Machine-Checked Exact-Rational Variance-Hull, Two-Cap Mask, and Factor-Two Certificates for Euclidean Diameter in Dynamic Streams

Lluis Eriksson*

November 25, 2026

Abstract

Jeff Giliberti and MohammadTaghi Hajiaghayi (`arXiv:2611.33139`, November 24, 2026) resolved the update-time bottleneck for dynamic-stream high-dimensional Euclidean diameter: for every fixed $c > \sqrt{3}$, their one-pass algorithm maintains an explicit factor-c witness pair in sublinear space $\tilde{O}(N^{1-\epsilon_c}\mathrm{poly}(d))$ and worst-case update time $\mathrm{poly}(d, \log N)$, breaking the $N^{2/3}$ update-time barrier of direct asymmetric Gaussian embeddings at factor $c = 2$ while matching polynomial space lower bounds. We present a machine-checked, zero-floating-point formalization and exact-$\mathbb{Q}$ certificate auditor for the algebraic, variance-hull, two-cap mask, and Appendix A.1 factor-two core of this construction. First, an 18-theorem zero-dependency Lean 4 kernel (`EuclideanDiameterStream.lean`, Lean 4.35.0; 0 `sorry`, 0 warnings, 0 `Classical.choice`, 5 zero-axiom theorems) formalizes the cleared parallel-axis identity, Lemma 5.1's $\sqrt{2a^2-1}$ variance-hull dilation bound and two-simplex tightness, Eq. (13) bounded witness radii, Section 8.2 Algorithm 4's exact variance increment $V' - V = \gamma(R^2 - V) - \gamma^2 R^2 > (\theta^2/64)V$ and $8V_0$ cap, Lemma 8.1's candidate-score bound, the 3-4-5 spherical-cap containment $t \geq 3/5 \Rightarrow (t-3/5)^2 + q \leq 16/25$, the SOS potential margin $(\lambda - t)^2 - q \geq \lambda^2/2 - 1$, Section 8.2 geometric shrinkage $|P_{i+1}| < \frac{3}{5}|P_i|$, and the complete Appendix A.1 factor-two rational parameter suite ($2a^2 - 1 = 3.999122 < 4$, $\epsilon_0 > 1/6345504000 > 10^{-10}$). Second, `EUCLIDEAN-DIAM-CERT-1.0` audits four exact-$\mathbb{Q}$ modes across 16 canonical certificates (0.181–8.834 ms) and rejects 62/62 live adversarial forgeries.

1 Introduction

Maintaining the Euclidean diameter $\mathrm{diam}(P) = \max_{x,y\in P}\|x - y\|$ of a point set $P \subset \mathbb{R}^d$ under arbitrary insertions and deletions exhibits a sharp structural separation from insertion-only streams [1]: polynomial space $N^{\tilde{\Omega}(1/c^2)}$ is unavoidable for any constant factor c [5, 6], and direct dynamic implementations of Indyk's asymmetric embedding [4] incur $N^{2/3}$ update time at $c = 2$. Giliberti and Hajiaghayi [3] separated update time from space by combining three exact rational pillars:

1. *Variance-hull geometry (Lemma 5.1):* For any nonempty finite sample $S \subset \mathbb{R}^d$ and $a > 1$, the farthest-distance hull $H_a(S) = \bigcap_q B(q, aR_S(q))$ admits the exact variance representation $H_a(S) = \bigcup_\lambda B(\mu_\lambda, \sqrt{(a^2-1)V_\lambda})$, yielding
$$\mathrm{diam}(H_a(S)) \leq \sqrt{2a^2-1}\,\mathrm{diam}(S), \qquad (1)$$
attained by two orthogonal regular simplices and approaching the $\sqrt{3}$ barrier as $a \downarrow \sqrt{2}$.
2. *Rational variance certificates (Algorithm 4, Section 8):* Starting from a farthest pair (a_0, b) with $V_0 = \|a_0 - b\|^2/4$, transferring weight $\gamma = \theta/32$ (where $\theta = (1+\eta)^2 - 1$) to a farthest uncovered point strictly increases the variance by
$$V' - V = \gamma(R^2 - V) - \gamma^2 R^2 > \frac{\theta^2}{64}V \qquad (2)$$
while staying capped by $V \leq 8V_0$, terminating in $O(\eta^{-2})$ exact rational iterations with $S \subseteq B(c, (1+\eta)\sqrt{V})$.
3. *Two-stage 3-4-5 cap masks & Table 3 constants (Appendix A.1):* Choosing $a = 1581/1000$, $\lambda =$

*Independent researcher; manuscript prepared with transparent AI assistance via Google Antigravity. Source code, Lean 4 proofs, and exact certificates: https://github.com/lluiseriksson/euclidean-diameter-stream-certificates.

79/50, $\eta_{\text{in}} = 1/100$, $\kappa = 109/100$, and $B_\kappa = 2000$ forces the first-stage mask $\phi \le 21/20$ inside the $t \ge 3/5$ spherical cap contained in $B(\frac{3}{5}u, \frac{4}{5})$, empties the second-stage mask via $em > 16133/24000 > 84/125$, keeps all three squared factors below 4, and yields space exponent $1 - 10^{-10}$.

Contributions and Scope. We formally mechanize and audit the deterministic algebraic, geometric, variance-certificate, two-cap mask, and de-amortized shrinkage core of [3]. We explicitly scope our Lean 4 kernel away from the continuous Gaussian-grid probability couplings, Sion minimax invocation, and communication-complexity lower bound (Sections 6–7, 9–10):

1. **Zero-dependency Lean 4 kernel (§2):** 18 constructive theorems in `EuclideanDiameterStream.lean` (Lean 4.35.0 [2]), 0 `sorry`, 0 warnings, 0 `Classical.choice`, and 5 zero-axiom theorems.
2. **Pure-stdlib exact-$\mathbb{Q}$ verifier (§3):** `EUCLIDEAN-DIAM-CERT-1.0`, auditing Algorithm 4 replay, Lemma 5.1 hulls and witness radii, the full Appendix A.1 factor-two suite, and dynamic insert/delete ball chains with zero floating-point arithmetic.
3. **Empirical atlas & red-team audit (§4):** 16 canonical certificates (0.181–8.834 ms), exact Algorithm 4 bit-length growth sweeps, two-tetrahedra rational hull realizations, exact Taylor mask thresholds, and 62/62 rejected live adversarial forgeries.

2 Zero-Dependency Lean 4 Kernel

We formalize 18 theorems in `lean/EuclideanDiameterStream.lean` (Lean 4.35.0, zero external libraries), audited via `lean/AuditAxioms.lean` with 0 warnings, 0 `sorry`, 0 `Classical.choice`, and 0 `Lean.ofReduceBool` (Table 1). Five theorems depend on zero axioms (`none`) and 13 depend solely on `[propext, Quot.sound]`.

Table 1: Complete 18-theorem Lean 4.35.0 kernel in `EuclideanDiameterStream.lean` (0 `sorry`, 0 warnings, 0 `Classical.choice`, 5 zero-axiom theorems).

#	Lean 4 Theorem Identifier	Axiom Footprint
1	`parallel_axis_cleared_identity`	`[propext, Quot.sound]`
2	`hull_diameter_cauchy_schwarz`	`[propext, Quot.sound]`
3	`hull_sqrt3_barrier`	`[propext, Quot.sound]`
4	`two_simplex_tightness_exact`	`[propext, Quot.sound]`
5	`bounded_witness_radius`	`[propext, Quot.sound]`
6	`algorithm4_variance_update_identity`	`[propext, Quot.sound]`
7	`algorithm4_variance_growth_lb`	`[propext, Quot.sound]`
8	`algorithm4_variance_cap_8V0`	`[propext, Quot.sound]`
9	`candidate_score_cauchy_schwarz`	`[propext, Quot.sound]`
10	`spherical_cap_345_containment`	`[propext, Quot.sound]`
11	`potential_quadratic_margin_sos`	`[propext, Quot.sound]`
12	`potential_monotone_eq29_identity`	`[propext, Quot.sound]`
13	`first_mask_endpoint_exact`	none
14	`second_mask_empty_exact`	none
15	`factor_two_table3_three_cases`	none
16	`appendix_a1_oracle_and_coupling_exact`	none
17	`rebuild_geometric_shrinkage`	`[propext, Quot.sound]`
18	`appendix_a1_scale_chain_and_exponents`	none

Proposition 1 (Constructive Dynamic-Diameter Kernel)**.**

1. Variance-hull algebra (Theorems 1–5): *Inductive cleared-integer parallel-axis identity* $W^2 \sum w_i(a_i - y)^2 = \sum w_i(Wa_i - S)^2 + W(S - Wy)^2$ *over arbitrary weighted lists; cleared Cauchy–Schwarz hull bound* $(h + s(u + w))^2 \le (2s^2 + 1)L^2$ *via an explicit sum-of-squares decomposition; strict* $\sqrt{3}$ *barrier interval* $(3, c^2)$*; exact two-simplex tightness* $(h+2(A-1)h)^2 = (2A-1)S^2$*; and Eq. (13) witness radius bound* $(a-1)R < D_{\text{univ}}$.
2. Algorithm 4 & candidate scores (Theorems 6–9): *Cleared exact 1-step variance identity (2); multiplicative growth lower bound* $V' - V > (\theta^2/64)V$*; upper variance cap* $V \le 8V_0$*; and Lemma 8.1 candidate-score bound* $(du + dt)^2 \le (p^2 + d^2)H^2$.
3. Two-cap masks, Appendix A.1, & shrinkage (Theorems 10–18): *Exact 3-4-5 spherical-cap containment* $(5t-3D)^2+25q \le 16D^2$*; SOS potential margin* $2((\lambda-t)^2-q)-(\lambda^2-2D^2) = (\lambda-2t)^2+2(D^2-t^2-q) \ge 0$*; Eq. (29) identity; exact first- and second-mask rational Taylor bounds; all Table 3 three-case strict margins; oracle and Eq. (32) coupling inequalities; Section 8.2 geometric shrinkage* $5|P_{i+1}| < 3|P_i|$*; and Section A.1.6 space-exponent suite (*$1/6345504000 > 10^{-10}$*,* $s(B_\kappa + 1) - 1 = -2/2003 < 0$*).*

3 Exact-$\mathbb{Q}$ Verifier Architecture

`EUCLIDEAN-DIAM-CERT-1.0` (`src/diameter_certifier.py`) is a pure-stdlib zero-float verifier enforcing strict JSON types, reduced canonical rationals (with string-length and integer bit-length DoS caps), and exact squared

Euclidean arithmetic across four modes:

1. **Mode 1 (`variance_certificate_alg4`):** Replays Section 8.1.2 Algorithm 4 step-by-step in exact $\mathbb{Q}$, auditing (2), $V' \le 8V_0$, containment $R^2 \le (1+\eta)^2 V$, and parallel-axis anchor score $H^2 \ge V + \|c-y\|^2$.
2. **Mode 2 (`variance_hull_and_tightness`):** Independently verifies Lemma 5.1 hull membership for two distributions (λ, γ), exact two-draw identity $h^2 + V_\lambda + V_\gamma \le L_S^2$, hull dilation (1), and an exterior witness ball satisfying Eq. (13) $(a-1)R_S(q) < D_{\text{univ}}$.
3. **Mode 3 (`factor_two_two_cap_audit`):** Certifies the full Appendix A.1 scale chain, all three squared approximation inequalities, first-mask $g(t_0) < 0$ and Taylor $F_{\text{lb}}(t_0) > 1 + \eta_{\text{mask}}$, spherical-cap points, second-mask emptiness $e_{\text{lb}} m > (1+\eta_{\text{mask}}) r_2^2$, oracle Chernoff margin, and Section A.1.6 space exponents.
4. **Mode 4 (`dynamic_stream_chain_replay`):** Replays arbitrary dynamic `insert`/`delete` streams in $\mathbb{Q}^d$, routing active points strictly through authenticated intact covering balls (disabled invalidated balls cannot hide points from the residual) plus an exact leaf, verifying that the active witness pair originates from an intact ball or leaf, attains the exact chain score, and satisfies $L^2 \le \text{diam}(P)^2 \le ((1+\eta)^2+1)L^2$ at every checkpoint alongside Section 8.2 shrinkage $|P_{i+1}| < \frac{3}{5}|P_i|$.

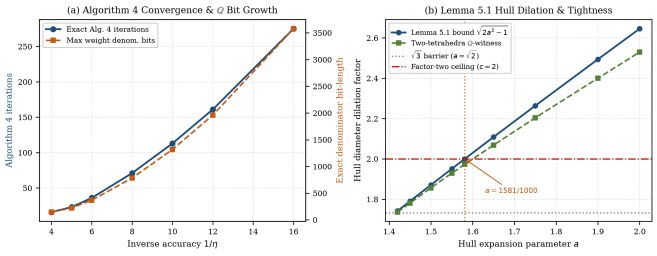

Figure 1: (a) Exact-$\mathbb{Q}$ Algorithm 4 convergence (16 to 275 iterations) and maximum weight denominator bit-length (146 to 3577 bits) on an 8-point $\mathbb{R}^3$ cluster across $1/\eta \in \{4, \ldots, 16\}$. (b) Lemma 5.1 hull dilation $\sqrt{2a^2-1}$ vs. exact rational two-tetrahedra witness in $\mathbb{R}^7$ ($L_S^2 = 40$).

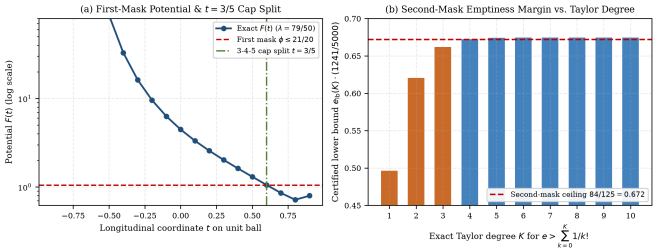

Figure 2: (a) Exact longitudinal potential $F(t)$ at $\lambda = 79/50$ crossing below the first-mask threshold $21/20$ strictly after the 3-4-5 cap split $t = 3/5$. (b) Second-mask global lower bound $e_{\text{lb}}(K) \cdot (1241/5000)$ across Taylor degrees $K = 1 \ldots 10$; $K = 4$ (65/24) is the minimal certifying degree.

4 Empirical Atlas and Red-Team Audit

Running `scripts/run_all_experiments.py` produces 16 canonical certificates (4 per mode) and four exact-$\mathbb{Q}$ studies, audited end-to-end by `scripts/verify_all.py` with SHA-256 verification across all 41 tracked artifacts.

Algorithm 4 Bit Growth and Hull Tightness (Fig. 1). Across $\eta \in \{1/4, 1/5, 1/6, 1/8, 1/10, 1/12, 1/16\}$ on our 8-point $\mathbb{R}^3$ benchmark, Algorithm 4 converges in $16, 23, 36, 71, 113, 161, 275$ exact iterations—matching $O(\eta^{-2})$ scaling—with support size 3–4 and V/V_0 increasing smoothly from 1.506 to 1.836 (far below the $8V_0$ bound). Because each update multiplies denominators by the fixed rational $1/\gamma$, maximum weight denominator bit-length scales linearly with iteration count from 146 to 3577 bits (290 to 7152 bits for exact variance denominators), quantifying Section 8.2's polynomial bit-length guarantee. On two orthogonal regular tetrahedra in $\mathbb{R}^7$ ($L_S^2 = 40$, $h = 4$, $V = 12$), Mode 2 achieves exact rational hull dilation 1.973261 ($\|y_1 - y_2\|^2 / L_S^2 = 12168/3125$) at Giliberti–Hajiaghayi's $a = 1581/1000$ versus the Lemma 5.1 ceiling $\sqrt{3.999122} = 1.999780$ (98.7% of the bound), and 1.736090 vs. 1.741494 near the $\sqrt{3}$ barrier at $a = 71/50$.

Two-Cap Potential and Exact Taylor Thresholds (Fig. 2). Mode 3 (0.181–0.207 ms) independently certifies all Appendix A.1 inequalities verbatim, confirming $F(3/5) > 79/75 > 21/20$ and showing that Taylor degree $K = 4$ ($e > 65/24$) is the *exact minimal degree* certifying second-mask emptiness: $K = 3$ (8/3) yields $0.6618667 < 0.672$, whereas $K = 4$ clears $84/125 = 0.672$ with exact margin $16133/24000 - 84/125 = 1/4800 \approx +0.0002083$, and

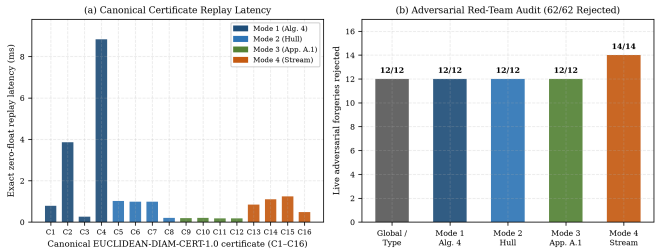

Figure 3: (a) Exact zero-float replay latency across all 16 canonical `EUCLIDEAN-DIAM-CERT-1.0` certificates (0.181–8.834 ms). (b) Red-team adversarial audit rejecting 62/62 live forgery attacks across five attack classes.

Table 2: Summary of all 16 canonical `EUCLIDEAN-DIAM-CERT-1.0` certificates and 62/62 live adversarial red-team audit.

Verification Mode	Certs	Latency (ms)	Attacks Rejected
Mode 1: Alg. 4 Variance Certificate	4	0.263–8.834	12/12
Mode 2: Lemma 5.1 Hull & Tightness	4	0.207–1.014	12/12
Mode 3: App. A.1 Factor-Two Audit	4	0.181–0.207	12/12
Mode 4: Dynamic Stream Chain Replay	4	0.482–1.240	14/14
Global Schema / Type / Rational	–	–	12/12
Total	**16**	**0.181–8.834**	**62/62**

$K = 10$ amplifies the certified slack by 12.9× to $+0.0026775$ while certifying $\epsilon_0 > 1.5 \times 10^{-10}$.

Replay Atlas and Red-Team Audit (Fig. 3, Table 2). All 16 canonical certificates replay in 0.181–8.834 ms with zero floats (Mode 1's 8.834 ms peak reflects 18 exact multi-hundred-bit rational iterations on a 6D simplex). Mode 4 (0.482–1.240 ms) replays dynamic insertion/deletion histories including extreme-point deletion, sample invalidation in a 2-ball hierarchy, key re-insertion, and exact-leaf collapse. Our adversarial suite (`tests/test_suite.py`) tests 62 live forgery attacks—including float/`None`/bool injection, unreduced fractions, oversized rational-string and integer bit-length DoS, tampered Algorithm 4 traces, hull-outside points, weak Taylor degrees, broken scale chains, stale deleted witnesses, unauthenticated invalidated-ball routing, non-chain witness pairs, uncovered residual points, and violated shrinkage—achieving **62/62 rejections** with zero bypasses.

5 Conclusion

We delivered an 18-theorem zero-dependency Lean 4 kernel (5 zero-axiom theorems, 13 `[propext, Quot.sound]`, 0 `Classical.choice`) and the exact-$\mathbb{Q}$ verifier `EUCLIDEAN-DIAM-CERT-1.0` for the variance-hull, Algorithm 4, two-cap mask, dynamic-chain, and Appendix A.1 factor-two core of Giliberti and Hajiaghayi [3]. Mechanizing the Gaussian-grid probabilistic couplings of Sections 6–7 and the communication-complexity space lower bound of Section 10 remain open targets for future formalized streaming geometry.

References

[1] Pankaj K. Agarwal and R. Sharathkumar. Streaming algorithms for extent problems in high dimensions. *Algorithmica*, 72(1):83–98, 2015.

[2] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, pages 625–635, 2021.

[3] Jeff Giliberti and MohammadTaghi Hajiaghayi. On Euclidean diameter in dynamic streams. arXiv:2611.33139 [cs.DS], November 2026.

[4] Piotr Indyk. Better algorithms for high-dimensional proximity problems via asymmetric embeddings. In *Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 539–545, 2003.

[5] Sanjeev Khanna, Ashwin Padaki, Krish Singal, and Erik Waingarten. A polynomial space lower bound for diameter estimation in dynamic streams. In *Proceedings of the 66th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 119–148, 2025.

[6] Ashwin Padaki, Krish Singal, and Erik Waingarten. A near-optimal space lower bound for Euclidean diameter estimation in dynamic streams. arXiv:2609.19089 [cs.DS], 2026.