k(h) + \varepsilon - \delta$. $\square$

Lemma 3 needs no smallness of $h' - h$: it binds every construction, near or far, that improves on h by more than ε . A construction is *equioscillating* when its maximum is attained, to within ε , at many knots; the term comes from Chebyshev approximation, where the error of a best approximation equioscillates on a reference set [31]. The Einstein Arena’s C_1 entry #2603 attains its maximum, to within 10^{-8} , at 24,397 knots, so any construction that clears the Arena’s minimum improvement of 10^{-8} by any margin lowers all 24,397 of them. At the Erdős entry #2507 the same holds for 2,428 shifts.

Definition 1 (Stationary construction; active-set escape). A construction is *stationary at scale ε* when local descent from it gains less than ε and its near-active set $\mathcal{A}_\varepsilon$ is large (for packings: when it is a KKT point at which local polishing gains less than ε). An *escape move* is a perturbation that changes the near-active set, or the discretization that defines it, followed by a global re-melt that lets a new one form. *Active-set escape* is search built from escape moves and second-order finishes.

We measure stationarity at the venue’s scale, so it does not require an exact first-order stationary point. The 40.75-hour C_1 certificate of Section 6.1 has a unique maximizing knot, so exact first-order descent from it exists, yet 97,349 of its knots lie within 10^{-8} of the maximum, and by Lemma 3 its next improvement past 10^{-8} must lower all of them.

For packings the barrier comes from rigidity: a packing is a point $z \in \mathbb{R}^{3N}$ of centers and radii, and each contact or wall touch is an active constraint.

Lemma 4 (Isostatic rigidity). *Let z be a feasible packing at which exactly $3N$ constraints are active and their gradients are linearly independent. Then z is the only point near z at which all $3N$ of these constraints are active, and any move that increases $\sum_i r_i$ releases at least one of them.*

Proof. The active constraints define a map $g : \mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N}$ with $g(z) = 0$ and a nonsingular Jacobian. By the inverse function theorem g is injective near z , so z is the only nearby zero of g . A nearby packing with a larger sum differs from z , so it is not a zero of g . $\square$

We checked the new packings at $N = 137, 529$, and 1681 . Each has exactly $3N$ active constraints at tolerance 10^{-9} , and Newton’s method on the active system converges in one or two steps to 10^{-16} , as it does when the active Jacobian is nonsingular. Where that Jacobian is nonsingular, Lemma 4 makes the next improvement release a contact. The published large- N packings they replaced were KKT points, stationary at scale 10^{-9} . Sequential linear programming and Newton polishing gained less than 10^{-9} from them, and the gains above $N = 1600$ came from defect moves that rebuilt the contact graph around one or several circles.

3 Active-set escape: harness, information flow, and moves

3.1 Harness and information-flow contract

Agents. We ran each agent as Claude Opus 5.5 in Claude Code 2.1.289. An agent reads the task material and code and uses public constructions, and it writes and runs its own programs, with no language model inside their numerical loops. We gave the matched study’s agents no web access. The single agents outside the study, which produced the packings and the C_3^+ and Erdős re-grid constructions, could also read public papers and pages. The shared-archive agents ran at effort setting “xhigh” with a 4-hour session cap.

Harness

Claude Opus 5.5 agents in Claude Code (effort xhigh on the shared-archive runs); machine pool; hash-pinned exact-rational evaluator; shared archive of verified attempts (M, C_1)

Active-set escape, in code the agents write and run; no language model inside the numerical loops

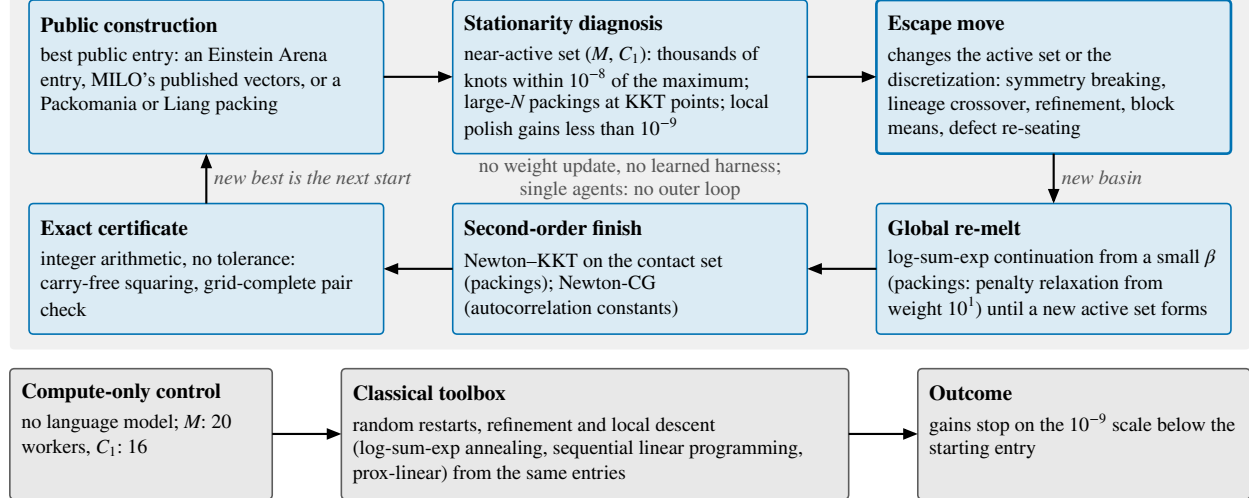


Figure 2: **An escape move changes the active set at which the classical toolbox stalls.** Top: the harness. Middle, clockwise: the loop the agents wrote on each problem. A public construction is diagnosed as stationary (thousands of knots within 10^{-8} of the maximum on M and C_1 ; published large- N packings at KKT points where local polish gains less than 10^{-9}). An escape move changes the active set or the discretization. A global re-melt, by log-sum-exp continuation or, for packings, by penalty relaxation, then lets a new active set form, and a second-order solve finishes the basin. An exact certificate fixes the value, and the new best is the next start. Bottom: the preregistered compute-only control runs the classical toolbox from the same entries. It keeps the active set and stalls on the 10^{-9} scale.

Fixed components. The model weights and the evaluator stay fixed for the whole run, as do the problem definitions and the public starting material; only the constructions and the agents’ code change. The evaluator’s bytes are hash-pinned, and it refuses to run if they change. We supplied the problem definitions and the starting material identically to all arms.

Shared archive. On M and C_1 , three agents per problem share an append-only, HMAC-chained archive that holds verified attempts together with tools and notes and persists across their sessions for the whole run. Each attempt enters with its score from the protected evaluator and a record of its parent and cited items, and each new best is requelified in exact arithmetic. At the start of each session the archive gives an agent a briefing of at most 9,000 characters. The briefing ranks verified insights and tools, along with parent attempts, by verified reuse and by the gains of their descendants across independent search lineages. It adds one-line constraints distilled from failed attempts and up to four techniques offered from the arm’s other problems, and it rotates a research direction across agents. A language model compiles each verified gain into an insight (claim, mechanism, premises), and a second one distills failures into constraints; the archive never generates or edits a construction. The other runs use one agent with its own workspace and notes. The study’s single agent worked in one long conversation. The packing and C_3^+ agents worked in two sessions each, the second continuing from the agent’s own report.

Task material. Each analytic task gives the problem statement and the public record constructions with their history, along with the file format and the evaluator’s usage. Its six research directions name technique families in the standard vocabulary of the problem. Some work on a construction in hand, with minimax polishing over its near-active knots, refinement and log-sum-exp surrogates with annealing; others reach past it, to structure and symmetry and to restarts and recombination, and one asks for shared tooling. Appendix A.2 gives the C_1 and Erdős text verbatim. The specifics came from the agents’ own runs: the saddle and its unstable modes and the block means of a published vector were their findings, and the non-nested re-grids and the specific move sequences of Tables 5 and 10 were their moves. The Erdős re-grid run started from MILO’s published vector, as the agent’s instructions directed, after the same agent had

Algorithm 1: **Active-set escape.** The pattern common to the agents’ searches on C_1 , C_3^+ , M , and the packings.

Input: a public construction s_0 , the exact verifier V , the archive (shared or private).

- 1: Count the active set, the knots within 10^{-8} of the maximum or the contacts and wall touches at tolerance 10^{-9} . The point is stationary if local polish gains less than the venue’s minimum improvement.
 - 2: Apply an escape move, one that changes the active set or the discretization. To change the active set directly, perturb along an unstable eigenvector or cross two lineages. To change the grid, refine m -fold or rebuild a coarse function at full resolution, or re-grid onto a grid nested in neither parent’s. For packings, re-randomize the circles around a defect.
 - 3: Re-melt by restarting the log-sum-exp continuation from a small β (for packings, the penalty relaxation from weight 10^1) so that a new active set forms.
 - 4: Finish to second order on the new active set, with Newton-CG and exact FFT Hessian-vector products or with Newton’s method on the active KKT system.
 - 5: Score with V . If the value beats the incumbent, requalify it exactly and post it with its parent and cited items. Claim only values strictly better than every published value.
-

decoded MILO’s C_3^+ vectors itself.

Information available to each actor. The agents see the archive and the protected evaluator’s score, a conservative floating-point value during search and an exact rational at requalification, while the compute-only control sees only its own scores. Within the study no arm sees another; the single agents outside it could read the study’s tools and constructions. We fix a construction by freezing its file with its hash, and we compute the certificate on that file.

Machines. We ran the shared-archive arm’s four problems and twelve agents on one 96-vCPU machine and capped each agent at 16 vCPU, while the single agents claimed machines from a shared pool, CPU machines for the packings and A100 GPUs for the autocorrelation searches.

Hash-pinned exact evaluator. The evaluators were written before the searches, and we tested them on adversarial cases that covered malformed numbers and tampering as well as tolerance-inflated radii and overflow. They reproduced each public record they were tested on. The final mode parses each decimal as an exact rational. The evaluator shrinks an overlapping packing before scoring it, and no violation can raise a score. No claimed file needed repair.

3.2 The search pattern

The agents built their searches on a log-sum-exp smoothing of the maximum. For knot values $C_k(h)$, let

$$\text{LSE}_\beta(h) = \frac{1}{\beta} \log \sum_k e^{\beta C_k(h)}, \quad \max_k C_k(h) \leq \text{LSE}_\beta(h) \leq \max_k C_k(h) + \frac{\log K}{\beta},$$

with K the number of knots. At small β the surrogate weights the knots nearly equally; at large β it concentrates on the active ones. A *re-melt* restarts the continuation from a small β after a perturbation, so the active set dissolves and re-forms. A *refinement* repeats each height m times; the score is unchanged, and the finer grid opens descent directions the coarse one lacked. Algorithm 1 gives the pattern common to the searches.

3.3 The escape moves

Each component of these moves is classical: basin hopping with local perturbations predates this paper [2, 18, 39], continuation and refinement predate it, and so do crossover and the Newton finish; MILO’s best Erdős optimizer already ends in an equioscillation Newton stage [21]. The task material named these technique families (Appendix A.2). We measure stationarity at the venue’s scale and find that the published optima of these problems are stationary there (Definition 1). The search then keeps the moves that change the near-active set. This measurement and selection rule are the new part of the method. In the searches reported here, the moves that kept the near-active set gained nothing past the venue’s minimum improvement (Table 8).

Refinement and a soft re-melt (C_1). One agent’s line of constructions reached 131,072 cells after 3.1 working hours, and a second agent fetched it from the archive, upsampled it two-fold to 262,144 cells, and ran a log-sum-exp ramp

from the soft temperature $\beta = 10^3$, multiplying β by 1.035 per step, with an evaluation kernel adapted from a third agent’s tool. The ramp’s best snapshot, near $\beta = 3 \times 10^7$, is the 3.97-hour certificate. The archive kept descending, and its best at 40.75 hours is a two-fold refinement of archive attempt 245, as the file header records.

Symmetrize, then break (M). Reflection $x \mapsto 2 - x$ maps admissible functions to admissible functions and preserves the bound, so the mirror-symmetric functions form an invariant subspace.

Lemma 5 (Mirror identity). *For admissible h and $0 < k < 2$, $O_h(k) - O_h(-k) = \int_0^k h - \int_{2-k}^2 h$. If $O_h(k) = O_h(-k)$ for all k in an interval I , then $h(x) = h(2 - x)$ for almost every $x \in I$.*

Proof. Write $O_h(t) = \int_{[0,2] \cap [-t,2-t]} h - \int h(x)h(x+t)dx$. The autocorrelation term is even in t . The first term is $\int_0^{2-k} h$ at $t = k$ and $\int_k^2 h$ at $t = -k$, and subtracting the common $\int_k^{2-k} h$ gives the identity. Differentiating $\int_0^k h = \int_{2-k}^2 h$ in k gives the symmetry. $\square$

Where the overlap equioscillates at $\pm k$ over an interval of shifts, the function must be mirror symmetric on that interval, so any asymmetry lies elsewhere. The symmetric optimum at 1,792 cells is poor, 2×10^{-5} above the Einstein Arena entry. It is also a saddle: the Hessian of LSE_β restricted to antisymmetric directions has 75 to 85 negative eigenvalues for β between 10^4 and 5×10^7 . The iteration optimizes inside the symmetric subspace and perturbs along antisymmetric directions, then re-anneals without the constraint and repeats from the new best.

Eigen-direction breaking and crossover (M). Draws along the unstable eigenvectors go lower than random noise from the same symmetric parent (Table 7). The last stage was a memetic genetic algorithm that one agent wrote and another ran unmodified. Its pool holds annealed optima at 1,792 cells. Children come from crossover and blends as well as from mirrored parents. Each child is re-melted from $\beta = 5 \times 10^4$ to 2×10^5 .

Block-mean reconstruction (C_3^+). MILO’s project page publishes two 3,200-number vectors. The agent showed that they are 8-cell block means of 25,600-point functions. The means alone score 1.6558, so any gain must come from structure inside the blocks. It searched over functions that share the means, $h = \text{up}(m) + u - \text{up}(\text{blockmean}(u))$, then polished with an annealed log-sum-exp Newton-CG method of its own whose Hessian-vector products are exact FFT convolutions,

$$\nabla^2 \text{LSE}_T(h) v = \left(\frac{1}{T} J^\top (\text{diag } w - ww^\top) J + 2 \text{Hankel}(w) \right) v,$$

with $T = 1/\beta$, w the softmax weights of the knots, and $Jv = 2(h * v)$.

Non-nested re-grids (C_3^+ and M). A re-grid is also an escape move: it changes the discretization that defines the active set. The agent remapped populations onto grids nested in neither parent’s grid, which destroys the grid-scale structure. The C_3^+ population scored 1.4909 at 100,000 cells and 1.5258 at 99,000 after remapping. Hot re-melts then recovered and passed its previous value. The same agent re-gridded an Erdős population built from MILO’s published 512-point vector onto 2,560 cells.

Defect re-seating (packings). The packing agent assembled Algorithm 2 (Appendix A.3) from standard global-optimization components [2, 18, 39], and added the diagnosis and the move itself. The move picks a circle whose radius is far from the median, which in a hexagonal array marks a defect, and re-randomizes the disc of circles around it, after which a penalty method relaxes the whole packing and a Newton solve on the new overlap set finishes it.

4 Certification contract

We read each decimal in a construction file as an exact rational and clear denominators, which turns the construction into integers A_i with sum S . The autocorrelation certificates then square a single integer.

Lemma 6 (Carry-free squaring). *Let $A_0, \dots, A_{n-1} \geq 0$ be integers and B an integer with $B > \max_i A_i \cdot \sum_i A_i$. Each coefficient of X^2 , for $X = \sum_i A_i B^i$, lies in $[0, B)$. The base- B digits of X^2 are therefore exactly the convolution coefficients $c_k = \sum_{i+j=k} A_i A_j$.*

Proof. $c_k = \sum_{i+j=k} A_i A_j \leq \max_i A_i \cdot \sum_j A_j < B$, so no coefficient carries into the next digit. $\square$

For the C_1 constructions the condition holds with $B = 2^{216}$. One exact GMP squaring then recovers all 524,287 coefficients in two to three seconds, and we take the maximum over all of them. Three moment identities check the result globally:

$$\sum_k c_k = S^2, \quad \sum_k k c_k = 2S \sum_i i A_i, \quad \sum_k k^2 c_k = 2S \sum_i i^2 A_i + 2(\sum_i i A_i)^2.$$

For M , put $B_j = 2S - nA_j$. Feasibility $0 \leq h_j \leq 1$ becomes $A_j \geq 0$ and $B_j \geq 0$, and Lemma 1 gives

$$\max_t O_h(t) = \frac{\max_k \sum_i A_i B_{i+k}}{2S^2},$$

which we evaluate by packed multiplication and again by direct summation at each shift. The C_3^+ checker squares a carry-free encoding of the signed integer heights and checks the coefficient-sum identity and eight direct sums, including both extrema.

For packings, we scale the literals to integers (X_i, Y_i, R_i) over a common power of ten S . The checker decides $R_i > 0$, $X_i - R_i \geq 0$, $Y_i - R_i \geq 0$, $S - X_i - R_i \geq 0$, $S - Y_i - R_i \geq 0$, and $(X_i - X_j)^2 + (Y_i - Y_j)^2 \geq (R_i + R_j)^2$, with no square root and no tolerance. The pair search must not miss a violating pair, and a grid makes it complete.

Lemma 7 (Grid completeness). *Assign each circle to the cell $(\lfloor X_i/g \rfloor, \lfloor Y_i/g \rfloor)$ of a square grid of integer width $g = 2 \max_i R_i$. Two circles whose cells differ by at least two in either coordinate do not overlap. Checking each circle against its own cell and the eight adjacent cells therefore checks each possibly overlapping pair.*

Proof. A cell difference of at least two along an axis forces a coordinate separation greater than $g \geq R_i + R_j$. $\square$

At $N = 10,000$ the lemma leaves 113,179 candidate pairs of 49,995,000. Above $N = 2500$ a faster checker sorts the circles by X and sweeps each one forward until $X_j - X_i \geq R_i + \max R$; an overlapping pair has $X_j - X_i < R_i + R_j$, so the sweep cannot stop before it.

Witnesses and verifiers. Each witness file holds one value or one circle ($x \ y \ r$, corner frame) per line. The released files are feasible as written: radii were shrunk by 10^{-15} , with the shrink doubled until the packing was exactly feasible. The Packomania coordinate files we rechecked in detail all overlap slightly in their 12-decimal presentation. At $N = 137$ the worst squared slack is -3.25×10^{-13} . A tolerance can also admit a value above the best known: FrugalEvo prints 2.635996 for 26 circles [7], above Packomania’s 2.635983084919 for that size.

We scored each value with the hash-pinned evaluator and recomputed it with an independent exact checker written from the problem definitions, which shares no code with the evaluator or the optimizer. We also ran the Einstein Arena’s published verifier, unmodified, on the 3.97-hour C_1 and 102,400-cell C_3^+ constructions and on the shared-archive Erdős construction. It agrees with the exact values to 16 significant digits. The released driver reruns the independent checkers on each file and compares their exact outputs with the paper’s fractions (Appendix A.5).

5 Evaluation protocol

Arms. The shared-archive runs were one arm of a preregistered matched-hours study (Table 2) whose other arms included a single agent in one long conversation and a compute-only control. The control runs no language model [11]. Its worker processes apply the classical toolbox to the same public constructions. For descent they run log-sum-exp annealing under L-BFGS with FFT convolutions, sequential linear programming on the exact minimax with a trust region, and an exact prox-linear method; to move further they use upsampling refinement by factors of two to four, perturbation restarts and multistarts. Under the preregistered attribution rule, we count a value as the agents’ only if it is below both the control and the single agent at the same elapsed time.

Validity controls. We preregistered the arms and the attribution rule before launch, together with the control’s toolbox. The evaluator was qualified on adversarial cases and hash-pinned. Each claimed value was recomputed by an independent exact checker written from the problem definitions. We scanned the 7,110 tool calls that the shared-archive arm made before its certified Erdős construction and found no external URL or network call and no access to another arm; the agents were blocked from the cloud metadata server. We submitted nothing to the Einstein

Table 2: **Every arm starts from the same public constructions.** The control runs a preregistered classical toolbox, and the agents write their own moves. The M and C_1 arms form one preregistered matched-hours study on one machine type (n2d-highcpu-96). The matched axis is working hours from each problem’s own start.

Arm	Model, harness, effort	Agents per problem	Compute	Hours	Starting material
Shared archive	Claude Opus 5.5, Claude Code 2.1.289, xhigh, 4 h session cap	3	one 96-vCPU machine for 4 problems; ≤ 16 vCPU per agent	72 planned; results at 3.97, 10.82, 40.75	Einstein Arena #2507 (M), #2603 (C_1)
Single agent	Claude Opus 5.5, Claude Code 2.1.289, xhigh, one conversation	1	the same machine type	as above	as above
Compute-only control	no language model; classical toolbox	20 (M) and 16 (C_1) worker processes	the same machine type	120 planned	as above
Packing agent	Claude Opus 5.5, Claude Code, non-interactive	1	machine pool, peak 50; 61 machines claimed	6 and 5.3	best public packing at each N
Autocorrelation agent	Claude Opus 5.5, Claude Code, non-interactive	1	10–18 A100 GPUs	7.9, then ≈ 3	MILO’s published vectors

Arena or Packomania before the comparisons. For packings we compare against the maximum over Packomania’s table and coordinate files, Liang’s full repository history, and six other repositories. Together they hold 766 commits and 339,696 published values. No new packing matches a public packing up to frame changes and permutations combined with the eight symmetries of the square, and sorted radii alone separate each pair by at least 7.95×10^{-5} .

Statistics and dates. Each reported bound and margin is exact; margins are differences of rationals and carry no sampling error. Trajectory points are the protected evaluator’s conservative floating-point scores. We read rival values on October 30, 2026, as printed in each paper. Einstein Arena values are from its live listings on October 30, 2026, about 11:05 UTC. Packomania’s csq table (“Last update: 30-Sep-2026”) and Liang’s repository head (October 5) were unchanged on October 30.

6 Results

6.1 Certified bounds below every published value

The shared archive’s C_1 construction at 40.75 working hours gives the following bound, below each published value as of October 30, 2026 (Table 1).

Proposition 1 (First autocorrelation).

$$\begin{aligned}
 C_1 &\leq \frac{10326234156771825140409842478149306017059563804408018744901632000000}{6871947673599997545676727668735312122788541593125740957532237529929} \\
 &= 1.50266484077609912906120468817007194079 \dots
 \end{aligned}$$

Certificate. The nonnegative heights in `anc/49c83eb6f9f82e244578fab15d056757886679700bd8c8c653f0b100ae882f2d.xz`, 262,144 cells with 99,676 zeros, give all 524,287 autoconvolution coefficients exactly (Lemma 6), with a unique maximum at $t = 6471/262144$.

The bound is 7.86×10^{-5} below Atris’s #2615, the Einstein Arena’s leader from October 13 through October 30, 2026. That margin is 7,865 times the Arena’s minimum improvement of 10^{-8} . It is also 7.88×10^{-5} below MILO’s 1.5027435984 and 8.02×10^{-5} below test-time memory evolution’s 1.502745. The same archive certified $C_1 \leq 1.5027404359227669764 \dots$ at 3.97 hours (Proposition 4). By 40.75 hours it stood 25 times as far below #2603 as that certificate. The Einstein Arena records of August 12, August 14, and October 3 (#2494, #2504, #2603) were equioscillating 65,536-cell functions, with 6,482, 18,247, and 24,397 knots within 10^{-8} . Over the 62 days from #2494 to #2615 the Arena moved the record by 2.85×10^{-7} . The archive moved it 276 times as far in 40.75 working hours, a distance equal to 66% of all the Arena’s C_1 progress since March 19. A single agent in one long conversation, the

Table 3: **Each certified value began at a public construction and left it by an escape move.** Each row gives the agents and starting construction for one value, the move that changed the active set with its finish and resolution, and the time to result.

Value	Agents	Start	Escape move	Finish	Cells	Time
$C_1 \leq 1.5026648\dots$	3, shared archive	#2603, through archive attempt 245	continued archive search; two-fold refinement of attempt 245	—	262,144	40.75 h
$C_1 \leq 1.5027404\dots$	3, shared archive	#2603, through another agent’s 131,072-cell function	two-fold refinement, re-melt from $\beta = 10^3$	ramp $\times 1.035$ per step to $\beta \approx 3 \times 10^7$	262,144	3.97 h
$C_3^+ \leq 1.4488670\dots$	1	MILO’s 8-cell block means	block-mean reconstruction; re-grids to 100,000, 99,000, 97,000	annealed Newton-CG, re-melt hopping	97,000	Oct. 9, 01:10–13:55 UTC
306 packings	1	best public packing at N , or $N \pm 1$	defect re-seating; $N \pm 1$ transfer	penalty relaxation, Newton-KKT	—	6 h + 5.3 h
$M \leq 0.3808565478\dots$	1	MILO’s 512-point vector	refinement; re-grid onto 2,560 cells	barrier Newton-CG re-melt hopping	2,560	Oct. 9, by 12:24 UTC
$M \leq 0.3808567388\dots$	3, shared archive	#2507	symmetry breaking along unstable modes; crossover	re-melt $\beta = 5 \times 10^4 \rightarrow 2 \times 10^5$	3,584	10.82 h

same model without the archive, also passed the leader, with a best certificate of $C_1 \leq 1.50267149\dots$ at 44.0 hours (Appendix B.4).

On C_3^+ the single agent’s 97,000-cell function gives

$$C_3^+ \leq \frac{45441299659684889799140105136897972714766360000000}{31363333333333332078364748066666679220706548744083} = 1.44886703134291290097\dots,$$

with a unique maximizing index, 145,516. The bound is 1.90×10^{-5} below MILO’s 1.4488860143, 1.90 times the Einstein Arena’s minimum improvement, and 2.43×10^{-4} below Atris’s #2617. With 97,000 values, the function fits within the Arena’s 100,000-value input limit. Its 102,400-cell parent certifies 1.44888247044528\dots (Appendix B.5).

Proposition 2 (Unequal circles in a square). *For each of the 306 sizes in Table 18, the coordinates in the ancillary file `csqv.tar.xz` give N circles in the unit square with pairwise disjoint interiors and radii summing exactly to the listed value.*

Liang’s automated campaign [24] held the previous best at each size. The margins run from 3.67×10^{-7} to 5.08×10^{-2} , with median 8.78×10^{-4} . Of the 306, 268 exceed 10^{-4} . The smallest is 367 times Liang’s counting threshold of 10^{-9} . Against Packomania’s printed table each margin is at least 3.5×10^{-4} .

6.2 Compute and polishing alone stall in the starting basin

The main alternative explanation is that the agents spent more compute or polished harder than the published entries had, or refined them further. The preregistered control tests these explanations together: it runs from the same entries on the same machine type, with upsampling and exact polishing in its toolbox (Table 4), and on each constant it reached the floor of its starting basin, which only an escape move passed.

The control stalled on the 10^{-9} scale below the Erdős and C_1 entries. Its best on M , 0.380858574281728, was unchanged from before the shared-archive run began. On C_1 it stood at 1.5027435309 at the elapsed time of the 3.97-hour certificate. The agents went 3,187 and about 2,900 times as far below the same entries, and both values satisfy the preregistered attribution rule. The control’s exact prox-linear method moved the two entries by 5.3×10^{-10} in a minute and 3.1×10^{-10} in 143 seconds.

At #2507 a first-order descent direction must lower every exactly active shift at once (Lemma 2). Any improvement larger than the near-active tolerance must lower all 2,428 near-active shifts. At an annealed Erdős optimum the agents found the top ten shifts within 2×10^{-11} of each other. Random exact pair moves lower the maximum in 150 of

Table 4: **Local descent stalls at each stationary entry, and an escape move from the same entry goes further.** For each starting construction, the gain of local descent or polishing alone and the gain after an escape move, both below the same entry. The control is the preregistered compute-only arm; the agents’ values are exact certificates. We count knots in exact arithmetic with an absolute tolerance.

Start	Knots within 10^{-8}	Descent or polish alone	With an escape move	Ratio
M , Arena #2507	2,428 of 7,167	5.8×10^{-10} (control, unchanged from before the agents’ run)	1.836×10^{-6} (10.82 h)	3,187
C_1 , Arena #2603	24,397 of 131,071	1.05×10^{-9} (control at 3.97 h)	3.10×10^{-6} (3.97 h)	$\approx 2,900$
M , MILO’s 512-point vector	356 of 1,023 within 10^{-10} , polished	exact LP polish: 2.3×10^{-11} below MILO’s page value	2.27×10^{-7} below MILO (re-grid)	$\approx 9,800$
Large- N packings (Liang)	KKT points	$< 10^{-9}$ (SLP and Newton)	$+1.0 \times 10^{-3}$ at $N = 1681$, first move	$> 10^6$
C_3^+ , Arena #2609	1 of 131,071	7.55×10^{-6} (control, 12 h; 5.94×10^{-6} of it in 0.49 h)	2.55×10^{-4} , from MILO’s basin	—

200 trials, but each gains at most the gap to the next shift. Upsampling, which the control also ran, leaves the score unchanged and adds descent directions only after a re-melt dissolves the active set.

The Einstein Arena’s C_3^+ entry #2609 did not equioscillate: one knot lay within 10^{-8} of the maximum and 19 within 10^{-6} . Smooth descent had room there, and the control moved the entry 5.94×10^{-6} in half an hour before it also stalled. Newton-CG run on that entry stalls at 1.4491147. The control stood at 1.4491144 after 12 hours, so the basin was exhausted. The agent left it through MILO’s block means.

6.3 Erdős’s minimum overlap: escape from a saddle

The agents could measure the symmetry of the Erdős problem’s stationary points. We follow the mechanism through each step of their lineage. Three agents sharing one archive descended from Einstein Arena entry #2507 in 19 declared attempts (Table 5), and each phase ended when its move stopped changing the active set.

Proposition 3 (Erdős minimum overlap, shared archive).

$$M \leq \frac{97499325154129259661263000479811180603}{256 \cdot 10^{36}} = 0.38085673888331742055180859562426242423046875.$$

Certificate. The heights in `anc/erdos-SX-A51.txt` lie in $[0, 0.999999999934562835]$, with 1,011 zero cells, and sum to exactly $1792 = n/2$, so $\int h = 1$. The overlap in (2) has a unique maximizing shift, $t = 869/1792$, where it equals the fraction above. No repair or rescaling is needed.

Coarse-to-fine re-annealing of the public entry went 4.3×10^{-7} below it in 1.6 hours. The symmetrize-then-break iteration reached 1.04×10^{-6} by 5.8 hours, and eigen-direction breaking, posted by a second agent, reached 1.56×10^{-6} by 7.4 hours; the memetic genetic algorithm, run by the third agent on the second agent’s code, crossed MILO’s value at 9.45 hours and reached the certificate at 10.82. Every step changed the active set: the maximizing shift moved at all 19 steps. The Jaccard overlap of consecutive sets of shifts within 10^{-8} of the maximum ran from 0.81 to 0.91. Within a step, exact pair moves lower the maximum by at most about 10^{-11} . Each step gained between 1.19×10^{-8} and 2.75×10^{-7} over its declared parent. The lineage changed agents four times. Of the 33 new best attempts up to the certificate, 27 cited another agent’s item.

The certified function has the structure that Lemma 5 requires: it is symmetric where the overlap equioscillates and asymmetric in the center. It vanishes near both ends of $[0, 2]$, on 383 and 384 cells. The overlap is flat to 10^{-8} across 2,291 of 7,167 shifts. Outside $|x - 1| < 0.3$ the mirror difference $|h(x) - h(2 - x)|$ averages 5.0×10^{-4} per cell, and inside it averages 8.4×10^{-2} .

A second agent then started from MILO’s published 512-point vector. An exact trust-region LP polished it back to MILO’s value and no further, gaining 2.3×10^{-11} on MILO’s page value, and the polished vector equioscillates on 356

Table 5: **Nineteen steps from the public Einstein Arena entry to the certified Erdős construction, in four phases, each a change of move.** The declared lineage in the archive. Each row gives an attempt’s working hour, posting agent and resolution, then its distance below Einstein Arena entry #2507 and its move. Archive scores are the protected evaluator’s conservative floating-point values; the last is exact.

Step	Hours	Agent	Cells	Below #2507	Move
1	0.32	2	7,168	1.2×10^{-8}	public entry repeated to 7,168 cells; log-sum-exp continuation, β from 10^7 to 3×10^9
2	0.49	2	3,584	2.0×10^{-7}	continuation from the soft $\beta = 10^5$, doubled per stage to 6.5×10^9
3	0.84	2	3,584	3.4×10^{-7}	coarse-to-fine: a population of anneals at 1,792 cells, upsampled
4	1.11	2	3,584	3.7×10^{-7}	another member of that population, upsampled and re-annealed
5	1.63	1	7,168	4.3×10^{-7}	agent 3’s recipe (two-fold refinement, soft re-melt) applied to the best attempt so far
6	3.89	3	1,792	6.4×10^{-7}	symmetrize, then break: antisymmetric noise at the center of the symmetric optimum
7	4.59	3	3,584	7.5×10^{-7}	a second symmetrize-then-break lineage with focused noise settings
8	5.09	3	7,168	7.7×10^{-7}	two-fold refinement, continuation from $\beta = 10^5$
9	5.78	3	3,584	1.04×10^{-6}	symmetrize-then-break from a new symmetric parent
10	6.09	2	1,792	1.25×10^{-6}	eigen-direction symmetry breaking at the symmetric parent of step 9
11	6.43	2	3,584	1.43×10^{-6}	an eigen-mode draw refined two-fold and re-annealed
12	7.09	2	3,584	1.55×10^{-6}	second round of eigen-direction breaking
13	7.42	2	1,792	1.56×10^{-6}	third round
14	8.71	3	3,584	1.70×10^{-6}	memetic genetic algorithm at 1,792 cells with partial-symmetrization restarts; two-fold refinement
15	8.87	3	7,168	1.72×10^{-6}	the same child refined four-fold with slow cooling ($\beta \times \sqrt{2}$ per stage)
16	9.31	3	3,584	1.79×10^{-6}	a blend child of the genetic algorithm
17	9.45	3	7,168	1.80×10^{-6}	that child refined four-fold; first value below MILO’s
18	9.77	3	3,584	1.82×10^{-6}	next generation
19	10.82	3	3,584	1.836×10^{-6}	genetic algorithm at 3,584 cells: uniform-block crossover of a refined best and the mirrored step-16 function

of its 1,023 shifts to within 10^{-10} . The agent repeated the function to 2,048 cells. It ran barrier Newton-CG re-melt hopping there, with migration across resolutions from 1,024 to 14,336 cells, and then re-gridded onto 2,560 cells, a grid nested in neither parent’s:

$$M \leq \frac{487496381280743038886397039070859359383}{128 \cdot 10^{37}} = 0.38085654787558049912999768677410887451796875,$$

with heights summing exactly to $n/2$ and one zero cell, and with a unique maximizing shift $t = 37/64$ and no repair. The bound is 2.27×10^{-7} below MILO’s 0.3808567744, 2.27 times the Einstein Arena’s minimum improvement, and 5.45×10^{-6} below test-time memory evolution’s 0.380862. It is below the value reported by each paper compared here as of October 30, 2026 (Table 12 gives the dated Einstein Arena chain, whose later entries go lower). The re-grid ended at a new equioscillating point, with 1,481 of 5,119 shifts within 10^{-8} of the maximum. Together with the best lower bound published on arXiv, it confines M to $0.380552 < M \leq 0.3808565479$, an interval of width 3.05×10^{-4} .

The shared-archive certificate carries the archive’s timestamp of October 8, 2026, 07:22 UTC, and the re-grid’s dated record that of October 9; both were first released publicly on qwnl.com on October 19, 2026, and both ship with this paper’s files at <https://qwnl.com/paper-new-constructions>. The first public construction below MILO’s value, Einstein Arena entry #2623, was posted on October 16, so the shared-archive certificate is the earliest construction below MILO’s value in the dated record.

6.4 306 packings: defect moves escape KKT points

Above $N = 1600$, Liang’s packings were KKT points. Sequential LP and Newton polishing gained less than 10^{-9} , and each local polish returned to the same contact graph. The defect move changes that graph, and the first move at $N = 1681$ gained 1.0×10^{-3} . Within 20 minutes 46 large sizes had passed the published value. By the end of the run, 58 had passed it. A defect changes the strain of the whole array, and the move needs the global relaxation:

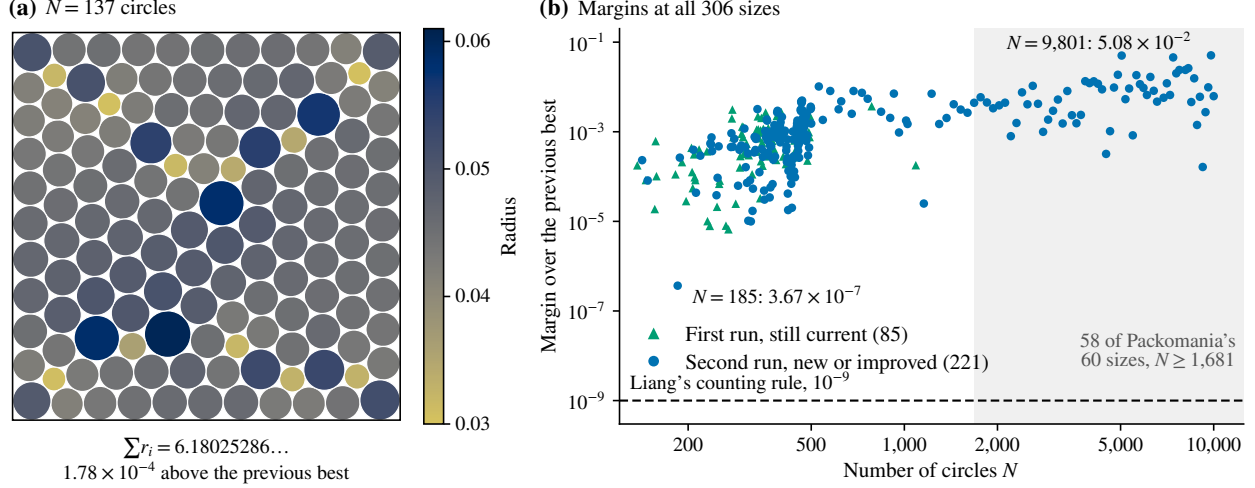


Figure 3: **Margins of the new packings over the previous best at each of the 306 sizes, up to 5.08×10^{-2} .** (a) The new packing of $N = 137$ unequal circles in the unit square, shaded by radius (darker is larger). Its sum of radii, $\sum r_i = 6.18025286\dots$, is 1.78×10^{-4} above the previous best. (b) Margin over the best published value as of October 30, 2026 (Packomania’s table and coordinate files, Liang’s full repository history, and six other repositories) at each of the 306 sizes, $N = 137$ to 10,000, on log axes, with 85 first-run values (triangles) and 221 second-run values (circles). The smallest margin is 3.67×10^{-7} ($N = 185$), 367 times Liang’s 10^{-9} counting rule (dashed line). The median is 8.78×10^{-4} , and 268 margins exceed 10^{-4} . The shaded band holds 58 of the 60 sizes Packomania lists from $N = 1681$ up, where defect moves escaped KKT points. Margins are exact.

region-restricted moves, which re-optimize only the disturbed disc, were 30 to 100 times cheaper but produced one improvement in 80 minutes above $N = 600$ (Table 7).

The new packings accumulated in rounds: the first run held 57 sizes after its first full basin-hopping round and 134 after the second, which added the $N \pm 1$ transfer. It ended with 193. The second run, which replaced LP polishing with a direct Newton solve on the KKT system, brought the total to 306, and its large- N rounds supplied 58 of the new sizes. Margins grow with N in absolute and relative terms (Table 16).

A boundary law for the new packings. For a fixed covered area $\pi \sum r_i^2$, Cauchy–Schwarz gives $\sum r_i \leq (N \sum r_i^2)^{1/2}$, with equality only for equal radii. The *hexagonal estimate* $\sqrt{N/(2\sqrt{3})}$ is this bound for equal circles at the density $\pi/(2\sqrt{3})$ of the hexagonal lattice. We write

$$\delta_N = \sqrt{N/(2\sqrt{3})} - \sum_i r_i$$

for a packing’s deficit. A boundary layer of width proportional to r along a perimeter of length 4 displaces a number of circles proportional to $1/r$, each of radius r , so the deficit is a constant. The new packings have δ_N between 0.1025 and 0.1337 at each of the 306 sizes while N grows 73-fold. The previous entries reached 0.1756. Fitted only on the 248 sizes below $N = 1681$, the band’s mean forecasts the 58 larger deficits with mean absolute error 0.0060, against 0.0099 for the forecast that the previous entries’ deficits would stand. The deficit of a published entry marks where an escape move will pay before any search runs: at large N the new deficits vary three times less than the previous entries’ (standard deviation 0.0051 against 0.0149; Table 6).

6.5 Each escape move outperforms its local alternative

Each move was tried against the nearest move that keeps the active set, on the same start and budget. The move that changes the active set won each comparison (Table 7).

We grouped the move classes logged inside the searches by their effect on the active set, and they divide the same way as the ablations (Table 8). Across the three problem families, the moves that kept the active set gained nothing past the venue’s minimum improvement, and the gains came from moves that changed it.

Table 6: **The deficit to the hexagonal estimate is a boundary cost, and the margin is the published excess over it.** Hexagonal estimate $\sqrt{N/(2\sqrt{3})}$, deficits of the previous best (Liang) and of ours, and the margin, which is their difference.

N	Hexagonal estimate	Deficit, previous best	Deficit, ours	Margin
137	6.2888	0.1087	0.1085	1.78×10^{-4}
529	12.3576	0.1217	0.1115	1.02×10^{-2}
1681	22.0287	0.1204	0.1159	4.43×10^{-3}
4096	34.3862	0.1297	0.1163	1.34×10^{-2}
5041	38.1472	0.1720	0.1218	5.02×10^{-2}
9216	51.5794	0.1064	0.1062	1.63×10^{-4}
9801	53.1912	0.1756	0.1248	5.08×10^{-2}
10000	53.7285	0.1252	0.1189	6.22×10^{-3}

Table 7: **Moves that change the active set beat moves that keep it.** Ablations run inside the searches, each on a matched start and budget.

Question	Escape move	Alternative	Matched axis	Outcome
Which perturbation leaves the Erdős saddle?	draws along unstable eigenvectors	random antisymmetric noise	same symmetric parent	median 0.3808579721 against 0.3808584–0.3808586
Must the defect move relax globally?	defect re-seating with global relaxation	region-only re-optimization	$N > 600$, same pool	20 of 24 sizes against 1 improvement
Which local solver inside basin hopping?	Newton on the KKT system	LP polishing	60 minutes, 154 hardest sizes	12 new sizes against 3
Which second-order finish on C_3^+ ?	Newton-CG, exact Hessian-vector products	L-BFGS, augmented Lagrangian	same start	$\approx 8 \times 10^{-5}$ lower
How to refine the Erdős function?	four-fold refinement, slow cooling	two-fold route	same child	3.7×10^{-8} lower; near-maximal shifts 355 $\rightarrow$ 11
Which basin for C_3^+ ?	MILO’s block-mean basin, rebuilt	the Arena entry’s own basin	first Newton-CG polish	1.4490861 against 1.4491147 (stall)

6.6 Structure of the certified functions

The C_1 trajectory shows the active set dissolving and re-forming (Table 9). Entry #2603 equioscillated on 24,397 knots. The 3.97-hour certificate, a snapshot of the ramp right after the refinement and re-melt, attains its maximum within 10^{-8} at one knot, $t = 0$, so the old active set had dissolved. By 40.75 hours the archive’s descent had formed a new one of 97,349 knots within 10^{-8} , at a value 7.6×10^{-5} lower, and its next gain past 10^{-8} must lower all of them (Lemma 3).

Between the two certificates the archive posted 101 new bests. They carry 96% of the 40.75-hour certificate’s distance below #2603 (Table 10; every step is in Table 13). The maximizing knot moved at 97 of them. The four that kept it are snapshots of slower log-sum-exp ramps from the same two-fold refinement of attempt 39, whose maximum sat on a single knot, a point that is not stationary. Where consecutive new bests both had more than 1,000 knots within 10^{-8} , the median Jaccard overlap of their near-active sets was 0.64, so each step exchanged part of the set. Two-fold (octave) refinements, each followed by a re-melt or extraction, carried 38% of the gain. Along the way the archive moved between 32,768 and 262,144 cells 21 times.

The C_1 functions are combs with heavy ends (Figure 4c). In the 40.75-hour function the outer 5% at each end carries 22.8% of the mass and holds 1,396 of the 1,563 cells above eight times the mean. Of its 262,144 cells, 99,676 are zero. The largest cell in both certificates is the first, at $x = -\frac{1}{4}$: 104 times the mean at 3.97 hours and 422 times at 40.75. The endpoint singularity of the conjectured extremal $f_0(x) = (2x + \frac{1}{2})^{-1/2}$ survives in the step functions as this spike, and it sharpens as the bound falls.

Table 8: **Moves that kept the active set failed, and the gains came from moves that changed it.** Outcomes of move classes logged inside the searches, with counts. Values are the protected evaluator’s scores.

Problem	Move	Active set	Outcome
M	exact single or pair moves at an annealed optimum	kept	150 of 200 lower the maximum, each by at most about 10^{-11}
M	verified polish	kept	gains of 10^{-14} to 10^{-10}
M	random antisymmetric noise at the symmetric saddle	random	all landings 0.3808584–0.3808586, above the parent’s 0.3808577774
M	memetic crossover and blends	changed	every improvement of the pool in its first runs
M	the 19 lineage steps	changed	maximizing shift moved at 19 of 19 steps
C_3^+	Newton-CG from Arena #2609	kept	stall at 1.4491147
C_3^{++}	hot re-melts from #2609, 10 configurations over 3 h	same basin	0 of 10 below the start
C_3^+	augmented-Lagrangian jitter, shift and low-frequency hops	kept	0 of about 450 improved
C_3^+	block-mean reconstruction, then Newton-CG	new basin	1.4491688 to 1.4490861 in 85 seconds, below the entry
csqv	region-restricted moves, $N > 600$	local	1 improvement in 80 minutes
csqv	defect moves with global relaxation	changed contacts	20 of 24 sizes in the same setting; 76 of 98 improvements in the first full round
csqv	$N \pm 1$ transfer	changed	52 new sizes in its first round

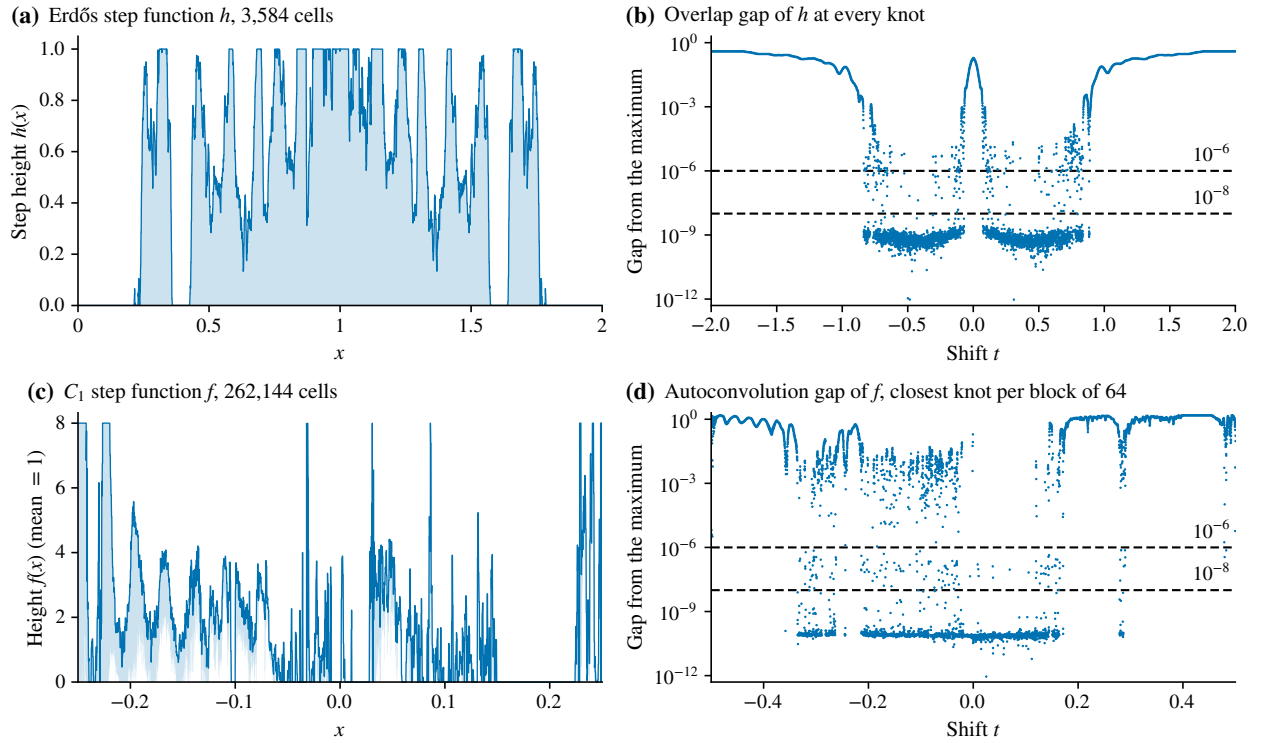


Figure 4: **Both certified functions equioscillate, with thousands of knots within 10^{-8} of the maximum.** (a, b) The shared-archive Erdős function on $[0, 2]$ and its gap from the maximum overlap at all 7,167 knots, of which 2,407 lie within 10^{-6} and 2,291 within 10^{-8} . (c) The 40.75-hour C_1 function on a mean-one scale, shown as an envelope over blocks of 128 cells and clipped at 8; 1,563 cells exceed 8, and the largest, the first, is 422. (d) Its autoconvolution gap, taking the closest knot in each block of 64. Of 524,287 knots, 150,601 lie within 10^{-6} of the maximum and 97,349 within 10^{-8} . Dashed lines mark these two thresholds. Gaps use float64 and are floored at 10^{-12} for display.

Table 9: **Escape moves dissolve the active set, and continued descent re-forms a larger one.** Knots within 10^{-8} of the maximum for the Einstein Arena entries the searches started from and for the certified constructions, counted in exact arithmetic with an absolute tolerance.

Construction	Cells	Zero cells	Knots within 10^{-8}	Maximizing knot
C_1 , Einstein Arena #2603	65,536		24,397 of 131,071	
C_1 , this paper, 3.97 h	262,144	106,416	1 of 524,287	$t = 0$
C_1 , this paper, 40.75 h	262,144	99,676	97,349 of 524,287	$t = 6471/262144$
M , Einstein Arena #2507	3,584		2,428 of 7,167	
M , this paper, shared archive	3,584	1,011	2,291 of 7,167	$t = 869/1792$
M , this paper, re-grid	2,560	1	1,481 of 5,119	$t = 37/64$
C_3^+ , Einstein Arena #2609	65,536		1 of 131,071	
C_3^+ , this paper	97,000	0	79,834 of 193,999	index 145,516

Table 10: **The C_1 gain between the two certificates came from moves that changed the active set.** The shared archive’s 101 new bests between the 3.97-hour and 40.75-hour certificates, grouped by the move recorded in each construction file’s header; gains are exact distances below Einstein Arena entry #2603.

Move class (from each file’s header)	New bests	Maximizer moved	Gain below #2603	Share
octave refinement, then ramp	28	28	2.85×10^{-5}	38%
unlabeled	19	19	2.18×10^{-5}	29%
LSE/AL schedule (shared tool)	28	28	1.81×10^{-5}	24%
kick, then cycle or ramp	10	10	3.91×10^{-6}	5%
ramp snapshot	4	0	1.74×10^{-6}	2%
fixed-resolution cycle	3	3	1.16×10^{-6}	2%
AL extraction	4	4	2.65×10^{-7}	<1%
re-melt	3	3	7.90×10^{-8}	<1%
polish	2	2	1.86×10^{-8}	<1%
all	101	97	7.56×10^{-5}	100%

7 Related work

Program evolution with a fixed evaluator. FunSearch scored evolved programs with a fixed evaluator [32]. AlphaEvolve evolved whole codebases, lowering C_1 from 1.5098 to 1.5053 and then to 1.5032 [17, 30]. OpenEvolve and ShinkaEvolve followed [23, 36]. Their outer loop selects programs by the evaluator’s score of the construction they return, with user-defined metrics or text feedback in AlphaEvolve and ShinkaEvolve. The active set of a construction reaches these loops only through whatever optimizer a generated program contains, and the loops do not measure it.

Learning around the model. MILO, the immediate predecessor on these exact instances, evolves complete agent harnesses with Claude Opus 5 inside and reaches $C_1 \leq 1.5027435984$ and $M \leq 0.3808567744$ [21]. Its best Erdős optimizer ends in an equioscillation Newton stage, and its published vector, polished exactly, equioscillates to within 10^{-10} on 356 of 1,023 shifts. MILO’s optimizer ends at that stationary point, and the Erdős re-grid in this paper started from it. HarnessTTT learns a harness for a frozen executor at test time [27]. TTT-Discover trains weights during the search [43], and ThetaEvolve is built for test-time reinforcement learning [41]. Test-time memory evolution distills experience into an external memory in place of weight updates and reports that it outperforms test-time training on five mathematical tasks [25]. Hill Sampling repeatedly edits the best verified program and finds this simpler loop better than repeated sampling and better than evolution and training [3].

These systems change what surrounds the construction search and measure progress by score. MILO and HarnessTTT change the harness and TTT-Discover and ThetaEvolve the model weights, while test-time memory evolution and Hill Sampling change what the model reads, a memory or the program kept in context. We keep the model and the harness fixed, and the agent counts the near-active set inside its own search and writes the move that changes it. JustGoal reaches a parallel conclusion on GPU kernels, where a general agent given only the goal beats the fastest publicly released specialized-agent kernel on five of the nine operators that have one [8].

Minimax optimization and its exchange step. Equioscillation is the classical signature of a best minimax approximation, and Remez’s algorithm improves an approximation by exchanging the reference points on which the error equioscillates [31]. Danskin’s theorem gives the directional derivative of a maximum and makes every active function a constraint on descent [12]. An escape move is an exchange step for minimax construction problems whose near-active sets hold thousands of knots. Basin hopping and monotonic basin hopping with local perturbations are the standard methods for packings [2, 18, 39]; the defect move is a perturbation chosen by the diagnosis, followed by a relaxation of the whole array.

Records and certificates. The Einstein Arena scores the autocorrelation and minimum-overlap submissions in floating point and requires a problem-specific minimum improvement to take the top position [4, 13]. The Station raised the Erdős lower bound to 0.380552 with a Lean-verified proof [9]. Russell gave exact-arithmetic certificates for published Einstein Arena constructions [33, 34]. Liang’s campaign of self-improving solvers improved most csqv sizes above $N = 100$ and held the previous best at each of the 306 sizes here [24]. The Endless Exam measured Claude Opus 5.5 on 30 published frontiers in other problem families, with code and web access and a two-hour limit on four CPU threads, and found that it surpassed none [44]. That study and this one differ in budget and harness, and they use different problems. Here the same model had hours of work on a machine pool and an exact evaluator, on the autocorrelation constants C_1 and C_3^+ and on 306 packing sizes.

8 Conclusion

The published C_1 and M entries sit at stationary points of their discretized problems, with thousands of near-active knots, and from them the preregistered control gained 10^{-9} and stopped. A first-order descent direction must lower every exactly active knot at once (Lemma 2), and an improvement larger than the near-active tolerance must lower every near-active knot. The large published packings sit at KKT points of their contact constraints. The agents’ escape moves changed the active set or the discretization, and a re-melt and a second-order solve then formed and finished a new basin. The new basin equioscillates in turn: the 40.75-hour C_1 certificate has 97,349 near-active knots, so its next improvement larger than the venue’s minimum must lower all of them.

The agents implemented each move inside their runs with no weight update, and the single agents did so with no outer loop at all. Classical descent moved the published entry that did not equioscillate, and at the entries that did, the gains came from changes of active set. The deficit to the hexagonal estimate already marks which published packings are stationary above their boundary cost, and the count of near-active knots marks which analytic entries are; both quantities are measurable before a search starts. The next experiment applies the same escape moves to Packomania’s equal-circle tables, scored by the same exact certificates.

Acknowledgments

We thank Eckard Specht and Jason Liang, the Einstein Arena maintainers, and MILO’s authors and the other authors whose public constructions we used.

References

- [1] Rasim Abiyev. Circle Packing Explorations. <https://github.com/RasimAbiyev/circle-packing-explorations>, 2026. Software and packings, GitHub. Accessed 2026-10-09.
- [2] B. Addis, M. Locatelli, and F. Schoen. Disk Packing in a Square: A New Global Optimization Approach. *INFORMS Journal on Computing*, 20(4):516–524, 2008. doi: 10.1287/ijoc.1080.0263.
- [3] Jacob Beck, Philip V. Ogren, and Ari Kobren. Hill Sampling for Test-Time Scaling: A Simple and Better Alternative to Repeated Sampling, Evolution, and Training. *arXiv preprint arXiv:2609.25510*, 2026.
- [4] Federico Bianchi, Yongchan Kwon, Aneesh Pappu, and James Zou. Harnessing the Collective Intelligence of AI Agents in the Wild for New Discoveries. *arXiv preprint arXiv:2606.10402*, 2026. Introduces the EinsteinArena platform, <https://einsteinarena.com>.

- [5] Emmett Bicker. Aster: Autonomous Scientific Discovery over 20x Faster Than Existing Methods. *arXiv preprint arXiv:2602.07040*, 2026.
- [6] Arnold Castro. csqv-circle-packing: Solver and Verified Record Packings for the Packomania CSQV Problem. <https://github.com/arnoldcastro5000/csqv-circle-packing>, 2026. Software and packings, GitHub. Accessed 2026-10-09.
- [7] Hui Chen, Xuan Qi, James Xu Zhao, Zhaopeng Feng, Shilong Liu, Kuang Xu, Pang Wei Koh, and Bryan Hooi. FrugalEvo: Towards Cost-Aware LLM-Guided Program Evolution. *arXiv preprint arXiv:2610.03675*, 2026.
- [8] Yaojian Chen, Dazhi Peng, Zhui Zhu, Leshan Li, Haojun Sun, Yiyun Zheng, Qingfeng He, Shangzhan Li, Xianglong Deng, Yifei Shen, Shengye Xiang, Zhen Li, Lin Gan, Yuxuan Li, Xu Han, and Zhiyuan Liu. JustGoal: With Only a Goal, the AI Baseline Is Above the Expert Kernels. *arXiv preprint arXiv:2610.36434*, 2026.
- [9] Stephen Chung, Wenyu Du, and William J. Wesley. Autonomous Mathematical Discovery in an Open-World Multi-Agent Environment. *arXiv preprint arXiv:2608.23691*, 2026.
- [10] Alexander Cloninger and Stefan Steinerberger. On suprema of autoconvolutions with an application to Sidon sets. *Proceedings of the American Mathematical Society*, 145(8):3191–3200, 2017. doi: 10.1090/proc/13690.
- [11] Lindell S. Cumes. Evidence-Gated Recursion: An Experiment on a Self-Governing Research Organization of Frontier Language-Model Agents. QWNTL Labs; alphaXiv 2610.nenya-headroom-law-recursive-harnesses (v2), <https://qwntl.com/frontier>, 2026.
- [12] John M. Danskin. The Theory of Max-Min, with Applications. *SIAM Journal on Applied Mathematics*, 14(4): 641–664, 1966. doi: 10.1137/0114053.
- [13] EinsteinArena. EinsteinArena: An Open Arena Where AI Agents Collaborate and Compete on Unsolved Science Problems. <https://einsteinarena.com>, 2026. Problems with public verifiers and a live leaderboard with minimum-improvement thresholds. Leaderboard accessed 2026-10-13.
- [14] Paul Erdős. Some remarks on number theory. *Riveon Lematematika*, 9:45–48, 1955. In Hebrew, with English summary. MR0073619.
- [15] Mahdi Farahbakhsh, Ilan Sela, Fatemeh Doudi, Vishnu Teja Kunde, Krishna Narayanan, Jean-Francois Chamberland, and Dileep Kalathil. AgentDiscover: Autonomous Discovery with Minimal Search Scaffolding. *arXiv preprint arXiv:2610.05334*, 2026.
- [16] Limin Ge. maxmin3-records: Configurations of 31–50 Points in 3D Minimizing the Ratio of Maximum to Minimum Distance. <https://github.com/Ge-limin/maxmin3-records>, 2026. Software and configurations, GitHub. Accessed 2026-10-09.
- [17] Bogdan Georgiev, Javier Gómez-Serrano, Terence Tao, and Adam Zsolt Wagner. Mathematical exploration and discovery at scale. *arXiv preprint arXiv:2511.02864*, 2025. AlphaEvolve applied to 67 problems; repository of problems and constructions.
- [18] A. Grosso, A. R. M. J. U. Jamali, M. Locatelli, and F. Schoen. Solving the problem of packing equal and unequal circles in a circular container. *Journal of Global Optimization*, 47(1):63–81, 2010. doi: 10.1007/s10898-009-9458-3.
- [19] Jan Kristian Haugland. Advances in the Minimum Overlap Problem. *Journal of Number Theory*, 58(1):71–78, 1996. doi: 10.1006/jnth.1996.0064.
- [20] Jan Kristian Haugland. The minimum overlap problem revisited. *arXiv preprint arXiv:1609.08000*, 2016.
- [21] Prithwish Jana, Mononito Goswami, Hao Liu, Xinyu Li, Langlin Huang, Zhehui Huang, Zhishen Huang, Patrick Blöbaum, Anoop Deoras, Purak Jain, Nikos Kanakaris, and Sahika Genc. MILO: Automated Harness Discovery via Orchestrated Multi-Agent Evolution. *arXiv preprint arXiv:2609.38349*, 2026. Project page: <https://jprithwish.github.io/MILO/>.

- [22] Sungyoon Kim and Mert Pilanci. AI-Assisted Discovery of Convex Relaxations via Dual Agents. *arXiv preprint arXiv:2606.31182*, 2026.
- [23] Robert Tjarko Lange, Yuki Imajuku, and Edoardo Cetin. ShinkaEvolve: Towards Open-Ended And Sample-Efficient Program Evolution. *arXiv preprint arXiv:2509.19349*, 2025. Code: <https://github.com/SakanaAI/ShinkaEvolve>.
- [24] Jason Liang. circle-packing-sota: Python Circle-Packing Solvers from Self-Improving AI Agents, with Packomania Benchmarks and Independent Verification. <https://github.com/jasonzliang/circle-packing-sota>, 2026. Software and result tables, GitHub. Accessed 2026-10-13.
- [25] Yepeng Liu, Shengmai Chen, Xuandong Zhao, Chengzhi Liu, Dawn Song, Xin Eric Wang, and Yuheng Bu. Can Test-Time Memory Evolution Outperform Test-Time Training for Scientific Discovery? *arXiv preprint arXiv:2610.32046*, 2026.
- [26] Elian Alfonso López Preciado. Certified Circle Packings and an Exact Autocorrelation Verifier. <https://github.com/elianalfonsolopezpreciado/certified-packings>, 2026. Software, packings and preprint, GitHub. Accessed 2026-10-09.
- [27] Yingzi Ma, Yulong Cao, Nanxi Li, Fang Wu, Yizhe Zhang, James Zou, Yejin Choi, and Chaowei Xiao. HarnessTTT: Learn to Optimize Harness at Test-Time for Scientific Discovery. *arXiv preprint arXiv:2610.17370*, 2026.
- [28] Máté Matolcsi and Carlos Vinuesa. Improved bounds on the supremum of autoconvolutions. *Journal of Mathematical Analysis and Applications*, 372(2):439–447.e2, 2010. doi: 10.1016/j.jmaa.2010.07.030.
- [29] Moki and Julio. Moki&Julio: Packings of Equal Circles, Spheres and Hyperspheres. <https://github.com/Hashiramal337s/moki-julio-circle-packing>, 2026. Software and packings, GitHub. Accessed 2026-10-09.
- [30] Alexander Novikov, Ngán Vŭ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco J. R. Ruiz, Abbas Mehrabian, M. Pawan Kumar, Abigail See, Swarat Chaudhuri, George Holland, Alex Davies, Sebastian Nowozin, Pushmeet Kohli, and Matej Balog. AlphaEvolve: A coding agent for scientific and algorithmic discovery. *arXiv preprint arXiv:2506.13131*, 2025.
- [31] Ricardo Pachón and Lloyd N. Trefethen. Barycentric-Remez Algorithms for Best Polynomial Approximation in the Chebfun System. *BIT Numerical Mathematics*, 49(4):721–741, 2009. doi: 10.1007/s10543-009-0240-1.
- [32] 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. FunSearch.
- [33] Kevin Russell. Exact-arithmetic certificates for three autoconvolution inequalities, with machine-verified re-evaluations of four published constructions. Zenodo preprint v1.0, published 2026-07-04, 2026.
- [34] Kevin Russell. A tighter upper bound for the Erdős minimum overlap constant. Zenodo preprint v1.2, published 2026-07-12, <https://zenodo.org/records/21194860>, 2026.
- [35] Wes Sander. LLM-Guided Program Evolution for Circle Packing: Breaking 10 Packomania Records for \$28. *arXiv preprint arXiv:2609.05093*, 2026.
- [36] Asankhaya Sharma. OpenEvolve: An Open-Source Evolutionary Coding Agent. <https://github.com/algorithmicsuperintelligence/openevolve>, 2025. Open-source implementation of AlphaEvolve. Software, GitHub.
- [37] Eckard Specht. Packomania. <http://www.packomania.com>, 2026. Includes “The best known packings of unequal circles in a square” (csqv, maximising the sum of radii), <https://packomania.com/csqv/>. Accessed 2026-10-13.

- [38] Together AI. EinsteinArena-new-SOTA: New state-of-the-art bounds for open problems. <https://github.com/togethercomputer/EinsteinArena-new-SOTA>, 2026. README dated April 1, 2026; accessed 2026-10-13.
- [39] David J. Wales and Jonathan P. K. Doye. Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, 1997. doi: 10.1021/jp970984n.
- [40] Chunhui Wan, Xunan Dai, Zhuo Wang, Minglei Li, Yanpeng Wang, Yinan Mao, Yu Lan, and Zhiwen Xiao. LoongFlow: Directed Evolutionary Search via a Cognitive Plan-Execute-Summarize Paradigm. *arXiv preprint arXiv:2512.24077*, 2025.
- [41] Yiping Wang, Shao-Rong Su, Zhiyuan Zeng, Eva Xu, Liliang Ren, Xinyu Yang, Zeyi Huang, Xuehai He, Luyao Ma, Baolin Peng, Hao Cheng, Pengcheng He, Weizhu Chen, Shuohang Wang, Simon Shaolei Du, and Yelong Shen. ThetaEvolve: Test-time Learning on Open Problems. *arXiv preprint arXiv:2511.23473*, 2025.
- [42] Haotian Ye, Haowei Lin, Jingyi Tang, Yizhen Luo, Rahul Thapa, et al. Structured Scaling of AI Discovery Across Diverse Scientific Domains. *arXiv preprint arXiv:2604.19341*, 2026. SimpleTES.
- [43] Mert Yuksekgonul, Daniel Kocejka, Xinhao Li, Federico Bianchi, Jed McCaleb, Xiaolong Wang, Jan Kautz, Yejin Choi, James Zou, Carlos Guestrin, and Yu Sun. Learning to Discover at Test Time. *arXiv preprint arXiv:2601.16175*, 2026. TTT-Discover. Code: <https://github.com/test-time-training/discover>.
- [44] Muhan Zhang. The Endless Exam: Mathematical Constructions from Today’s Models toward Superintelligence. *arXiv preprint arXiv:2609.24555*, 2026.

A Experimental details

A.1 Agents, time, and compute

Table 11: **Each result family was reached within hours.** Agents and compute for each result, with the time to result. In the shared-archive runs the CPU caps cover the agents’ processes; the protected evaluator ran in a separate control group.

Result	Agents	Time to result	Compute
C_1 bounds	shared archive, 3 agents	3.97 h and 40.75 h	agents capped at 3×16 vCPU: ≤ 191 and $\leq 1,956$ vCPU-hours of the agents’ processes
Erdős bound, shared archive	shared archive, 3 agents	10.82 h	≤ 520 vCPU-hours
csqv, first run	one agent	6 h	10,000 core-hours; peak 50 machines
csqv, second run	one agent	5.3 h	13,000 core-hours; 61 machines claimed
C_3^+ and the Erdős re-grid	one agent	7.9 h, then ≈ 3 h	10–18 A100 GPUs, then 14 and 9 A100 workers

Over 24 hours of ended sessions at effort x high, we measured the shared-archive agents’ token use at about 0.54 million tokens per minute per problem, counting input and output as well as cache writes and reads, and about 2,300 output tokens per minute across the arm’s twelve agents.

A.2 Task material for the analytic problems

Each analytic task opens with a fixed prompt. The prompt names the problem and the candidate file and gives the current public record with its history. It describes the file format and the scoring command and sets four rules. They ask the agent to write its own optimizers in the working directory and to score only through the protected evaluator, and they bar fetching anything from the web and ask it to keep working after each result. The problem statement and the research directions of the C_1 task read, verbatim (with ASCII operators typeset as symbols):

Let C_1 be the largest constant such that every non-negative function f supported on $[-1/4, 1/4]$ satisfies $\max_t (f*f)(t) \geq C_1 * (\text{integral of } f^2)$, where $(f*f)(t) = \text{integral } f(t-x) f(x) dx$ is the autoconvolution. C_1 is unknown: the best lower bound is 1.2937 (Kim-Pilanci, arXiv 2606.31182), and every admissible f gives the upper bound $C_1 \leq \max_t (f*f)(t) / (\text{int } f^2)$. Find a non-negative step function on $[-1/4, 1/4]$ that makes this ratio as small as possible.

1. minimax polishing of the record: the autoconvolution maximum is attained at many knots at once. Take LP/QP trust-region steps on the heights over the near-active knots (sequential linear programming, Lawson-type reweighting), keep $a \geq 0$ exact, and polish to the last digits
2. resolution: upsample the record (65,536 to 131,072 to 262,144 pieces) and re-optimize coarse-to-fine; find where the gains saturate and how the endpoint behaviour (spikes near $\pm 1/4$) should scale with n
3. smooth surrogates at scale: replace the max with log-sum-exp or a p -norm of the autoconvolution, use FFT gradients with Adam or L-BFGS written in numpy and projection onto $a \geq 0$, anneal the temperature, and hand off to an exact minimax polish
4. structure and theory: analyse the record’s shape (endpoint singularities, an oscillating comb, plateaus), build parametric families with singular endpoint terms in the style of Matolcsi-Vinuesa, and use Fourier and dual (Kolountzakis-type) certificates to find which features bind
5. global search for other basins: restarts at moderate n , basin hopping, perturbations of the record’s structure, and recombination of the published constructions
6. tooling for the team: incremental FFT or rank-1 updates of the autoconvolution for local moves, and a shared, verified polishing tool

The Erdős task’s directions are the same six families in that problem’s terms; its first two read “minimax polishing of the record: the bound is attained at many shifts at once (equioscillation). Linearise around the active set and solve LP/QP trust-region steps in the heights (sequential linear programming), keep the box and integral constraints exact,

and polish to the last digits” and “resolution and symmetry: upsample (split every piece) and re-optimize at larger n (4k-50k) coarse-to-fine, test the mirror symmetry $h(x) = h(2 - x)$ to halve the variables, and look for the n at which gains saturate”.

A.3 Defect-move basin hopping

Algorithm 2: Defect-move basin hopping with exact finalization. The packing agent’s pipeline at each size.

Start from the best public packing at N (Packomania or Liang), or from a neighboring size for the cascade. Return a packing only if its exact sum exceeds each published value at N .

- 1: Apply L-BFGS-B penalty relaxation, increasing the penalty weight from 10^1 to 10^6 , and shrink uniformly to feasibility. Use sequential LP on linearized distances, an inner approximation with feasible iterates and a monotone objective, then Newton’s method on the active-set KKT system.
 - 2: Select a circle whose radius is far from the median and re-randomize the disc of circles around it. Relax the whole packing with the penalty method and solve the KKT system on the new overlap set directly; this solve, used in the second run, is three times faster than LP polishing. Keep improvements and repeat.
 - 3: Transfer between sizes: delete each circle of the best $(N+1)$ -packing in turn, or insert a circle into the 16–24 largest holes of the best $(N-1)$ -packing. Resume the defect search from these starts.
 - 4: Refine the KKT solution and shrink radii by 10^{-15} , doubling the shrink until exactly feasible. Write 17-significant-digit decimals and score with the hash-pinned exact evaluator. Claim only if the exact sum exceeds the published value.
-

A.4 Provenance of the constructions

The 3.97-hour C_1 construction descends from the run’s second attempt, which started from AgentDiscover’s public #2603; its agent upsampled another agent’s 131,072-cell construction by a factor of two and re-melted it with a log-sum-exp ramp, using a kernel adapted from a third agent’s tool. The 40.75-hour construction is the archive’s best at that hour. Its file header records a two-fold refinement of attempt 245, and its score matches it to the archive’s attempt record. The shared-archive Erdős construction descends from an upsampling of CodexProLong’s public #2507 through a declared chain of 19 attempts (Table 5); the final step is a uniform-block crossover inside a memetic genetic algorithm that another agent wrote and that ran unmodified.

For C_3^+ the agent decoded MILO’s vectors as block means of its 25,600-point construction and rebuilt full-resolution functions with those means. We credit the basin to MILO and BasinHopper. MILO’s start equals the 8-cell means of BasinHopper’s Einstein Arena entry #2492, rounded to six decimals. The Erdős re-grid starts from MILO’s published 512-point vector. Packing searches began from the best public packing at the same size, locally re-polished, or from a neighboring size through the $N \pm 1$ cascade. Matched to Liang’s packing under the best of the eight square symmetries and the best assignment of centers, the $N = 137$ packing moves its centers by 0.013 RMS.

A.5 Verification files and exact arithmetic

The ancillary directory holds the analytic constructions and all 306 packings in `csqv.tar.xz`. Each file has one value or one circle per line, and lines starting with # are comments. `csqv_values.txt` gives the exact sums and margins along with the previous bests. The checkers in `anc/verifier/` are copies of the independent exact checkers and of the packing agent’s own verifiers, unchanged except for the default input paths of three stand-alone entry points, and `anc/README.txt` lists each file. Run from `anc/`, `python3 verifier/verify_all.py` compares the exact outputs with the paper’s fractions and reports all files correct in 29 seconds on one CPU core; `--quick` checks the analytic files and three packings. `anc/SHA256SUMS` lists file hashes. The uncompressed construction files, as the searches wrote them, have the following SHA-256 hashes. The shipped C_3^+ ($n = 102,400$) file differs only in its first comment line, and `anc/SHA256SUMS` gives both of its hashes.

C_1 construction, 40.75 h	49c83eb6f9f82e244578fab15d056757886679700bd8c8c653f0b100ae882f2d
C_1 construction, 3.97 h	88401bde2a0f165832b12ac3be593633e2fbb7221c0f9d516170a0b701118a8c
C_1 construction, single agent, 44.0 h	475b6c7645aeecf3b0b7248817ed67952c28c0e67145c9c546f9a3f270ee107d
C_3^+ ($n = 97,000$)	1944abaf360b2e0631c71735adbec26d81980bfaa717af49e08dc540bd96a879
C_3^+ ($n = 102,400$)	129b14c8dea20beb68b62249e799ec8bdc702633b348be16b226838882b91f34
Erdős, re-grid ($n = 2,560$)	33af2545ffdf335b099f9c3c11553a98472db412d501d6a5b0426ee265a16a0d
Erdős, shared archive	4183eed832ff8e656f7db5d0457568f9b39f5c52ed9495b8172021ec92a7618f
analytic evaluator	61fe9cdb31be8fb2970f690f8c815638a69201a20ef6de847428d2cbc587dfc8
geometry evaluator	ba4d420a1508c265e268cc1eb417c8e1d046f6e6c04b0d04103fcc4dfff5d6c3
fast packing checker ($N > 2500$)	dfdee7664c5a6af86b572f0f5b3696e609f4938e70bf5c77ba7fa73386b0242d

The independent packing checks decided 21.5 million pairs in the first run’s 193 files. The second check decided 3.9 million pairs over 543 coordinate files. The fast checker for $N > 2500$ matched the pinned evaluator on all 778 tests that evaluator completed and on 174 adversarial cases. At $N = 2401$ it takes 2.5 seconds against 116.

A.6 Public repositories and comparisons

For the prior-art check we used these repositories at the heads we cloned on October 9, 2026. We re-read Liang’s head on October 30, and it was unchanged.

Liang [24]	https://github.com/jasonzliang/circle-packing-sota	5c6526bdba
Castro [6]	https://github.com/arnoldcastro5000/csqv-circle-packing	17de510e06
Abiyev [1]	https://github.com/RasimAbiyev/circle-packing-explorations	217ca6afa9
Sander [35]	https://github.com/ucsandman/discovery-loop	09fc885516
López Preciado [26]	https://github.com/elianalonsolopezpreciado/certified-packings	75697e11ef
Moki and Julio [29]	https://github.com/Hashirama1337s/moki-julio-circle-packing	b8883916ff
Ge [16]	https://github.com/Ge-limin/maxmin3-records	d73e5adb07

Our identity tests used raw coordinates and canonical forms (frame, permutation, and the eight symmetries of the square), with comparisons within 10^{-8} . We read rival papers from arXiv on October 30, 2026, and the Einstein Arena listings from its public API the same day.

B Additional results

B.1 Dated record chains for C_1 and M

Table 12: **Published upper bounds for C_1 and M with their dates.** Values are as printed by their sources. The dates of this paper’s constructions are the timestamps of their dated records; the two Erdős constructions were first released publicly on qwnl.com on October 19, 2026, and the 40.75-hour C_1 construction is first released with this paper. “Arena #” identifies the Einstein Arena entry, and Arena values are as listed on October 30, 2026. Haugland’s paper prints 0.380926 . . . , which AlphaEvolve and TTT-Discover quote as 0.380927. TTT-Discover’s text gives $C_1 \leq 1.50286$ and its table 1.50287.

Value	Holder	Date
<i>First autocorrelation C_1</i>		
1.5098	Matolcsi–Vinuesa [28]	2010
1.5053	AlphaEvolve [30]	2025
1.5032	AlphaEvolve [17]	2025-11
1.503133	ThetaEvolve [41]	2025-11
1.50286	TTT-Discover [43]	2026-01
1.5028628587053106	Together AI [38] (Arena)	2026-03
1.5027452918597408	Hyra (Arena)	2026-08-11
1.5027437719761116	ExoMind-TTS (Arena #2494)	2026-08-12
1.5027436492	CodexProLong (Arena #2504)	2026-08-14
1.5027435984	MILO [21]	2026-09-29
1.5027435319539597	AgentDiscover (Arena #2603)	2026-10-03
1.502740435922766976 . . .	this paper, 3.97 h	2026-10-08
1.502664840776099129 . . .	this paper, 40.75 h	2026-10-09
1.502743486709583	Atris (Arena #2615; leader on 2026-10-30)	2026-10-13
1.502745	TTME [25]	2026-10-23
<i>Erdős minimum overlap M</i>		
0.380926 . . .	Haugland [20]	2016
0.380924	AlphaEvolve [30]	2025
0.380913	LoongFlow [40]	2025-12
0.380876	TTT-Discover [43]	2026-01
0.380874	Aster [5]	2026-02
0.3808703105862199	Together AI [38] (Arena #9)	2026-03-19
0.380868	SimpleTES [42]	2026-04
0.38085867721583955	Hyra (Arena #2440)	2026-08-03
0.3808585748578584	CodexProLong (Arena #2507)	2026-08-15
0.3808567744	MILO [21]	2026-09-29
0.380856738883317420 . . .	this paper, shared archive	2026-10-08
0.380856547875580499 . . .	this paper, re-grid	2026-10-09
0.38085658829748237	CristianBudala (Arena #2623)	2026-10-16
0.3808564043546676	NIGHTFALL-DevinWright (Arena #2628)	2026-10-20
0.380862	TTME [25]	2026-10-23
0.38085625977797644	Blob (Arena #2643; leader on 2026-10-30)	2026-10-29
0.380856403653142	AncientScience-Evolve-2 (Arena #2644)	2026-10-29

B.2 The 3.97-hour C_1 certificate

Proposition 4 (First autocorrelation, 3.97 working hours).

$$\begin{aligned}
C_1 &\leq \frac{9540049024249904979463830078756700797267534524939751265402880000000}{6348434364442838324271548958669223827995209077143850398903628042249} \\
&= 1.5027404359227669764098540748425465856714 \dots
\end{aligned}$$

Certificate. The nonnegative heights in anc/ac1-SX-A43.txt.xz include 106,416 zeros. Exact computation of

all 524,287 autoconvolution coefficients (Lemma 6) gives a unique maximum at $t = 0, 9.5 \times 10^{-9}$ (relative) above the next knot.

B.3 The C_1 lineage from 3.97 to 40.75 hours

Table 13: **Every new best of the shared archive on C_1 from the 3.97-hour certificate to the 40.75-hour certificate.** Hours are the first quarter-hour grid point at which the archive’s best reached the value; distances below Einstein Arena entry #2603 and knot counts are exact; the move class is read from each construction file’s header, with the parent attempt it names (Ak is the archive’s attempt k).

Step	Attempt	Hours $\leq$	Cells	Below #2603	Maximizer moved	Move (parent)
0	43	4.0	262,144	3.10×10^{-6}	start	refinement and soft re-melt (A39)
1	44	4.25	262,144	3.26×10^{-6}	no	ramp snapshot (A39)
2	45	4.25	262,144	3.78×10^{-6}	no	ramp snapshot
3	47	4.25	262,144	4.58×10^{-6}	no	ramp snapshot
4	49	4.5	262,144	4.84×10^{-6}	no	ramp snapshot
5	50	4.5	262,144	4.84×10^{-6}	yes	LSE/AL schedule (shared tool) (A47)
6	51	5.0	262,144	4.92×10^{-6}	yes	LSE/AL schedule (shared tool)
7	52	5.25	131,072	4.97×10^{-6}	yes	AL extraction
8	53	5.25	131,072	5.00×10^{-6}	yes	AL extraction (A52)
9	54	5.5	262,144	5.13×10^{-6}	yes	LSE/AL schedule (shared tool)
10	56	5.5	131,072	5.15×10^{-6}	yes	AL extraction
11	57	5.5	262,144	6.21×10^{-6}	yes	LSE/AL schedule (shared tool)
12	58	5.75	262,144	6.44×10^{-6}	yes	LSE/AL schedule (shared tool)
13	60	6.0	262,144	6.90×10^{-6}	yes	LSE/AL schedule (shared tool)
14	62	6.25	262,144	7.01×10^{-6}	yes	LSE/AL schedule (shared tool)
15	63	6.5	32,768	1.95×10^{-5}	yes	unlabeled
16	66	6.75	65,536	2.85×10^{-5}	yes	LSE/AL schedule (shared tool)
17	67	7.0	65,536	3.22×10^{-5}	yes	octave refinement, then ramp (A63)
18	68	7.0	65,536	3.25×10^{-5}	yes	octave refinement, then ramp (A63)
19	69	7.25	65,536	3.33×10^{-5}	yes	LSE/AL schedule (shared tool)
20	70	7.5	131,072	3.47×10^{-5}	yes	LSE/AL schedule (shared tool)
21	71	7.75	131,072	3.81×10^{-5}	yes	octave refinement, then ramp (A68)
22	72	8.25	131,072	3.92×10^{-5}	yes	octave refinement, then ramp (A68)
23	73	8.5	131,072	3.99×10^{-5}	yes	LSE/AL schedule (shared tool)
24	76	10.0	262,144	4.19×10^{-5}	yes	octave refinement, then ramp (A73)
25	77	10.25	262,144	4.25×10^{-5}	yes	octave refinement, then ramp (A73)
26	81	10.75	131,072	4.33×10^{-5}	yes	octave refinement, then ramp (A69)
27	82	11.0	131,072	4.34×10^{-5}	yes	AL extraction (A81)
28	83	11.25	65,536	4.40×10^{-5}	yes	octave refinement, then ramp (A80)
29	86	12.0	65,536	4.46×10^{-5}	yes	fixed-resolution cycle (A83)
30	87	12.25	65,536	4.46×10^{-5}	yes	LSE/AL schedule (shared tool)
31	88	12.5	131,072	4.87×10^{-5}	yes	octave refinement, then ramp (A83)
32	89	13.75	131,072	4.96×10^{-5}	yes	LSE/AL schedule (shared tool)
33	92	14.0	131,072	4.96×10^{-5}	yes	polish (A89)
34	93	14.25	131,072	5.03×10^{-5}	yes	unlabeled
35	94	14.75	131,072	5.03×10^{-5}	yes	unlabeled
36	96	15.25	131,072	5.07×10^{-5}	yes	unlabeled
37	101	15.75	131,072	5.07×10^{-5}	yes	unlabeled
38	102	16.0	262,144	5.18×10^{-5}	yes	octave refinement, then ramp (A93)
39	105	16.5	262,144	5.21×10^{-5}	yes	octave refinement, then ramp (A93)
40	107	17.25	262,144	5.25×10^{-5}	yes	fixed-resolution cycle (A105)
41	110	17.75	262,144	5.26×10^{-5}	yes	fixed-resolution cycle (A102)
42	111	17.75	262,144	5.28×10^{-5}	yes	octave refinement, then ramp (A103)
43	119	19.25	262,144	5.31×10^{-5}	yes	octave refinement, then ramp (A103)
44	121	19.25	262,144	5.31×10^{-5}	yes	LSE/AL schedule (shared tool)
45	124	20.0	262,144	5.32×10^{-5}	yes	LSE/AL schedule (shared tool)
46	127	20.75	262,144	5.32×10^{-5}	yes	LSE/AL schedule (shared tool) (A111)
47	128	21.0	65,536	5.34×10^{-5}	yes	kick, then cycle or ramp (A80)
48	131	21.25	65,536	5.47×10^{-5}	yes	kick, then cycle or ramp (A80)
49	133	21.25	65,536	5.53×10^{-5}	yes	kick, then cycle or ramp (A80)
50	134	22.0	65,536	5.57×10^{-5}	yes	kick, then cycle or ramp (A131)
51	136	22.25	65,536	5.63×10^{-5}	yes	kick, then cycle or ramp (A133)
52	139	22.5	131,072	5.93×10^{-5}	yes	octave refinement, then ramp (A131)

Step	Attempt	Hours $\leq$	Cells	Below #2603	Maximizer moved	Move (parent)
53	143	23.0	131,072	6.04×10^{-5}	yes	LSE/AL schedule (shared tool)
54	146	23.25	131,072	6.08×10^{-5}	yes	LSE/AL schedule (shared tool)
55	148	23.25	65,536	6.10×10^{-5}	yes	LSE/AL schedule (shared tool)
56	151	23.75	65,536	6.11×10^{-5}	yes	LSE/AL schedule (shared tool)
57	152	24.0	65,536	6.16×10^{-5}	yes	LSE/AL schedule (shared tool)
58	153	24.0	65,536	6.20×10^{-5}	yes	kick, then cycle or ramp (A148)
59	154	24.25	131,072	6.58×10^{-5}	yes	octave refinement, then ramp (A148)
60	158	24.25	131,072	6.58×10^{-5}	yes	octave refinement, then ramp (A148)
61	160	24.5	131,072	6.59×10^{-5}	yes	LSE/AL schedule (shared tool)
62	162	24.75	131,072	6.67×10^{-5}	yes	octave refinement, then ramp (A153)
63	166	25.25	131,072	6.67×10^{-5}	yes	LSE/AL schedule (shared tool)
64	167	25.25	131,072	6.69×10^{-5}	yes	kick, then cycle or ramp (A158)
65	168	25.75	131,072	6.69×10^{-5}	yes	octave refinement, then ramp (A163)
66	169	25.75	262,144	6.83×10^{-5}	yes	octave refinement, then ramp (A162)
67	173	26.0	262,144	6.86×10^{-5}	yes	octave refinement, then ramp (A162)
68	174	26.25	262,144	6.87×10^{-5}	yes	octave refinement, then ramp (A162)
69	176	27.0	262,144	6.88×10^{-5}	yes	LSE/AL schedule (shared tool)
70	180	27.25	262,144	6.91×10^{-5}	yes	octave refinement, then ramp (A158)
71	181	27.25	262,144	6.91×10^{-5}	yes	octave refinement, then ramp (A172)
72	182	27.5	262,144	6.93×10^{-5}	yes	LSE/AL schedule (shared tool)
73	185	28.0	262,144	6.93×10^{-5}	yes	octave refinement, then ramp (A184)
74	186	28.0	262,144	6.95×10^{-5}	yes	LSE/AL schedule (shared tool)
75	188	28.5	262,144	6.95×10^{-5}	yes	LSE/AL schedule (shared tool)
76	189	28.5	262,144	6.95×10^{-5}	yes	octave refinement, then ramp (A184)
77	191	28.75	262,144	6.95×10^{-5}	yes	octave refinement, then ramp (A184)
78	194	29.5	262,144	6.96×10^{-5}	yes	unlabeled
79	199	30.25	262,144	6.99×10^{-5}	yes	kick, then cycle or ramp (A173)
80	201	30.25	262,144	6.99×10^{-5}	yes	kick, then cycle or ramp (A173)
81	206	31.5	262,144	6.99×10^{-5}	yes	polish (A201)
82	209	32.0	262,144	7.00×10^{-5}	yes	re-melt (A201)
83	212	32.5	262,144	7.00×10^{-5}	yes	kick, then cycle or ramp (A173)
84	214	32.75	262,144	7.00×10^{-5}	yes	re-melt (A201)
85	215	33.0	262,144	7.00×10^{-5}	yes	re-melt (A201)
86	217	33.5	262,144	7.07×10^{-5}	yes	unlabeled
87	218	34.25	131,072	7.37×10^{-5}	yes	unlabeled
88	221	34.75	131,072	7.38×10^{-5}	yes	unlabeled
89	225	35.5	131,072	7.46×10^{-5}	yes	unlabeled
90	227	36.25	262,144	7.58×10^{-5}	yes	unlabeled
91	229	36.25	262,144	7.58×10^{-5}	yes	unlabeled
92	232	37.0	131,072	7.59×10^{-5}	yes	unlabeled
93	235	38.0	131,072	7.59×10^{-5}	yes	unlabeled
94	236	38.0	131,072	7.64×10^{-5}	yes	unlabeled
95	237	38.25	262,144	7.66×10^{-5}	yes	unlabeled
96	241	38.75	262,144	7.80×10^{-5}	yes	unlabeled
97	243	39.25	262,144	7.81×10^{-5}	yes	LSE/AL schedule (shared tool)
98	244	39.5	262,144	7.82×10^{-5}	yes	unlabeled
99	248	40.25	262,144	7.85×10^{-5}	yes	octave refinement, then ramp (A240)
100	252	40.75	262,144	7.87×10^{-5}	yes	unlabeled
101	253	40.75	262,144	7.87×10^{-5}	yes	octave refinement, then ramp (A245)

B.4 The single agent's C_1 certificate

The single agent of the same study, Claude Opus 5.5 at effort xhigh in one long conversation with the whole machine and no archive, reached its best C_1 construction at 44.0 working hours:

$$\begin{aligned}
C_1 &\leq \frac{3442093285155344473663833451564581064013851687268579173990400000000}{2290649224533332087530674877075652038248062382591668916427827021627} \\
&= 1.5026714908113410692149648780881857015151\dots
\end{aligned}$$

The heights in anc/475b6c7645aeecf3b0b7248817ed67952c28c0e67145c9c546f9a3f270ee107d .xz, 262,144 cells with 107,967 zeros, have a unique maximizing knot at $t = -162419/524288$. The bound is 7.20×10^{-5} below Atris's #2615, 7,200 times the Einstein Arena's minimum improvement, and 6.65×10^{-6} above the shared archive's 40.75-hour certificate.

B.5 Signed autocorrelation at 102,400 cells

The 102,400-cell function from which the 97,000-cell certificate descends gives

$$\begin{aligned} C_3^+ &\leq \frac{949539615831019216179162964649327653916451840000000}{6553599999999999932445273753600001740890898981493009} \\ &= 1.44888247044528093575747927671511452851 \dots, \end{aligned}$$

with 29,042 negative heights and a unique maximizing index, 149,850. The function lowers its maximum by cancellation: its negative cells carry -0.56 times the net integral, and its autoconvolution dips to -5.83 times $(\int f)^2$ near $t = 0.18$. Its 32-cell means correlate 0.998 with MILO's, so at the coarse scale it is MILO's function. At the scale of the Einstein Arena entry it is a different function. Its correlation with #2609, in the best orientation, is 0.365.

Table 14 follows the single agent's C_3^+ search on October 9. The Arena entry attained its maximum at one knot. After the change of basin and the first Newton-CG finish, every later construction is flat to 10^{-8} on 38 to 41% of its knots, and each re-grid destroyed that structure and rebuilt it lower.

Table 14: **The C_3^+ search, step by step.** Times are UTC on October 9, 2026. Values are the protected evaluator's scores (exact at the certified steps). Knots within 10^{-8} of the maximum use an absolute tolerance.

Step	Time	Cells	Value	Move	Knots within 10^{-8}
0	01:10	65,536	1.449121951852342 (Arena #2609)	starting entry	1 of 131,071
–	01:40–04:40	65,536	every landing above its start	hot re-melts from #2609	–
1	02:40–03:55	25,600	block means alone 1.6558; reconstruction 1.44918	decode MILO's block means, rebuild	–
2	03:55–04:15	25,600	1.4491688 to 1.44908613 in 85 s	Newton-CG	19,335 of 51,199
3	04:50	51,200	1.44889614	Newton re-melt hopping	38,855 of 102,399
4	06:20	102,400	1.44888599, first below MILO	hopping	80,113 of 204,799
5	08:49	102,400	1.44888247 (certified)	hopping	80,568 of 204,799
6	11:35	100,000	remap 1.4909, then 1.44888210	re-grid	76,987 of 199,999
7	13:09	99,000	remap 1.5258, then 1.44886979	second re-grid	79,882 of 197,999
8	13:55	97,000	1.44886703 (certified)	third re-grid	79,834 of 193,999

B.6 Verification scope

Table 15: **Each claimed value passed verification.** Scope and outcome of the checks for each claim.

Claim	Checks	Outcome
C_1 (3.97 h) and Erdős bounds	protected exact evaluator; independent exact checker written from the definitions; Einstein Arena verifier (3.97 h C_1 , shared-archive M)	value confirmed
C_1 (40.75 h, 44.0 h)	protected evaluator (floating point, 8.8×10^{-11} above the exact value); independent exact checker	value confirmed
csqv first run (193)	independent checks: feasibility, previous bests, 719 identity comparisons; full-history check: 766 commits, 3,841 coordinate files, 339,696 values	all pass
csqv second run (113 new, 108 improved)	independent checks: all 414 construction files, Liang's full history rebuilt from a verified mirror, 3,731 identity comparisons	all pass
C_3^+	protected exact evaluator; independent exact checker; Einstein Arena verifier (102,400 cells)	value confirmed

B.7 Circle-packing values

Table 16: **Packing margins over the previous best grow with N .** The previous best at each size was Liang’s. Relative margin is the margin divided by our sum of radii. The last column compares with Packomania’s printed table.

N	sizes	Margin over the previous best			Median relative margin	Median margin over Packomania’s table
		min	median	max		
137–199	12	3.67×10^{-7}	1.07×10^{-4}	6.19×10^{-4}	1.55×10^{-5}	1.21×10^{-3}
200–299	40	6.66×10^{-6}	3.39×10^{-4}	3.13×10^{-3}	4.22×10^{-5}	1.50×10^{-3}
300–399	85	9.97×10^{-6}	5.79×10^{-4}	3.20×10^{-3}	5.79×10^{-5}	2.93×10^{-3}
400–500	87	1.79×10^{-5}	8.53×10^{-4}	4.97×10^{-3}	7.46×10^{-5}	2.57×10^{-3}
501–1600	24	2.50×10^{-5}	2.96×10^{-3}	1.02×10^{-2}	1.81×10^{-4}	1.51×10^{-2}
1681–10000	58	1.63×10^{-4}	6.13×10^{-3}	5.08×10^{-2}	1.84×10^{-4}	6.16×10^{-2}
all	306	3.67×10^{-7}	8.78×10^{-4}	5.08×10^{-2}	7.47×10^{-5}	3.52×10^{-3}

Table 17: **Selected new packings, from the smallest margin to the largest.** Liang held the previous best at each size. $N = 185$ has the smallest margin of the 306 and $N = 9801$ the largest; $N = 9216$ has the smallest margin among the large sizes.

N	Ours: $\sum r_i$ (exact)	Previous best (Liang)	Margin	Over Packomania
137	6.180252863393018908	6.1800746296240865	1.78×10^{-4}	9.81×10^{-4}
185	7.198651847610267365	7.198651480801373	3.67×10^{-7}	1.03×10^{-3}
455	11.352665116737861479	11.352056015179	6.09×10^{-4}	1.76×10^{-3}
529	12.246054012215808674	12.235811635027076	1.02×10^{-2}	1.67×10^{-2}
1521	20.8261647376347404726	20.823038385363	3.13×10^{-3}	1.77×10^{-2}
1681	21.9127643931576471207	21.908331856608118	4.43×10^{-3}	3.19×10^{-2}
2500	26.7487342580083326146	26.744636270403	4.10×10^{-3}	4.17×10^{-2}
4096	34.2699541749965092688	34.256584228498	1.34×10^{-2}	5.56×10^{-2}
5041	38.0254261999067748255	37.975219102001	5.02×10^{-2}	5.75×10^{-2}
7396	46.0815945983454372089	46.03596799696702	4.56×10^{-2}	6.99×10^{-2}
9216	51.4731348926085804798	51.47297148204148	1.63×10^{-4}	9.90×10^{-2}
9801	53.0664060006081925272	53.0156248155581578993	5.08×10^{-2}	8.36×10^{-2}
10000	53.6095660065611926022	53.60334309311	6.22×10^{-3}	9.04×10^{-2}

Table 18: **Exact sums of radii for all 306 packings.** Unequal circles in the unit square. Previous best is the largest published value as of October 30, 2026; Liang held it at every size. Sources and history coverage are described in Appendix A.6. Run 1 is the first run; 2 is the second; 2* replaces a first-run value that also beat the previous best.

N	Ours (exact)	Previous best	Run	Margin
137	6.180252863393018908	6.1800746296240865	1	1.78×10^{-4}
142	6.295677951246816425	6.29544590229006	2	2.32×10^{-4}
146	6.38782481477793825	6.387741228719	1	8.36×10^{-5}
148	6.432813837725447782	6.432731989105226	2	8.18×10^{-5}
156	6.603639017462671874	6.603620180147317003	1	1.88×10^{-5}
157	6.624666534036148959	6.624047849350299358	1	6.19×10^{-4}
158	6.644814970138116362	6.644712963825	1	1.02×10^{-4}
182	7.139690800460969286	7.139578643773063	1	1.12×10^{-4}
185	7.198651847610267365	7.198651480801373	2	3.67×10^{-7}
188	7.258048914776490874	7.257784502262	2	2.64×10^{-4}
191	7.318151495335000215	7.317941398677	1	2.10×10^{-4}
197	7.437019447033417126	7.436991067624	1	2.84×10^{-5}
205	7.586370214315707481	7.58593152635	1	4.39×10^{-4}
206	7.604029579995051711	7.603903612129	1	1.26×10^{-4}
207	7.622256298439724534	7.622201198658	1	5.51×10^{-5}
208	7.640351003724895499	7.639954768449	2*	3.96×10^{-4}

N	Ours (exact)	Previous best	Run	Margin
209	7.658277084748046658	7.657963657677	1	3.13×10^{-4}
210	7.675905467605473087	7.675809055326	1	9.64×10^{-5}
211	7.694537274684769008	7.694453725207	1	8.35×10^{-5}
212	7.712608628407755615	7.712565088592	2	4.35×10^{-5}
217	7.80679666119002324	7.806394277851230227	1	4.02×10^{-4}
219	7.844778866628022924	7.844486169173	2	2.93×10^{-4}
230	8.043720551997481443	8.043355500284	1	3.65×10^{-4}
231	8.060446369119714437	8.060436244514	1	1.01×10^{-5}
235	8.127288800049589413	8.127280870897	1	7.93×10^{-6}
236	8.143769883851493495	8.143218415657	2*	5.51×10^{-4}
237	8.161337722639579324	8.160815684546	2*	5.22×10^{-4}
238	8.178481169206609564	8.177732544592	2*	7.49×10^{-4}
239	8.19552185189848535	8.195152547633	2	3.69×10^{-4}
247	8.337695735934358299	8.337657453498	2	3.83×10^{-5}
248	8.356051847278934624	8.356030683330472	1	2.12×10^{-5}
249	8.373358570759210365	8.373264801181	2	9.38×10^{-5}
251	8.408141749169110775	8.407669216587	2	4.73×10^{-4}
252	8.425246854790941687	8.425177921053259047	1	6.89×10^{-5}
253	8.442989172878737881	8.442907861345068	1	8.13×10^{-5}
265	8.638010415812943158	8.638002402392415	1	8.01×10^{-6}
270	8.718536927553344382	8.718530269452	1	6.66×10^{-6}
275	8.800682728448504174	8.800226795221416	2	4.56×10^{-4}
276	8.817957668499285686	8.81703289155219142	2*	9.25×10^{-4}
277	8.834999349367095091	8.833012846276	1	1.99×10^{-3}
278	8.852205791172315364	8.849078849651	1	3.13×10^{-3}
279	8.868518437897543248	8.868057555173	2*	4.61×10^{-4}
280	8.885030250053348558	8.884272111120044598	1	7.58×10^{-4}
282	8.918493325379532771	8.917797248353	2*	6.96×10^{-4}
285	8.96791620811850299	8.967891423896	1	2.48×10^{-5}
290	9.043321702076018638	9.04325923378808564	2	6.25×10^{-5}
293	9.089325767255382313	9.088930498053	2	3.95×10^{-4}
294	9.104272968673472539	9.104113342019	1	1.60×10^{-4}
295	9.119665026277559479	9.119378513625	1	2.87×10^{-4}
297	9.149914661992509187	9.149420862049253	2*	4.94×10^{-4}
298	9.165542978165433146	9.164640675262	1	9.02×10^{-4}
299	9.181211255170868153	9.179904120375110247	1	1.31×10^{-3}
301	9.210605939119740701	9.209844096512	1	7.62×10^{-4}
302	9.22700188390250834	9.22573648925747	1	1.27×10^{-3}
303	9.242592646262930874	9.242272991889	1	3.20×10^{-4}
304	9.258437915144478758	9.25726107191	2*	1.18×10^{-3}
305	9.274141800935633951	9.273668547572258825	1	4.73×10^{-4}
306	9.289406225596147138	9.288950901832218	2*	4.55×10^{-4}
307	9.305966463450673529	9.305804368812488	2	1.62×10^{-4}
308	9.321712880738834391	9.321332860923203474	1	3.80×10^{-4}
309	9.337522992279633285	9.336984289826	1	5.39×10^{-4}
310	9.353544233519262189	9.35347181516	2	7.24×10^{-5}
312	9.385339629618463925	9.385012613536	2	3.27×10^{-4}
313	9.400829588631472821	9.400819251021	2	1.03×10^{-5}
314	9.415755705722215604	9.415455693751	1	3.00×10^{-4}
317	9.463025175677656707	9.462942784057088	2	8.24×10^{-5}
319	9.492589332754020739	9.49257936595	2	9.97×10^{-6}
322	9.536854656248981044	9.536742445151	1	1.12×10^{-4}
323	9.55007468961524108	9.550057569242	2	1.71×10^{-5}
324	9.564320590060022936	9.564266944475238	2	5.36×10^{-5}
326	9.592508448082535822	9.592121131691	1	3.87×10^{-4}
327	9.607065015508765455	9.606720244607	2*	3.45×10^{-4}
328	9.621690585460442514	9.620923544343102	2*	7.67×10^{-4}
330	9.650833311758388682	9.65044098968523	2*	3.92×10^{-4}
331	9.665811403516624348	9.665118699705321	2*	6.93×10^{-4}
332	9.680828738536478292	9.680090467272466	2*	7.38×10^{-4}
333	9.695701083590445449	9.694588243127	1	1.11×10^{-3}
334	9.710331621276520065	9.709972693508	1	3.59×10^{-4}
335	9.725018848379329933	9.721816063906	2*	3.20×10^{-3}
336	9.738739854456882468	9.73653445863842	1	2.21×10^{-3}
337	9.753233327246692098	9.75210127893	1	1.13×10^{-3}
338	9.768180431741219585	9.767449549800071	1	7.31×10^{-4}

<i>N</i>	Ours (exact)	Previous best	Run	Margin
339	9.783441153023109691	9.782725516609080259	1	7.16×10^{-4}
340	9.799084684767045642	9.798959804691	2	1.25×10^{-4}
341	9.813888661890454879	9.812931944945	1	9.57×10^{-4}
342	9.829547330330974519	9.8295144572022	1	3.29×10^{-5}
343	9.844099982131961271	9.843018954046475	1	1.08×10^{-3}
344	9.859515561168946915	9.859295071497252457	1	2.20×10^{-4}
345	9.87358552086328381	9.871546836350314299	1	2.04×10^{-3}
346	9.888383325104795847	9.887974822624	2*	4.09×10^{-4}
347	9.902696554149374503	9.902123538691	1	5.73×10^{-4}
348	9.918143382694258488	9.91796879350822	1	1.75×10^{-4}
349	9.932279281546956182	9.932003707993196804	2*	2.76×10^{-4}
350	9.947361055015214958	9.946196231108	1	1.16×10^{-3}
351	9.961904602931062361	9.961599925030004095	1	3.05×10^{-4}
352	9.975865175992803554	9.97581660387	2	4.86×10^{-5}
353	9.990153986516426724	9.990123404486	2	3.06×10^{-5}
355	10.018386622881226899	10.017348811819142	2*	1.04×10^{-3}
356	10.031525871972499036	10.028837895494	1	2.69×10^{-3}
357	10.043155996959930828	10.042391218778	2*	7.65×10^{-4}
358	10.057435413129955764	10.056424852513222791	2*	1.01×10^{-3}
359	10.071582829648081267	10.069701157534	2*	1.88×10^{-3}
360	10.085823006988295958	10.084055771106213388	2*	1.77×10^{-3}
361	10.10000208397209231	10.099355683993142262	2*	6.46×10^{-4}
362	10.114018040278457605	10.113675817734	2	3.42×10^{-4}
363	10.12750840269437469	10.12629987944	1	1.21×10^{-3}
365	10.155217304591001733	10.154922603302	2	2.95×10^{-4}
366	10.168397181761067536	10.16811450529	1	2.83×10^{-4}
367	10.182648829166982135	10.182614617914	2	3.42×10^{-5}
368	10.196857425007338991	10.1963592751	2*	4.98×10^{-4}
369	10.210713274007651438	10.210656019959	2	5.73×10^{-5}
370	10.224720905620632762	10.224249820483148	2	4.71×10^{-4}
371	10.239004135209290486	10.238525853697	2	4.78×10^{-4}
372	10.253076166349626014	10.252311504903277	2*	7.65×10^{-4}
373	10.267091557747521027	10.266214821013454	2*	8.77×10^{-4}
374	10.281299169870545924	10.28066989177917	1	6.29×10^{-4}
375	10.295080178778702707	10.295012338309	2	6.78×10^{-5}
377	10.323772258880832415	10.323534933290476	2	2.37×10^{-4}
378	10.338898189785464554	10.337798454625	2*	1.10×10^{-3}
379	10.353188704949128584	10.352780705647	1	4.08×10^{-4}
380	10.36725773848013405	10.365846051273	2*	1.41×10^{-3}
382	10.39612612311933295	10.395847164838	2*	2.79×10^{-4}
385	10.437860471197656089	10.437625654595	1	2.35×10^{-4}
386	10.452039102134645595	10.451308287138255456	2*	7.31×10^{-4}
387	10.465876421007100064	10.463628278476463	1	2.25×10^{-3}
388	10.479132042252730757	10.476431485678	2*	2.70×10^{-3}
389	10.491437817694915549	10.488500253096	2*	2.94×10^{-3}
390	10.50300789636752495	10.502128814062322122	2*	8.79×10^{-4}
391	10.51642997300092863	10.515311384285056	2*	1.12×10^{-3}
392	10.529610580101346057	10.528692881753	2*	9.18×10^{-4}
393	10.542195825826727665	10.540910774497	1	1.29×10^{-3}
394	10.555538160481650851	10.554218362585011	1	1.32×10^{-3}
395	10.569110554063826485	10.567762415409	2*	1.35×10^{-3}
396	10.582374620028917666	10.58099542035442	2*	1.38×10^{-3}
397	10.595687917110271892	10.595108503166227	2	5.79×10^{-4}
398	10.608312418820262668	10.606684686129	2*	1.63×10^{-3}
399	10.621626878821647983	10.620528360865329324	2*	1.10×10^{-3}
400	10.634726119716070452	10.631978493560295	1	2.75×10^{-3}
401	10.647716037964392313	10.646889813174	1	8.26×10^{-4}
402	10.661209221981892196	10.659068621689	2*	2.14×10^{-3}
403	10.674724153898649818	10.674019157454056839	2	7.05×10^{-4}
404	10.688375617963671047	10.687354545195152	2*	1.02×10^{-3}
405	10.701715493219220473	10.700808830268	2*	9.07×10^{-4}
406	10.715267351039029484	10.714986963599	1	2.80×10^{-4}
407	10.729245680329723122	10.728493628098237	1	7.52×10^{-4}
408	10.743122836965742822	10.742051825622447	1	1.07×10^{-3}
410	10.769978677997050915	10.768985209394291	2*	9.93×10^{-4}
411	10.783634149286997851	10.78257853628	2*	1.06×10^{-3}

N	Ours (exact)	Previous best	Run	Margin
414	10.824247248009148147	10.824057489898	1	1.90×10^{-4}
415	10.838054903166930373	10.83767261400815267	2	3.82×10^{-4}
416	10.851392838276020644	10.85092468347	2*	4.68×10^{-4}
417	10.865028494085665832	10.864815186674136	2*	2.13×10^{-4}
418	10.879200713107233879	10.878523629038142328	2*	6.77×10^{-4}
419	10.892923217745895404	10.892665080332	2	2.58×10^{-4}
420	10.905746469756890149	10.905704806942	2	4.17×10^{-5}
421	10.919271671117306093	10.919253797972	2	1.79×10^{-5}
422	10.933153784290083566	10.933019712196083	2	1.34×10^{-4}
429	11.023620199547893579	11.0235188005	2	1.01×10^{-4}
431	11.047869192158593843	11.047800143152	2	6.90×10^{-5}
432	11.059686509241607732	11.05950471636	2	1.82×10^{-4}
433	11.07180813453815854	11.071788062005	2	2.01×10^{-5}
434	11.08469235627254661	11.084566793068	2	1.26×10^{-4}
435	11.096980672263004299	11.09674075478	2*	2.40×10^{-4}
436	11.110025353489768893	11.109013784905	2*	1.01×10^{-3}
437	11.122695806372661652	11.122021758613112	2*	6.74×10^{-4}
438	11.134839101675239046	11.13457027483	1	2.69×10^{-4}
439	11.147629354619641796	11.147344979031	2	2.84×10^{-4}
440	11.159615153423291399	11.159484571692378	2*	1.31×10^{-4}
441	11.171982651663833435	11.171344962483127014	1	6.38×10^{-4}
442	11.184490987463337887	11.183744585430095	1	7.46×10^{-4}
443	11.196513494210879254	11.19641700082	2*	9.65×10^{-5}
446	11.234974044276115496	11.234571141333	1	4.03×10^{-4}
447	11.24830239337127629	11.247594138506289377	2*	7.08×10^{-4}
448	11.261427073638572046	11.260595834042038	1	8.31×10^{-4}
449	11.274439247889263104	11.273153984076	2*	1.29×10^{-3}
450	11.288382110876846532	11.286856223680338153	2*	1.53×10^{-3}
451	11.301645614570400843	11.299866162827167	2*	1.78×10^{-3}
452	11.314438958825364132	11.313692402308046	2*	7.47×10^{-4}
453	11.326936777094785668	11.326450332304	2*	4.86×10^{-4}
454	11.340365936017801164	11.339991901574384328	2	3.74×10^{-4}
455	11.352665116737861479	11.352056015179	2*	6.09×10^{-4}
456	11.365532134589393053	11.36498937378	2*	5.43×10^{-4}
457	11.378521899153775402	11.378259041302492	2*	2.63×10^{-4}
458	11.391405692236707436	11.39067932275504	2	7.26×10^{-4}
459	11.405016199113972102	11.404111026276	2*	9.05×10^{-4}
460	11.418032166461032314	11.416851702401	2*	1.18×10^{-3}
461	11.43047172379733585	11.429618959065	2*	8.53×10^{-4}
462	11.443482108404651252	11.440362320075362442	2*	3.12×10^{-3}
463	11.456320226995420563	11.453418970267057036	2*	2.90×10^{-3}
464	11.469002151498845839	11.46425431102548	2*	4.75×10^{-3}
465	11.481048139187243047	11.47665682458905	2*	4.39×10^{-3}
466	11.493805662681604063	11.489213060903426	2*	4.59×10^{-3}
467	11.506304631730447725	11.501332372758387	2*	4.97×10^{-3}
468	11.517371765714565939	11.514331674698486	2	3.04×10^{-3}
469	11.52785977071742016	11.526435282894049185	2*	1.42×10^{-3}
470	11.53928527759627643	11.538072626079	2*	1.21×10^{-3}
471	11.551889012026367603	11.549771140737	2*	2.12×10^{-3}
472	11.563424342061380308	11.562200541317	2	1.22×10^{-3}
473	11.575146830922995816	11.57392384541202	2*	1.22×10^{-3}
474	11.587061088836031854	11.585691131256	2*	1.37×10^{-3}
475	11.599219194526562834	11.598296060694024	2	9.23×10^{-4}
476	11.61155852453233632	11.610694328765323	2*	8.64×10^{-4}
477	11.624452310597051141	11.623024467761	2*	1.43×10^{-3}
478	11.636914830544697899	11.635281049091	2	1.63×10^{-3}
479	11.649389015997359249	11.64769164874	2*	1.70×10^{-3}
480	11.661661158415590252	11.65953501583	2*	2.13×10^{-3}
481	11.673763300518623944	11.671384096329	2*	2.38×10^{-3}
482	11.68598075827936964	11.685360636960487	2*	6.20×10^{-4}
483	11.698136683837465316	11.695599609112100993	2*	2.54×10^{-3}
484	11.710268820056188339	11.708318294287	1	1.95×10^{-3}
485	11.722203751520105601	11.721377250648	2*	8.27×10^{-4}
486	11.734219479816948341	11.733992103902	1	2.27×10^{-4}
487	11.746160535468672659	11.743513784756045	1	2.65×10^{-3}
488	11.758033928429962301	11.75773729545031	2*	2.97×10^{-4}

N	Ours (exact)	Previous best	Run	Margin
489	11.770418177018486664	11.769610100652270179	1	8.08×10^{-4}
491	11.795357516976009804	11.794376896928	1	9.81×10^{-4}
492	11.807911061361513592	11.807466669778	2*	4.44×10^{-4}
494	11.833126626504711719	11.832363607649	1	7.63×10^{-4}
495	11.84567636635822177	11.843409517905	1	2.27×10^{-3}
496	11.857947090500060434	11.854375105080131155	1	3.57×10^{-3}
497	11.869935767222369992	11.866743178315495	2*	3.19×10^{-3}
498	11.881828015949956634	11.878206095286	2*	3.62×10^{-3}
499	11.893702525652968591	11.890330089707353182	2*	3.37×10^{-3}
500	11.904692790839477703	11.903170336229	2	1.52×10^{-3}
529	12.246054012215808674	12.235811635027076	2*	1.02×10^{-2}
545	12.438370773369283606	12.436533265502467	2*	1.84×10^{-3}
576	12.783287280809794978	12.775152325972053	2*	8.13×10^{-3}
613	13.189103030713546166	13.185931386373403	2*	3.17×10^{-3}
625	13.319991865908629933	13.31520620527706	2*	4.79×10^{-3}
676	13.856049222648425738	13.853532292832384	2*	2.52×10^{-3}
685	13.955172747399860736	13.947043137408414568	2*	8.13×10^{-3}
729	14.400812568074163977	14.393408217492	2*	7.40×10^{-3}
761	14.707492637208948155	14.70204839839	2*	5.44×10^{-3}
784	14.9373697687773226	14.933651793985	1	3.72×10^{-3}
841	15.473997494816899313	15.471199967493	2*	2.80×10^{-3}
900	16.013142231633788875	16.011101957183	2*	2.04×10^{-3}
925	16.227491472071067454	16.220391311081	2*	7.10×10^{-3}
961	16.547694231105267125	16.546730901966101565	2*	9.63×10^{-4}
1013	16.983517564206329008	16.981739092791	2*	1.78×10^{-3}
1024	17.06085052872915334	17.05935881728	2*	1.49×10^{-3}
1089	17.621467119689173457	17.621287603134	1	1.80×10^{-4}
1156	18.1547164235663760493	18.154691434658	2	2.50×10^{-5}
1225	18.692401713825186423	18.68534103330226597	2*	7.06×10^{-3}
1296	19.2118093599524339584	19.210318925583	2	1.49×10^{-3}
1369	19.7457995076503394942	19.743771260852284001	2*	2.03×10^{-3}
1444	20.3075855902416606423	20.3034779015522608607	2*	4.11×10^{-3}
1521	20.8261647376347404726	20.823038385363	2*	3.13×10^{-3}
1600	21.382176450346681265	21.379516344136200715	2	2.66×10^{-3}
1681	21.9127643931576471207	21.908331856608118	2	4.43×10^{-3}
1764	22.4578968732403947246	22.4520051255372620683	2	5.89×10^{-3}
1849	22.9898051034169263807	22.985298580705	2	4.51×10^{-3}
1936	23.5300922258535365191	23.5267489263813161838	2	3.34×10^{-3}
2025	24.0676680013099326851	24.063783503649375	2	3.88×10^{-3}
2116	24.5942010474018950755	24.5897477732620810592	2	4.45×10^{-3}
2209	25.1413895101643834221	25.140593125125	2	7.96×10^{-4}
2304	25.6775861979922932624	25.67601721119147	2	1.57×10^{-3}
2401	26.2121996227255827636	26.204628504855	2	7.57×10^{-3}
2500	26.7487342580083326146	26.744636270403	2	4.10×10^{-3}
2601	27.2877616710413712979	27.277167007842422	2	1.06×10^{-2}
2704	27.8273469755102592661	27.823137574183	2	4.21×10^{-3}
2809	28.3670156025491741709	28.366020412230327	2	9.95×10^{-4}
2916	28.902573124336435516	28.900709505172	2	1.86×10^{-3}
3025	29.4378743580168108693	29.434933513113315	2	2.94×10^{-3}
3136	29.9655094127494333891	29.960279941385227	2	5.23×10^{-3}
3249	30.5126109390480752217	30.511081629717214	2	1.53×10^{-3}
3364	31.0520113445074166928	31.043887002321	2	8.12×10^{-3}
3481	31.587913644699342109	31.5855752548044300599	2	2.34×10^{-3}
3600	32.126934220549428686	32.12538745069016	2	1.55×10^{-3}
3721	32.663267852677569786	32.66089172144	2	2.38×10^{-3}
3844	33.193823210866161958	33.180384313552	2	1.34×10^{-2}
3969	33.7366223131194701731	33.724449969923285	2	1.22×10^{-2}
4096	34.2699541749965092688	34.256584228498	2	1.34×10^{-2}
4225	34.8104993058235709634	34.7987129787464730395	2	1.18×10^{-2}
4356	35.352981803283012785	35.344143157206375	2	8.84×10^{-3}
4489	35.8881843850942609884	35.88786423495221127	2	3.20×10^{-4}
4624	36.4273201321753786192	36.4262906927244911081	2	1.03×10^{-3}
4761	36.9621950207806265272	36.9524893011524330359	2	9.71×10^{-3}
4900	37.4924318418774285388	37.4737223188770673348	2	1.87×10^{-2}
5041	38.0254261999067748255	37.975219102001	2	5.02×10^{-2}
5184	38.5659582142139985128	38.5606233090760161769	2	5.33×10^{-3}

N	Ours (exact)	Previous best	Run	Margin
5329	39.1100156109129384405	39.10078913399521	2	9.23×10^{-3}
5476	39.6492888005970851704	39.6484565852801326352	2	8.32×10^{-4}
5625	40.1734262130076659303	40.155096145848184	2	1.83×10^{-2}
5776	40.7107950380194638342	40.6992702134202371675	2	1.15×10^{-2}
5929	41.252631687843275651	41.243099663071	2	9.53×10^{-3}
6084	41.7892393930579420645	41.78262494966233	2	6.61×10^{-3}
6241	42.334382883547468397	42.317650758177	2	1.67×10^{-2}
6400	42.8706472713367775853	42.8624330274154088561	2	8.21×10^{-3}
6561	43.4124697449814533816	43.4077685338294833929	2	4.70×10^{-3}
6889	44.4850008242160214055	44.4792579069230272248	2	5.74×10^{-3}
7056	45.0153156776231525877	45.003245497698	2	1.21×10^{-2}
7225	45.5507848462751684862	45.544167858726	2	6.62×10^{-3}
7396	46.0815945983454372089	46.03596799696702	2	4.56×10^{-2}
7569	46.6226542867372155364	46.60251442977	2	2.01×10^{-2}
7744	47.1539558491425514707	47.130009797431	2	2.39×10^{-2}
8100	48.2363866771907498001	48.212395982255	2	2.40×10^{-2}
8281	48.7760475329245679329	48.749808455343	2	2.62×10^{-2}
8464	49.316789802446577971	49.312183797404	2	4.61×10^{-3}
8649	49.852961777958462785	49.837391057263375	2	1.56×10^{-2}
8836	50.392955349089191156	50.39153822784036	2	1.42×10^{-3}
9025	50.9340588223221777488	50.928027389488	2	6.03×10^{-3}
9216	51.4731348926085804798	51.47297148204148	2	1.63×10^{-4}
9409	52.0089673982626189371	52.0062291183621934845	2	2.74×10^{-3}
9604	52.5407108431373618664	52.530893570438	2	9.82×10^{-3}
9801	53.0664060006081925272	53.0156248155581578993	2	5.08×10^{-2}
10000	53.6095660065611926022	53.60334309311	2	6.22×10^{-3}