

Supplementary material to “cert-machine: independent three-valued certification of machine-proposed numerical claims in exact and interval arithmetic”[★]

Carlos Toledo^{a,*}, Pedro Veras Guimarães^b

^a*Independent researcher, the cert-machine project, Brazil*

^b*Department of Mechanical Engineering, Universidade Federal de Santa Catarina, Florianópolis, Brazil*

S1. Proofs of the function lemmas

The lemmas are stated in Appendix A of the main text with the table of functions; each proof below restates what it proves. Throughout, a series is evaluated by the natural extension on the enclosed argument, the tail bound is evaluated in interval arithmetic and its upper end taken, and the tail is added by an outward-rounded addition.

Lemma A.1 (exp). Let x be a double, k an integer, and $R \ni x - k \ln 2$ the interval computed from x , k and the enclosure of $\ln 2$ by the outward-rounding proposition, with $\rho = \max_{r \in R} |r| < 1$ and $-1074 \leq k \leq 1023$ (the call is refused otherwise). Then for every $r \in R$,

$$e^r \in \sum_{i=0}^{24} \frac{r^i}{i!} + [-T, T], \quad T = \frac{\rho^{25}}{25!} \left(1 - \frac{\rho}{26}\right)^{-1},$$

and $e^x = 2^k e^{x-k \ln 2}$ exactly, 2^k being a double.

Proof. Taylor’s theorem gives the tail $\sum_{i \geq 25} r^i / i!$, bounded in magnitude by $\sum_{i \geq 25} \rho^i / i! \leq (\rho^{25} / 25!) \sum_{j \geq 0} (\rho / 26)^j$ since $i! \geq 25! 26^{i-25}$ for $i \geq 25$; the geometric series sums to $(1 - \rho / 26)^{-1}$. The implementation takes k as the nearest integer to $x / \ln 2$ in floating point, which only affects ρ , evaluates

[★]Supplement v0.8 of 2026-10-21, built from repository commit 16849460e608 with the same macro files as the main text. S1 proves the function lemmas stated in the main text’s Appendix A; S2–S4 give the thresholds, the certifiers in detail and the decision loop; S5 is the reproduction table.

*Corresponding author.

Email addresses: `carlos@carlostoledo.co` (Carlos Toledo),
`pedro.veras.guimaraes@ufsc.br` (Pedro Veras Guimarães)

the sum by the natural extension on R and T in interval arithmetic with its upper end, and for a box wider than 10^{-6} encloses by the images of its ends, exp being increasing. $\square$

Lemma A.2 (log). Let $x > 0$ be a normal double, $x = m 2^e$ with $m \in [1, 2)$ (exact), and $V \ni v = (m - 1)/(m + 1)$ the interval computed outward, with $\nu = \max V < 0.34$ (refused otherwise). Then

$$\ln x = e \ln 2 + 2 \operatorname{atanh} v, \quad \operatorname{atanh} v \in \sum_{k=0}^{39} \frac{v^{2k+1}}{2k+1} + [0, T], \quad T = \frac{\nu^{81}}{81(1-\nu^2)}.$$

Proof. $\ln m = 2 \operatorname{atanh} v$ for $m > 0$, and the series of atanh has positive terms for $v \geq 0$, so the tail $\sum_{k \geq 40} v^{2k+1}/(2k+1)$ is nonnegative and at most $(\nu^{81}/81) \sum_{j \geq 0} \nu^{2j}$. A subnormal argument is refused; a box is enclosed by the images of its ends, log being increasing. $\square$

Lemma A.3 (sin, cos). Let x be a double, H the enclosure of $\pi/2$, n an integer and $R \ni x - n\pi/2$ the interval computed outward from x , n and H , with $\rho = \max_{r \in R} |r| \leq 2$. Then $\sin x$ and $\cos x$ are $\pm \sin r$ or $\pm \cos r$ according to $n \bmod 4$, and for $|r| \leq 2$

$$\sin r \in \sum_{j=0}^{15} \frac{(-1)^j r^{2j+1}}{(2j+1)!} + \left[-\frac{\rho^{33}}{33!}, \frac{\rho^{33}}{33!} \right], \quad \cos r \in \sum_{j=0}^{15} \frac{(-1)^j r^{2j}}{(2j)!} + \left[-\frac{\rho^{34}}{34!}, \frac{\rho^{34}}{34!} \right].$$

For a box X of width below 2π , $\sin(X)$ lies in the hull of the images of its ends, extended to 1 (to -1) whenever some $\pi/2 + 2k\pi$ (some $3\pi/2 + 2k\pi$) may lie in X , the test made on the enclosures of these points and answering “may” whenever it cannot separate; a box of width at least 2π is enclosed by $[-1, 1]$. The same holds for $\cos$ with the critical points $2k\pi$ and $\pi + 2k\pi$.

Proof. The quadrant identities are exact. The Lagrange remainder of the Taylor polynomial of degree 31 (resp. 30) is bounded by $\rho^{33}/33!$ (resp. $\rho^{34}/34!$) since every derivative of $\sin$ and $\cos$ is bounded by one in magnitude. On a box, $\sin$ is monotone between consecutive critical points, so its range is the hull of the end values unless a critical point lies inside, where it attains ± 1 ; a test that says “may lie inside” whenever it cannot prove otherwise can only widen the result. The implementation takes n as the nearest integer to x over the midpoint of H , which only affects ρ , and refuses $\rho > \pi/2 + 0.02$. $\square$

Lemma A.4 (Γ by Spouge’s formula). Let $a = 20$ and $w \geq 0$ real. Then, with $c_0 = \sqrt{2\pi}$ and $c_k = (-1)^{k-1}(a-k)^{k-1/2}e^{a-k}/(k-1)!$ for $1 \leq k \leq a-1$,

$$\Gamma(w+1) = (w+a)^{w+1/2} e^{-(w+a)} \left[c_0 + \sum_{k=1}^{a-1} \frac{c_k}{w+k} + \varepsilon_a(w) \right], \quad |\varepsilon_a(w)| \leq a^{-1/2} (2\pi)^{-(a+1/2)}$$

[1]. The implementation encloses each c_k , the sum and the prefactor by the natural extension, with $(w + a)^{w+1/2}$ as $\exp((w + \frac{1}{2}) \log(w + a))$, and widens the bracket by $\bar{\varepsilon}$, the upper end of $10 a^{-1/2} (2\pi)^{-(a+1/2)}$ computed in interval arithmetic from the enclosure of π —a tenfold margin, which cannot cost soundness. For $0.4 < z < 1$ it uses $\Gamma(z) = \Gamma(z + 1)/z$; for a box of width above 10^{-3} it uses that Γ is log-convex on $(0, \infty)$ with its minimum in $[1.4616, 1.4617]$, so the maximum over a box is at an end and the minimum at the interior minimum when the box straddles it, else at the nearer end; a value beyond the doubles is refused.

Lemma A.5 ($\ln \Gamma$, ψ , ψ'). For real $z \geq 20$ and $m = 8$, with B_{2k} the Bernoulli numbers,

$$\begin{aligned} \ln \Gamma(z) &= (z - \tfrac{1}{2}) \ln z - z + \tfrac{1}{2} \ln 2\pi + \sum_{k=1}^m \frac{B_{2k}}{2k(2k-1)z^{2k-1}} + R_0, & |R_0| &\leq \frac{|B_{2m+2}|}{(2m+2)(2m+1)z^{2m+1}}, \\ \psi(z) &= \ln z - \frac{1}{2z} - \sum_{k=1}^m \frac{B_{2k}}{2k z^{2k}} + R_1, & |R_1| &\leq \frac{|B_{2m+2}|}{(2m+2)z^{2m+2}}, \\ \psi'(z) &= \frac{1}{z} + \frac{1}{2z^2} + \sum_{k=1}^m \frac{B_{2k}}{z^{2k+1}} + R_2, & |R_2| &\leq \frac{|B_{2m+2}|}{z^{2m+3}}. \end{aligned}$$

Proof. Binet's second formula [2, §5.9] writes $\ln \Gamma(z) - (z - \frac{1}{2}) \ln z + z - \frac{1}{2} \ln 2\pi$ as $2 \int_0^\infty \arctan(t/z)/(e^{2\pi t} - 1) dt$, and its derivatives give ψ and ψ' with the integrands $t/((t^2 + z^2)(e^{2\pi t} - 1))$ and $t/((t^2 + z^2)^2(e^{2\pi t} - 1))$. Expanding $\arctan(s)$, $1/(1 + s)$ and $1/(1 + s)^2$ to m terms leaves remainders bounded by $s^{2m+1}/(2m+1)$, s^m and $(m+1)s^m$ for every $s \geq 0$, and $\int_0^\infty t^{2k-1}/(e^{2\pi t} - 1) dt = |B_{2k}|/(4k)$; the bounds follow term by term, and for real positive argument the remainder after m terms is bounded by the first omitted term [2, §5.11(ii)]. The implementation lifts an argument below 20 by $\Gamma(z + 1) = z\Gamma(z)$, $\psi(z + 1) = \psi(z) + 1/z$, $\psi'(z + 1) = \psi'(z) - 1/z^2$, each step in interval arithmetic, so every tail is below 10^{-23} and is added as a pad. Over a box, ψ is increasing and ψ' decreasing, so each is enclosed by its ends; $\ln \Gamma$ is convex (Bohr–Møllerup), decreasing to its minimum at $x^* = 1.4616\dots$ and increasing after it, so a box on one side of x^* is enclosed by its ends, and a box across x^* has, for any $p > 0$ and every x in it, $\ln \Gamma(x) \geq \ln \Gamma(p) + \psi(p)(x - p) \geq \ln \Gamma(p) - |\psi(p)| \max(p - \underline{x}, \bar{x} - p)$: a lower end proved from the enclosures at p , which the implementation takes at a double near x^* . $\square$

Lemma A.6 (the regularized incomplete gamma). For $a > 0$ and $z > 0$, $P(a, z) = \frac{z^a e^{-z}}{\Gamma(a+1)} \sum_{n \geq 0} t_n$ with $t_0 = 1$ and $t_{n+1} = t_n z/(a + n + 1)$. If $r_N = z/(a + N + 1) < 1$ then $0 \leq \sum_{n > N} t_n \leq t_N r_N/(1 - r_N)$.

Proof. The ratios $t_{n+1}/t_n = z/(a+n+1)$ decrease in n , so for $n \geq N$ they are at most r_N and the tail is bounded by the geometric series. The implementation stops at the first N with $r_N < 0.95$ and tail below 10^{-18} of the sum (or refuses after 40,000 terms), rescales the running term and sum by 10^{250} when the sum passes it, takes the prefactor in logarithms by Lemma A.1, Lemma A.2, Lemma A.5, and states a value whose logarithm is certainly below -708 as $[0, 10^{-307}]$ rather than underflowing. Over a box, P increases in z and decreases in a , so it is enclosed by two corners; $P^{-1}(a, p)$ is enclosed by a bisection in which an end moves only where the sign of $P(a, z) - p$ is certain, and increases in both a and p . $\square$

Lemma A.7 (Bessel J_ν). For $\nu > -0.6$ and $x > 0$, $J_\nu(x) = \sum_{m \geq 0} t_m$ with $t_0 = (x/2)^\nu / \Gamma(\nu + 1)$ and $t_{m+1} = -t_m (x/2)^2 / ((m+1)(\nu+m+1))$. If $q = (x/2)^2 / ((M+2)(\nu+M+2)) < 1$ then $|\sum_{m > M} t_m| \leq |t_{M+1}| / (1-q)$. For $0 \leq x \leq 1$ a box containing zero is enclosed by $[\max(0, 1 - x^2/4), 1]$ when $\nu = 0$ and by $[0, (x/2)^\nu / \Gamma(\nu + 1)]$ when $\nu > 0$.

Proof. Beyond $M+1$ the magnitude ratios are at most q , so the tail is a geometric series; for the box containing zero, the series alternates with decreasing terms when $x^2 \leq 4(\nu+1)$, which $x \leq 1$ guarantees, so the first two partial sums bracket it. The implementation uses $M = 40$, takes t_0 by Lemma A.4 and the exact ratios, and refuses when $q \geq 1$. For J'_ν the terms carry the factor $(2m+\nu)/2$, whose ratio of consecutive values is $1 + 2/(2m+\nu) \leq 1.026$ for $m \geq M = 40$ and $\nu > -0.6$, so with $q \leq 0.9$ the magnitude ratios are at most $1.026q$ and $1/(1-1.026q) \leq 2/(1-q)$; the implementation uses that bound and refuses $q > 0.9$. $\square$

Lemma A.8 (erf, erfc). For $|z| \leq 2$ and $N = 40$,

$$\operatorname{erf} z = \frac{2}{\sqrt{\pi}} \left[\sum_{k=0}^N \frac{(-1)^k z^{2k+1}}{k! (2k+1)} + T \right], \quad |T| \leq a_{N+1} = \frac{|z|^{2N+3}}{(N+1)! (2N+3)}.$$

For $x \geq 1$, $\sqrt{\pi} e^{x^2} \operatorname{erfc} x = 1/(x + \frac{1}{2}/(x + \frac{3}{2}/(x + \dots)))$, Laplace's continued fraction [2, 7.9.2], whose consecutive convergents bracket its value.

Proof. The magnitudes $a_k = |z|^{2k+1} / (k! (2k+1))$ satisfy $a_{k+1}/a_k = z^2 (2k+1) / ((k+1)(2k+3)) < 1$ for $k \geq 2$ when $|z| \leq 2$, so the alternating series bound applies after $N \geq 2$ terms. The implementation adds $(2/\sqrt{\pi}) a_{N+1}$ as a symmetric pad, with $\sqrt{\pi}$ from the enclosure of π . The continued fraction has positive partial numerators and denominators for $x > 0$, so its even convergents increase, its odd convergents decrease, and the value lies between any two consecutive ones [3, Thm. 1.5.1 and §13.2]; the implementation evaluates the 80th and 81st convergents in interval arithmetic and takes their hull. $\Phi(z) = \frac{1}{2} \operatorname{erfc}(-z/\sqrt{2})$ is pieced from the two: the series for $|z/\sqrt{2}| \leq 2$ and the fraction beyond 1 in magnitude, the overlap decided

by the series; Φ^{-1} is enclosed by a bisection that moves an end only on a certified sign. $\square$

Derivatives and sums. The derivatives of these functions that the certifiers use (∇G in the mean-value form, the Hessians of Proposition 2) are obtained by forward differentiation over the same arithmetic or, for the gamma family, by the series of Lemma A.5 written in the shape parameter so that nothing is divided by it near the lognormal limit (S3). Long sums are enclosed by Sum2 with the bound of the Sum2 bound.

S2. Thresholds

Proof of the threshold proposition (main text, Proposition 4). Soundness gives $T(m)$ in the first case and $F(m)$ in the second, and the monotone-truth assumption extends them to the half-lines. In the third case the two proved witnesses show that neither $T(m)$ nor $F(m)$ holds; by the monotone-truth assumption, $T(p)$ fails for every $p \geq m$ and $F(p)$ fails for every $p \leq m$, so $p_T \leq m \leq p_F$. $\square$

An operating choice. Where p enters additively, $g(x; p) = p + h(x)$ on a fixed box, no bisection is needed: an enclosure $[\underline{h}, \bar{h}]$ of h over B gives $T(p)$ for every $p \leq c - \bar{h}$ and $F(p)$ for every $p > c - \underline{h}$, and the gate reports these two values rounded outward to the precision printed (a tenth of a bar for the pump pressure), every step in directed rounding: the subtractions stepped outward, the product by ten stepped outward, then the floor of the lower bound and the ceiling of the upper one. The gap between them is the width of the enclosure.

A measurement that narrows the box. Let p be the lower end of the declared interval of one input, so that $B(p') \subseteq B(p)$ whenever $p \leq p'$: raising p is what a measurement showing “the roughness is at least p ” would do. Here the half of the monotone-truth assumption for T fails: as the box shrinks with p , truth propagates upward, not downward, and the threshold is an infimum, resting on soundness and nesting alone.

Proposition 1 (nested boxes). *Let the hypothesis of the soundness proposition hold. If the claim is certified over $B(p_0)$, then $T(p)$ holds for every $p \geq p_0$, and $p^* = \inf\{p : T(p)\} \leq p_0$. The same holds for an epistemic refutation and F ; a universal refutation is not kept, because its counterexample may lie outside the smaller box.*

Proof. By the soundness proposition, φ holds at every point of $B(p_0)$, and $B(p) \subseteq B(p_0)$ for $p \geq p_0$. For an epistemic refutation the argument is the same with $\neg\varphi$. $\square$

If the claim is certified at the declared upper end, the gate bisects for sixteen steps over the two-valued question “is the claim certified over $B(m)$?”: a certified midpoint lowers the upper end of the search, and any other outcome—a refutation, a split or an unresolved refusal—raises its lower end, so the search never stalls. The upper end, rounded up to the precision printed, is a proved sufficient measurement by Proposition 1; the lower end proves nothing, since the gate decides each $B(m)$ afresh and a midpoint can fail through the depth budget alone. The rock-physics engine bisects for ten steps, one-sidedly, on whether a proved violating witness is found as the parameter moves, and reports the end at which one was proved. A two-sided bracket of p_T and p_F by bisection in the operating-choice case is not implemented and not described.

How each instance verifies the hypotheses. The flow-line gate’s box is fixed by declaration and its rule is $P_d + h(x) \leq c$ with h the friction and hydrostatic terms, which do not depend on the pump pressure, so the monotone-truth assumption holds by construction and the closed form applies; its roughness and viscosity thresholds raise a lower end, so the boxes are nested by construction and Proposition 1 applies with no monotonicity needed. The rock-physics engine moves one declared end and decides each box afresh, nested by construction. The likelihood certifier and the universal branch-and-bound declare no parameter and report no threshold.

S3. The certifiers in detail

A zero, exactly one.

Theorem 1 (Krawczyk [4, 5, 6]). *Let $f : D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuously differentiable, $X \subseteq D$ a box, $\hat{x} \in X$, Y a nonsingular matrix and $F'(X)$ an interval extension of the Jacobian over X . If $K(X) = \hat{x} - Yf(\hat{x}) + (I - YF'(X))(X - \hat{x}) \subseteq \text{int } X$, then f has exactly one zero in X .*

The certifier takes $\hat{x}$ from a floating-point search and Y as an approximate inverse of the Jacobian at $\hat{x}$, evaluates $K(X)$ with every datum enclosed by its neighbouring doubles, and inflates X for at most a fixed number of rounds (twelve for a likelihood) and below a radius cap; if the inclusion never holds the zero is refused (“no contraction”). A sibling certifier for fixed points in weighted sequence spaces uses the radii-polynomial form of the same argument [7] and requires the contraction factor to be proved below one.

A maximum in a box.

Proposition 2. *Let ℓ be twice continuously differentiable on a convex box X , let $\nabla \ell$ have exactly one zero θ^* in X , and let the Hessian of ℓ be negative definite at every point of X . Then θ^* is the unique maximiser of ℓ on X .*

Proof. A function whose Hessian is negative definite on a convex set is strictly concave there, and a critical point of a strictly concave function is its unique maximiser on the set. $\square$

Negative definiteness over the box is proved either by Sylvester’s criterion on the interval Hessian $\mathbf{H}(X)$, i.e. by enclosures of the leading principal minors satisfying $(-1)^k \det \mathbf{H}_{1:k,1:k}(X) > 0$ for $k = 1, \dots, n$, for up to three parameters, with the two enclosures of each symmetric entry intersected; or, where the minors are too wide, at the candidate point together with a proof that every symmetric Hessian over X is nonsingular ($|I - A\mathbf{H}(X)|_\rho < \rho$ componentwise, with $\rho > 0$ the half-widths of X and A the approximate inverse used in the Krawczyk step), so that, X being connected and the eigenvalues of symmetric matrices continuous, no Hessian over X can change inertia. The statement is local: that no better maximum exists elsewhere remains the search’s claim, and every fit is stated with that scope.

A supremum at a limit.

Proposition 3. *Let $\ell(\mu, \sigma, Q)$ be continuous on $B \times [0, q_1]$ and differentiable in Q on $B \times (0, q_1]$, with $\partial\ell/\partial Q < 0$ there. Then $\ell(\mu, \sigma, Q) < \ell(\mu, \sigma, 0)$ for every $(\mu, \sigma) \in B$ and $Q \in (0, q_1]$.*

The proof is the mean value theorem. In Prentice’s form of the generalized gamma [8], with $Q = 1/\sqrt{\alpha} > 0$ for the shape α , the family reaches the lognormal at $Q = 0$, an ordinary point in these coordinates, and the limit is approached from $Q > 0$ only; a proof that the Q -score is negative over a neighbourhood of the lognormal fit refuses the family at its edge *with the proof*, and the neighbourhood is recorded because the conclusion is local. For $Q \leq 0.2$ the log-likelihood is written through the entire functions $(e^t - 1)/t$ and $(e^t - 1 - t)/t^2$ of $t = Qw$ and Binet’s series for the gamma-function part (Lemma A.5) as polynomials in Q , so nothing is divided by Q , ℓ is smooth at $Q = 0$ and equals the lognormal log-likelihood there; each series is truncated with its tail bound, $|t|^J e^{|t|}/(J+1)!$ and the first omitted Bernoulli term, added outward.

A statement for every input in a box. Three reductions are used. *Bisection with a proposer:* the box is split until each piece’s enclosure decides, a floating-point heuristic choosing the coordinate and never a verdict, under a depth budget (the termination proposition, Algorithm 1); the engineering gate is the main instance. *Monotone implicit functions:* where a model quantity is the root of an equation monotone in its parameters, its range over the box is bracketed by the roots at two corners, each counted only when the sign of the equation just beyond it is proved; for Colebrook’s friction law, $h(x) = x + (2/\ln 10) \ln(a + bx)$ is increasing in x , a and b , so the root decreases in a and b and every root over the box lies in $[x^*(a_+, b_+), x^*(a_-, b_-)]$. *Multilinear forms:*

Lemma 1. *Let f be multilinear on a box $X \subset \mathbb{R}^n$. Then $f(X) = [\min_v f(v), \max_v f(v)]$, the minimum and maximum taken over the 2^n vertices v of X .*

Proof. f is affine in each coordinate with the others fixed, so its extremes over an interval of one coordinate are at that coordinate's ends; induction over the coordinates places both extremes at vertices, and X is connected, so every value between is attained. $\square$

The kernel checks multilinearity syntactically and evaluates the corners in exact rationals; the enclosure is the exact range ($q = 0$ in the convergence-order inequality), and the two extreme corners are the witnesses of a split refusal.

A sign, exactly. Orientation predicates $\text{orient}(a, b, c) = \text{sign det} \begin{pmatrix} b_x - a_x & b_y - a_y \\ c_x - a_x & c_y - a_y \end{pmatrix}$ are evaluated first in double precision and trusted only when the computed magnitude clears a forward-error bound proved for the inputs [9, Thm. 2.2 and §3.1]: with every coordinate a decimal literal of magnitude at most $M = 128$ and its double within relative error $u = 2^{-53}$ of it, the computed determinant is within $48uM^2 \approx 8.7 \times 10^{-11}$ of the determinant of the literals, the conversion included, up to terms of order u^2M^2 that the threshold absorbs; a computed magnitude above 10^{-10} fixes the sign, and inputs beyond M go to integers directly. Below the threshold the sign is recomputed in integers, each decimal literal scaled to a fixed sixty-four places; a literal needing more is refused. A point on an edge is reported ON, an answer a floating-point test does not have [10].

A polynomial, nonnegative. A claim $p \geq \gamma$ is certified by an identity $p(x) - \gamma = \sum_i c_i q_i(x)^2$ with $c_i \in \mathbb{Q}_{\geq 0}$ and $q_i \in \mathbb{Q}[x]$, checked by exact coefficient comparison; finding the decomposition is a semidefinite program [11, 12] outside the trust path. A separate refuter evaluates p at an exact witness w and refutes if $p(w) < \gamma$. The checker has so far been exercised on worked examples and red controls only; no semidefinite solver is yet wired to it as a proposer.

A minimum of a cosine polynomial. Since $\cos(ax) = T_a(\cos x)$ for the Chebyshev polynomial T_a , $\min_{x \in \mathbb{R}} \sum_{a \in A} \cos(ax) = \min_{y \in [-1, 1]} P_A(y)$ with $P_A = \sum_{a \in A} T_a$. The certifier isolates the critical points with a Sturm chain of the square-free part of P'_A on $[-1025/1024, 1025/1024]$, refines each by interval Newton on 192-bit dyadic balls around exact Horner values (with bisection as fallback), encloses P_A on each critical interval using the bound $|T''_a| \leq a^2(a^2 - 1)/3$ on $[-1, 1]$ (V. Markov), adds the endpoints ± 1 exactly and returns $[\min_i \underline{v}_i, \min_i \bar{v}_i]$ over the candidate enclosures, rounded outward. A stronger certificate identifies the minimum as an algebraic number—both endpoint values above the enclosure, so the minimum is a

critical value; an exact sign change of the resultant across the enclosure and a Sturm count of one; irreducibility modulo a prime at which the degree does not drop—and is certified only when all of these hold.

A formal proof, built and read. A formal certificate is two claims. That the proof is accepted is decided by building it at the pinned commit, with the toolchain and library versions the claimant pinned, and asking the kernel which axioms the result depends on; where the claimant’s checker allows it, the proof is replayed by a second kernel written independently of the first [13, 14]. That the statement proved is the one claimed is not a computation: the statement is read hypothesis by hypothesis against the published theorem and the library’s definitions, the reading is recorded with its slips and weak points, and it never issues a verdict of its own—the verdict is the kernel’s, scoped by what the reading found absent.

S4. The decision loop

In the gate the epistemic refutation is decided by a second pass over the box only when no satisfying corner was proved; the universal branch-and-bound keeps no corners and refutes nothing, since its claims are bounds that either certify or exhaust the budget.

Algorithm 1 Deciding a claim (B, φ) with $\varphi \equiv \bigwedge_j g_j(x) \leq c_j$ (sketch).

Require: box B , interval extensions G_j of g_j , reading $\in \{\text{universal, epistemic}\}$, budget $d_{\max}$

- 1: pin every input by SHA-256 and refuse on mismatch; record the digests of the certifying modules
- 2: decide the 2^n corners of B (each the declared decimal enclosed) or proposer points as thin boxes; keep one proved satisfying point as x^+ and one proved violating point as x^- , or $\perp$ $\triangleright$ the gate; the branch-and-bound has no corner step
- 3: **if** universal and $x^- \neq \perp$ **then return** REFUTED with witness x^-
- 4: **end if**
- 5: $\mathcal{Q} \leftarrow \{B\}; \mathcal{X} \leftarrow \emptyset$
- 6: **while** $\mathcal{Q} \neq \emptyset$ **do**
- 7: pop X from $\mathcal{Q}$
- 8: **if** $\sup G_j(X) \leq c_j$ for all j **then**
- 9: $\mathcal{X} \leftarrow \mathcal{X} \cup \{X\}$
- 10: **else if** $\inf G_j(X) > c_j$ for some j **then**
- 11: **if** universal **then return** REFUTED with witness X
- 12: **end if**
- 13: count X as violating $\triangleright$ epistemic: φ false on all of X
- 14: **else if** $\text{depth}(X) < d_{\max}$ **then**
- 15: split X (heuristic choice of coordinate; never a verdict); push the halves
- 16: **else**
- 17: mark X unresolved
- 18: **end if**
- 19: **end while**
- 20: **if** $x^- = \perp$ and $\mathcal{X}$ covers B **then return** CERTIFIED with certificate (digests, hull of enclosures, counts, falsifier)
- 21: **end if**
- 22: **if** epistemic and every box was refuted **then return** REFUTED
- 23: **end if**
- 24: **if** epistemic and $x^+ \neq \perp$ and $x^- \neq \perp$ **then return** REFUSED (split) with witnesses x^+, x^- , and thresholds if a parameter is declared (the main text)
- 25: **end if**
- 26: **return** REFUSED (unresolved) with its reason

S5. Reproduction

Each headline result re-derives with the command beside it in Table 1, on the laptop of the benchmark of the main text unless the row says otherwise; each command that fetches a claimant’s file refuses it unless it matches the pinned digest, and each battery fails unless its red controls fire.

result	command	expected	time
the benchmark, re-derived	<code>node tools/run-selfaudit-mutations.js --check</code>	the campaign’s totals and every mutant’s outcome identical to the record	17 min
the records under two profiles	<code>node tools/run-selfaudit.js --check</code>	every record identical to its pin outside stamps	26 min
the concordance	<code>node tools/run-selfaudit-concord.js --check</code>	91 comparisons, none in disagreement	seconds
the kernel against Arb	<code>node tools/audit-kernel-arb.js --check</code>	no violation; every malformed input refused; both planted kernels caught	211 s for the record
planted claims	<code>node tools/run-method-eval.js --check</code>	every false claim refuted, every true one certified, every malformed one refused	under a second
the register as of 2026-10-21 (132 rows)	the rerun kit, RERUN.md	0 rows moved, 0 failed	27 min
kissing, $d = 25$, 27	<code>python3 tools/run-kissing-wave.py --only qiushi-25,qiushi-27</code>	CERTIFIED; the constructions rebuilt from the repository’s files	as recorded
kissing, $d = 32$ –39	<code>python3 tools/run-kissing-configs.py <d></code>	CERTIFIED; the pairs and contacts of the record	as recorded
kissing, $d = 49$ –55	<code>tools/runpod-qiushi-lifting.sh</code> , on a CUDA machine	the calibration on 32 and 39 first, then CERTIFIED	6.8 GPU-hours
the registry’s certificates	<code>python3 tools/verify_sumdiff.py;</code> <code>node tools/verify_fei.js</code>	CERTIFIED; their red controls refuted or refused	seconds
the registry’s queue	<code>python3 -I tools/run-optq-ledger.py</code>	7 CERTIFIED by computation, 1 REFUSED	161 s
Navier–Stokes	the build that corpus/navier-stokes/build.json records	both challenge theorems accepted; only the standard axioms	140 min
the release’s challenges	<code>.github/workflows/openai-math-kernel.yml</code>	398 accepted by both kernels, 18 refused for resources	hosted runners
the engineering gate	<code>node apps/contraprova/gate/battery.js</code>	24 checks, 9 red controls fired, the reference inside the enclosures at 176 inputs	seconds
the likelihood certifier	<code>node instruments/hseva/battery.js</code>	1,678 checks; the atlas and the ledger written by the code on disk	minutes

Table 1: Each headline result, the command that re-derives it, what it must print, and what it took (the times are those the records hold).

References

- [1] J. L. Spouge, Computation of the gamma, digamma, and trigamma functions, *SIAM Journal on Numerical Analysis* 31 (3) (1994) 931–944. doi:10.1137/0731050.
- [2] NIST, NIST digital library of mathematical functions, Release 1.2.4 of 2025-03-15, F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds. (2025). URL <https://dlmf.nist.gov/>
- [3] A. Cuyt, V. B. Petersen, B. Verdonk, H. Waadeland, W. B. Jones, *Handbook of Continued Fractions for Special Functions*, Springer, 2008. doi:10.1007/978-1-4020-6949-9.
- [4] R. Krawczyk, Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehler-schranken, *Computing* 4 (1969) 187–201. doi:10.1007/BF02234767.
- [5] A. Neumaier, *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990. doi:10.1017/CB09780511526473.
- [6] S. M. Rump, Verification methods: rigorous results using floating-point arithmetic, *Acta Numerica* 19 (2010) 287–449. doi:10.1017/S096249291000005X.
- [7] J. B. van den Berg, J.-P. Lessard, Rigorous numerics in dynamics, *Notices of the American Mathematical Society* 62 (9) (2015) 1057–1061. doi:10.1090/noti1276.
- [8] R. L. Prentice, A log gamma model and its maximum likelihood estimation, *Biometrika* 61 (3) (1974) 539–544. doi:10.1093/biomet/61.3.539.
- [9] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd Edition, SIAM, Philadelphia, 2002. doi:10.1137/1.9780898718027.
- [10] C. K. Yap, Towards exact geometric computation, *Computational Geometry* 7 (1–2) (1997) 3–23. doi:10.1016/0925-7721(95)00040-2.
- [11] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Mathematical Programming* 96 (2) (2003) 293–320. doi:10.1007/s10107-003-0387-5.
- [12] H. Peyrl, P. A. Parrilo, Computing sum of squares decompositions with rational coefficients, *Theoretical Computer Science* 409 (2) (2008) 269–281. doi:10.1016/j.tcs.2008.09.025.
- [13] Lean FRO, Comparator, Commit d03acab1 (2026). URL <https://github.com/leanprover/comparator>
- [14] ammkrn, nanoda_lib: a type checker for Lean 4 export files, Commit 3a240721, with the patch of issue #44 (2026). URL https://github.com/ammkrn/nanoda_lib