CERTINR: CERTIFIED VERIFICATION AND COMPRESSION OF IMPLICIT NEURAL FIELDS

Anonymous authors
Paper under double-blind review

ABSTRACT

Implicit neural representations (INRs) use neural networks to represent continuous functions over spatial or spatiotemporal coordinates. After compression, the decoded network can differ from the reference in ways that evaluations at sampled coordinates may miss. We introduce CertINR, a framework with two coupled components: a verifier that provides provable upper bounds on the maximum reference–decoded error over the continuous domain, and a certificate-aware compressor that uses these bounds to repair and select compressed representations. The key verification idea is to analyze the reference and decoded networks jointly at the same coordinates, so that variation shared by the two networks cancels before their difference is bounded. We realize this idea through paired Taylor bounds that preserve this cancellation through nonlinear layers and admit refinement-based certification guarantees. Alongside paired Taylor, we develop complementary global spectral and local residual/Jacobian certificates and combine all three in a verification portfolio. On the compression side, CertINR searches a pool of compressed candidates, uses the local error models to repair the final affine layer, verifies the actual decoded streams, and returns the smallest representation that satisfies the target error tolerance. Experiments on analytic and pretrained sinusoidal fields show that no single certificate dominates across network regimes: paired Taylor is strongest on shallow analytic models, whereas spectral bounds are stronger on larger pretrained models. The CertINR bounds also resolve cases left uncertified by general-purpose neural-network verifiers. On six pretrained signed distance fields (SDFs), CertINR achieves 41.21% mean per-shape storage savings at tolerance 0.01. More generally, we show that for SDFs, a certified uniform error preserves inside–outside decisions and, under explicit gradient conditions, can also guarantee topology preservation and bound surface displacement.

1 INTRODUCTION

Implicit neural representations (INRs) represent continuous functions using neural networks that take coordinates as input. Sinusoidal representation networks (SIRENs), for example, are well suited to representing functions with fine spatial detail (Sitzmann et al., 2020). Compressing INR parameters can substantially reduce storage (Strümpler et al., 2022; Schwarz et al., 2023; Jayasundara et al., 2025), but the decoded network may differ from the reference at locations that are not covered by sampled error measurements. Such localized differences matter when the field value itself determines a downstream decision. Signed distance fields (SDFs) provide an important example: negative values indicate the interior of a shape, positive values the exterior, and the zero values define its surface, or *zero set*, so an undetected compression error can reverse an inside–outside decision or alter the represented geometry. This creates a gap between achieving good sampled reconstruction after compression and knowing that the compressed field remains uniformly close to the reference throughout the domain.

We introduce **CertINR**, which couples continuous-domain verification with certificate-aware compression. Its verifier computes a provable upper bound on the maximum difference between a reference INR and its decoded compressed counterpart over a domain Ω, rather than estimating this difference only at sampled coordinates. Its compressor uses these certificates to adjust the output head of compressed candidates when beneficial, verify the stored and decoded networks, and select the smallest candidate that satisfies a prescribed error tolerance. For SDFs, we show that a uniform error

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

certificate preserves the reference inside–outside decision wherever the reference field has sufficient margin; near the zero set, where value error alone does not ensure geometric preservation, we derive additional gradient conditions that certify topology preservation and bound surface displacement. We make three contributions:

Certified compression objective. We formulate INR compression as minimizing the storage cost of a compressed representation subject to a provable bound on the maximum reference–decoded error over the entire continuous domain, with verification applied to the actual stored and decoded parameters (Section 3).

Complementary continuous-domain certificates. We develop three verification bounds for this objective: a global spectral bound, a local residual/Jacobian bound, and a paired Taylor bound that analyzes the reference and decoded networks jointly so that shared variation cancels before their difference is bounded. We characterize the paired Taylor remainder and establish refinement conditions under which candidates with true error below tolerance can eventually be certified (Section 4).

Certificate-aware compression. We integrate these certificates into a compressor that searches over compressed candidates, adjusts the final affine layer when beneficial, verifies the resulting stored and decoded networks, and returns the lowest-storage candidate meeting the target error tolerance. We also show how the resulting uniform error guarantees translate into geometric guarantees for signed distance fields (Sections 5 and 6).

2 Related work

Existing INR compression methods reduce storage through quantization, retraining, and entropy coding (Strümpler et al., 2022), sparse weight-vector coding (Jayasundara et al., 2025), or rate–distortion optimization across data types (Schwarz et al., 2023). These methods primarily evaluate reconstruction quality at sampled coordinates and do not provide a continuous-domain certificate on the error introduced by compression relative to the reference network. Appendix B gives a smooth example in which mean squared error vanishes while a sign error with fixed reference margin persists. CertINR addresses this complementary problem: existing INR codecs can supply compressed candidates, while CertINR verifies their decoded representations over the continuous domain, optionally adjusts the final affine layer, and selects the lowest-storage candidate that satisfies a prescribed uniform error tolerance.

Our verification problem compares two networks evaluated at the same input with different parameters, connecting CertINR to weight-perturbation certification (Weng et al., 2020) and differential or relational verification. ReluDiff and NeuroDiff track differences between network executions rather than analyzing them independently (Paulsen et al., 2020a;b), while DiffRNN extends this to recurrent networks with smooth activations (Mohammadinejad et al., 2021); input-relational verification instead compares outputs across related inputs (Banerjee et al., 2024). For quantized networks, QEBVerif bounds quantization errors (Zhang et al., 2023), Quadapter selects property-preserving quantization schemes (Zhang et al., 2025), and QNNRepair repairs networks over discrete datasets using mixed-integer linear programming (Song et al., 2023). General-purpose verification based on convex relaxations and branch-and-bound provides another route to bound outputs over continuous regions (Shi et al., 2025). CertINR instead focuses on the continuous-domain difference between a fixed reference INR and its decoded counterpart, exploiting their shared inputs to directly certify compression error throughout the domain.

Continuous-domain bounds have also been used to reason about neural-field geometry. Sharp & Jacobson (2022) use range analysis and affine arithmetic to bound outputs over spatial regions, while others constrain field variation via Lipschitz bounds (Coiffier & Béthune, 2024; McGinnis et al., 2026), derive reachability-based safety constraints (Michaux et al., 2023), or impose topological constraints during training (Jignasu et al., 2024). These works primarily analyze or constrain the field itself; we instead focus on what is preserved when a fixed, trained field is compressed. Uniform value bounds also connect to classical persistence stability, which controls changes in persistence diagrams across value thresholds (Cohen-Steiner et al., 2007). Our geometric result targets the represented zero set directly: for signed distance fields, uniform value control and explicit gradient conditions guarantee a continuous deformation between reference and decoded surfaces, bounding their displacement.

3 PROBLEM SETUP

Each compressed representation approximates a fixed scalar reference field $f_\theta : \Omega \to \mathbb{R}$ with parameters $\theta \in \mathbb{R}^P$, where P is the parameter count. The verification domain $\Omega \subset \mathbb{R}^d$ is a nonempty compact box or a finite union of such boxes, and d is the coordinate dimension. For sine networks, θ contains the frequency-scaled coefficients defined in Section 4.1.

Each compressed candidate is represented by a finite binary string $s \in \{0,1\}^*$, from which the decoder $\mathcal{D}$ reconstructs parameters $\hat{\theta} = \mathcal{D}(s)$ with the same architecture. We define its storage cost as $R(s) := |s|$ bits. We call the decoded field minus the reference the *compression residual* g_s and define it and its maximum absolute error δ, respectively, as:

$$g_s(x) := f_{\mathcal{D}(s)}(x) - f_\theta(x), \quad \delta(s;\Omega) := \sup_{x\in\Omega} |g_s(x)|. \tag{1}$$

Because the maximum error $\delta(s;\Omega)$ is difficult to compute, we derive a computable upper bound $U(s;\Omega)$ satisfying $\delta(s;\Omega) \le U(s;\Omega)$. This *error certificate* guarantees error at most ε throughout Ω whenever $U(s;\Omega) \le \varepsilon$. A larger bound leaves acceptance unresolved; it does not prove a violation. For tolerance $\varepsilon \ge 0$ and candidate family $\mathcal{C}$, we minimize certified storage:

$$R_{\text{cert}}(\varepsilon;\Omega) = \inf_{s\in\mathcal{C}} \{R(s) : U(s;\Omega) \le \varepsilon\}, \tag{2}$$

which is $+\infty$ when no candidate is certified. Appendix A summarizes notation and scope.

4 ERROR BOUNDS

We construct the upper bound in Equation (2) by tracking the compression residual through the network. To obtain a guarantee over the continuous domain, we partition Ω into a finite collection $\mathcal{P}$ of axis-aligned boxes, and bound the residual over each box. We call each box $B \in \mathcal{P}$ a spatial cell. A cell is specified by its center $c \in \mathbb{R}^d$ and half-widths $r \in \mathbb{R}^d_+$, $B = \left\{x \in \mathbb{R}^d : |x_i - c_i| \le r_i,\ i = 1,\dots,d\right\}$. CertINR uses three complementary upper bounds. A global spectral bound controls error propagation through layer operator norms without subdividing the domain. A local residual/Jacobian bound combines the compression residual at a cell center with bounds on how rapidly that residual can vary within the cell. Finally, our paired Taylor bound analyzes the reference and decoded networks jointly to retain cancellation in both their linear and nonlinear variation. We take the smallest valid bound available on each cell. Appendix H gives the certificate-arithmetic details.

4.1 LAYERWISE DIFFERENCES

Matching architectures let us compare corresponding hidden features. We first derive the bounds for SIRENs, which use sinusoidal activations with fixed layer frequencies. The network has L hidden layers, with $n_0 = d$ input coordinates. Layer $\ell = 1,\dots,L$ has width n_ℓ, weights $W_\ell \in \mathbb{R}^{n_\ell \times n_{\ell-1}}$, biases $b_\ell \in \mathbb{R}^{n_\ell}$, and fixed sine frequency $\omega_\ell > 0$. We include each frequency in the effective coefficients $A_\ell = \omega_\ell W_\ell$ and $a_\ell = \omega_\ell b_\ell$. The final affine head has weights $A_{L+1} \in \mathbb{R}^{1\times n_L}$ and scalar bias a_{L+1}. Equation (3) defines the preactivations z_ℓ and hidden features h_ℓ, applying sine componentwise:

$$h_0 = x, \quad z_\ell = A_\ell h_{\ell-1} + a_\ell, \quad h_\ell = \sin z_\ell, \quad f_\theta = A_{L+1}h_L + a_{L+1}. \tag{3}$$

Frequencies remain fixed, with $\omega_\ell = 1$ when storage already includes their scaling. The coefficient changes $\Delta A_\ell = \widehat{A}_\ell - A_\ell$ and $\Delta a_\ell = \widehat{a}_\ell - a_\ell$ therefore capture the complete difference between the reference and decoded parameters, with hats denoting decoded quantities, including reconstruction and frequency-scaling roundoff.

These coefficient changes induce preactivation differences $\Delta z_\ell = \widehat{z}_\ell - z_\ell$ and activation differences $e_\ell = \widehat{h}_\ell - h_\ell$. Subtracting corresponding computations gives exact recurrences for both:

$$\Delta z_\ell = \widehat{A}_\ell e_{\ell-1} + \Delta A_\ell h_{\ell-1} + \Delta a_\ell, \tag{4}$$

$$e_\ell = 2\sin(\Delta z_\ell/2) \odot \cos(z_\ell + \Delta z_\ell/2), \tag{5}$$

where $\odot$ denotes componentwise multiplication, and $e_0 = 0$. Equation (4) separates propagated activation error from new weight and bias errors. Equation (5) retains the factor $\sin(\Delta z_\ell/2)$, which vanishes with the preactivation difference even when the reference preactivation varies widely.

The same decomposition gives the output residual as $g_s = \widehat{A}_{L+1}e_L + \Delta A_{L+1}h_L + \Delta a_{L+1}$. To turn this into a cell certificate, we must bound how the residual varies across the cell. The residual's spatial derivatives provide this control.

4.2 Residual derivatives

We differentiate the residual identities before bounding them, retaining the factors that vanish with the parameter change (Appendix C.2). Each hidden-layer Jacobian records derivatives of its features with respect to input coordinates; the scalar output Jacobian is the gradient's transpose.

For a cell $B = \prod_{i=1}^{d}[c_i - r_i, c_i + r_i]$, let $c \in \mathbb{R}^d$ be its center and $r = (r_1, \ldots, r_d)^\top$ its half-width vector. Intervals $G_B \supseteq g_s(B)$ and $G_c \ni g_s(c)$ contain the residual range and center value, respectively. With derivative bounds $M_i(B) \geq \sup_{x\in B} |\partial_i g_s(x)|$, we take the smaller of a direct range bound and a center-based variation bound:

$$U_{\mathrm{RJ}}(B) = \min\Big\{ \mathrm{mag}(G_B), \mathrm{mag}(G_c) + \sum\nolimits_{i=1}^{d} r_i M_i(B) \Big\}, \tag{6}$$

where $\mathrm{mag}([\underline{u}, \overline{u}]) = \max(|\underline{u}|, |\overline{u}|)$ is the largest absolute value in an interval. The residual/Jacobian certificate in Equation (6) combines these two bounds. Theorem 4 establishes soundness provided the interval arithmetic is inclusion-preserving; in finite precision, this is enforced by rounding interval endpoints outward.

We next consider when subdivision makes the certificate progressively tighter. Assume that the network architecture and sine frequencies are fixed, the input domain is compact, and the parameters lie in a bounded set. In exact arithmetic, the residual and derivative bounds are $O(\rho)$ as $\rho = \|\widehat{\theta} - \theta\|_\infty \to 0$, where ρ is the largest change in an effective coefficient.

Global and local bounds. We first consider two certificates that do not use the paired Taylor construction. The spectral bound is global: it propagates parameter perturbations through the network using spectral norms of the weight matrices, producing a single bound U_{global} over Ω without spatial subdivision. The residual/Jacobian bound in Equation 6 is local and can tighten as the domain is subdivided. Appendix C gives the corresponding recurrences and soundness arguments.

4.3 Paired Taylor bounds

To see this cancellation, consider $f(c+y) = a + by$ and $\widehat{f}(c+y) = a + \Delta a + (b + \Delta b)y$ on $|y| \leq r$. Independent interval subtraction gives $|\Delta a| + (|b| + |b + \Delta b|)r$, counting the shared slope twice. Subtracting first cancels this slope, giving the exact maximum error $|\Delta a| + |\Delta b|r$.

We extend this cancellation to nonlinear layers using local coordinates $y = x - c$. Omitting layer indices, let $\zeta = \Delta z$ denote the preactivation difference. We write both the reference preactivation and its difference as a center value, a linear term, and a nonlinear remainder:

$$z = z_c + L_z y + \xi_z, \qquad \zeta = \zeta_c + L_\zeta y + \xi_\zeta, \qquad |\xi_z| \leq R_z, \quad |\xi_\zeta| \leq R_\zeta.$$

Here, $z_c = z(c)$ and $\zeta_c = \zeta(c)$ are the center values; the Jacobians L_z and L_ζ have one row per neuron and d columns. Using the same y retains their shared linear variation. The nonnegative remainder bounds R_z and R_ζ give variation bounds $v = |L_z|r + R_z$ for the reference and $w = |L_\zeta|r + R_\zeta$ for the difference over $|y_i| \leq r_i$, with absolute values taken componentwise. Passing these models through sine introduces a nonlinear remainder. To retain cancellation in its bound, we analyze the activation difference $q(z, \zeta) = \sin(z + \zeta) - \sin z$, whose second derivatives satisfy

$$|q_{zz}| \leq \min\{2, |\zeta|\}, \qquad |q_{z\zeta}| \leq 1, \qquad |q_{\zeta\zeta}| \leq 1.$$

Since $|\zeta| \leq |\zeta_c| + w$, the componentwise bound $\kappa_q = \min\{2, |\zeta_c| + w\}$ controls $|q_{zz}|$ on the cell and along the Taylor expansion segments. The center derivatives are $\widehat{c}_c = \cos(z_c + \zeta_c)$ and $\Delta c_c = \cos(z_c + \zeta_c) - \cos z_c$. Then, Taylor's theorem gives the exact linear coefficient L_e and remainder bound R_e for the activation difference:

$$L_e = \widehat{c}_c L_\zeta + \Delta c_c L_z, \quad R_e := |\widehat{c}_c| R_\zeta + |\Delta c_c| R_z + \tfrac{1}{2}\kappa_q v^2 + vw + \tfrac{1}{2}w^2, \tag{7}$$

where activation vectors scale Jacobian rows; other products and powers are componentwise. The reference activation has coefficient $L_h = \cos(z_c)L_z$ and remainder bound $R_h := |\cos(z_c)|R_z + v^2/2$. Absolute weight matrices and their perturbations propagate these bounds through the next affine layer.

Equation (7) retains the compression perturbation through κ_q and w, while the reference variation v can amplify the remainder. Smaller spatial cells reduce this variation. For a fixed stream, we write $g = g_s$ and quantify this dependence in Proposition 1.

Proposition 1 (Paired Taylor remainder)**.** *Fix the architecture and frequencies, and let the input domain be compact. With parameters restricted to a bounded neighborhood, the exact-arithmetic output remainder radius satisfies $R_g(B) = O(\rho\|r\|_1^2)$ as $\rho, \|r\|_1 \to 0$. We therefore obtain*

$$|g(x)| \le |g(c)| + \sum_{i=1}^{d} |\partial_i g(c)| r_i + R_g(B), \quad \forall x \in B. \tag{8}$$

Proof. See Appendix D.1. □

Proposition 1 shows that the nonlinear bound shrinks locally with both parameter change and squared cell radius. Adding it to the affine bound gives the Taylor certificate U_T. We take the smallest valid upper bound on each cell from a *certificate portfolio*: $U(B) = \min\{U_{\mathrm{RJ}}(B), U_T(B), U_{\mathrm{global}}\}$. Appendix D.5 explains how we include numerical uncertainty in the local models.

4.4 Domain certification

Given a partition $\mathcal{P}$ of Ω, the domain-wide certificate is

$$U(s;\Omega) = \max_{B\in\mathcal{P}} U(B).$$

We accept the candidate if $U(s;\Omega) \le \varepsilon$. Otherwise, interval evaluations at selected points provide a lower bound L_Ω on the true maximum error $\delta(s;\Omega)$. If $L_\Omega > \varepsilon$, the candidate violates the tolerance; if neither condition holds, the candidate remains unresolved. For an unresolved candidate, we bisect the cell with the largest upper bound along its longest side, or use sensitivity-based splitting that eventually refines every unresolved direction. Each child uses the smaller of its new bound and its parent's bound, so $U(s;\Omega)$ cannot increase during refinement.

We call subdivision *fair* when the maximum unresolved-cell diameter tends to zero. With fair subdivision, exact or sufficiently refined arithmetic, and no budget limit, Theorem 5 in Appendix C.4 guarantees certification whenever the true error lies strictly below tolerance. To relate refinement to cell size, Theorem 2 bounds how far the Taylor certificate can exceed the true cell error.

Theorem 2 (Taylor certificate gap)**.** *Suppose $g(c+y) = a + b^\top y + e(y)$ on $|y_i| \le r_i$, with $|e(y)| \le R_B$. In exact arithmetic, set $U_T(B) = |a| + \sum_i |b_i| r_i + R_B$. Then*

$$0 \le U_T(B) - \sup_{x\in B} |g(x)| \le 2R_B.$$

Proof. See Appendix D.2. □

Theorem 2 limits overestimation to twice the remainder bound. Corollary 6 uses this limit to give sufficient cell sizes when the true error lies below tolerance. Appendix D.2 extends the analysis to finite precision.

5 Compression and Geometric guarantees

The local models in Section 4 also guide compression. We use them to adjust the final affine layer (output head), compensating for compression error while keeping the quantized hidden layers (body, or backbone) fixed.

Head repair. The output depends linearly on the head coefficients, allowing us to optimize them with the body fixed. Let $\psi(x) = (\widehat{h}_L(x)^\top, 1)^\top \in \mathbb{R}^m$ collect the decoded features and bias coordinate, where $m = n_L + 1$.

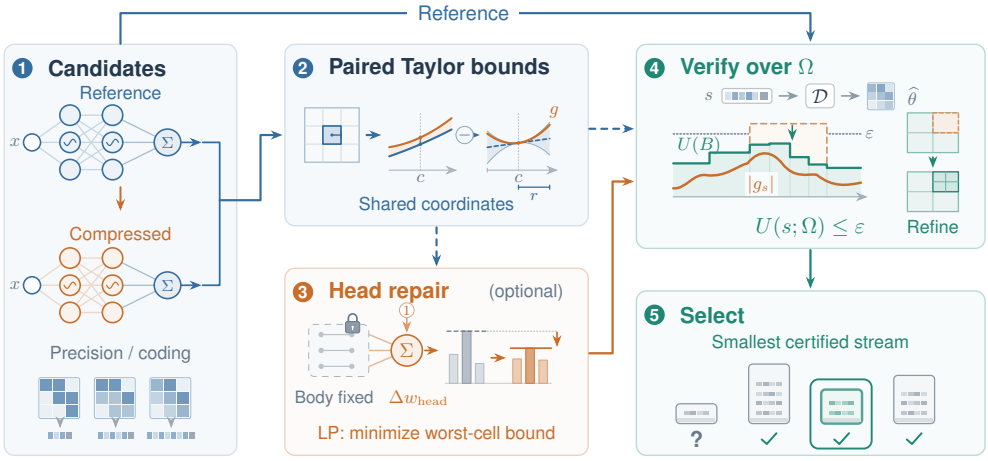


Figure 1: CertINR combines paired Taylor bounds, optional head repair, and adaptive verification to select the smallest certified stream found. Graphics are schematic; the residual is magnified, and ? denotes an unresolved candidate.

Starting from the reference head, let g denote the unrepaired compression residual. A head correction $\Delta w_{\text{head}} \in \mathbb{R}^m$ adds $\Delta w_{\text{head}}^\top \psi(x)$ to this residual. We expand both g and ψ in the same local coordinates $y = x - c$ to retain their shared spatial variation:

$$g(x) = a + b^\top y + e_g(x), \qquad \psi(x) = H + Jy + e_\psi(x), \qquad |e_g| \le R_g, \quad |e_\psi| \le R_\psi \tag{9}$$

At the cell center, $a = g(c) \in \mathbb{R}$ and $b = \nabla g(c) \in \mathbb{R}^d$ describe the residual, while $H = \psi(c) \in \mathbb{R}^m$ and $J = D_x\psi(c) \in \mathbb{R}^{m\times d}$ describe the decoded features. The nonnegative radii R_g and $R_\psi \in \mathbb{R}^m$ bound their nonlinear remainders. The bias feature has zero derivative and remainder. We hold these local models fixed while optimizing the head. Combining the shared terms gives the repaired-error bound as:

$$|g(x) + \Delta w_{\text{head}}^\top \psi(x)| \le |a + H^\top \Delta w_{\text{head}}| + \sum_i r_i |b_i + (J^\top \Delta w_{\text{head}})_i| + R_g + R_\psi^\top |\Delta w_{\text{head}}|. \tag{10}$$

Paired-head repair chooses one correction for all spatial cells and minimizes the largest repaired-error bound. Introducing auxiliary variables for the absolute-value terms yields a sparse linear program, given in Appendix E.1.

Acceptance depends on the stored head, so we serialize the corrected head (including its bias) in FP32, recover the decoded correction $\Delta w_{\text{head}}^{\text{dec}}$, and verify the joint bound with parameter and bound-evaluation rounding included. The repair portfolio also computes new paired Taylor, residual/Jacobian, and spectral bounds directly from the repaired network, plus a bound that treats the residual and correction separately (Appendix E). If repair increases the portfolio bound, we retain the original reference head, preserving previously certified candidates.

Candidate selection. After verification, CertINR selects the lowest-storage candidate whose certificate satisfies the target tolerance. Because higher bit precision does not necessarily imply lower error, each candidate is verified independently. Figure 1 summarizes the overall procedure; pseudocode is provided in Appendix E.3.

Signed distance fields. A selected stream satisfying $\|g_s\|_\infty \le \varepsilon$ preserves the reference sign wherever $|f_\theta(x)| > \varepsilon$. For a signed distance field (SDF), this preserves the reference inside–outside decision away from its zero set. Preserving the topology of the zero set requires additional control near the surface. Consider a field-value band $|f_\theta(x)| \le a$ with $a > \varepsilon$, contained in the interior of the verification domain. For every spatial cell B intersecting this band, let m_B be a lower bound on the reference gradient magnitude $\|\nabla f_\theta\|_2$ and let κ_B be an upper bound on the residual gradient magnitude $\|\nabla g_s\|_2$. We require the positive gradient gap $m_B - \kappa_B > 0$.

Under the stated smoothness and nonemptiness conditions, these value and gradient bounds guarantee a continuous deformation of the surrounding space that maps the reference zero set to the decoded zero

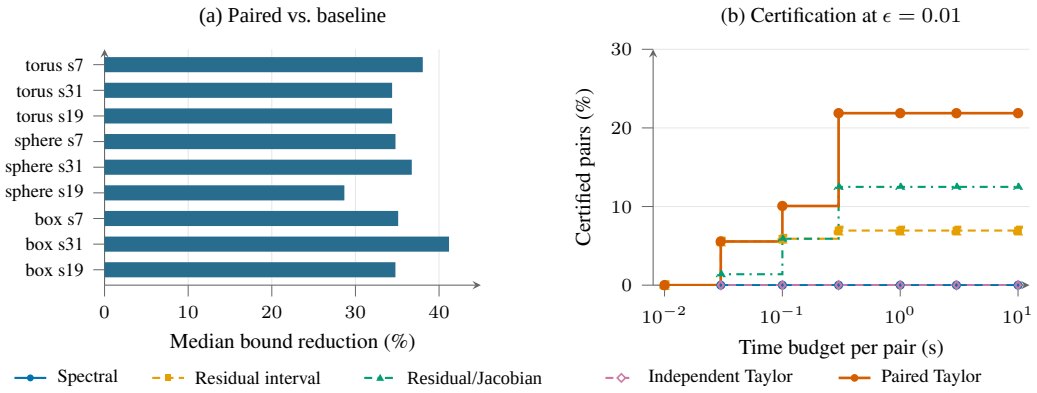


Figure 2: Paired Taylor bounds on 288 small-model pairs. Left: Median percentage reductions from the minimum of the spectral, residual-interval, and residual/Jacobian bounds (positive indicates tighter paired bounds; s denotes seed). Right: Percentage certified by wall-time deadline at $\varepsilon = 0.01$. Spatial methods have a 512-cell limit; plateaus indicate no further certificates.

set, with continuously differentiable maps and inverses. This *ambient isotopy* preserves connected components and holes. The same construction also bounds the *Hausdorff distance* between the two zero sets, i.e., the largest distance from a point on either surface to its nearest point on the other. Appendix F gives the formal conditions, and Appendix G proves the result. Section 6.1 evaluates these conditions in practice.

6 Experiments

We test when paired bounds improve certification, whether head repair reduces certified storage, and how these effects change on larger networks. Our analytic benchmark contains nine fixed reference checkpoints: 3-16-16-1 sine networks with 353 parameters, trained on the signed distances of a sphere, rounded box, and torus over $[-1, 1]^3$ using seeds 7, 19, and 31 (Appendix N). For each reference, 32 compressed bodies combine hidden-layer precisions $\{6, 8, 10, 12\}^2$ with direct or layerwise orthonormal DCT coding. Reference parameters and output heads remain in FP32, giving 288 reference–decoded pairs, aggregated by checkpoint. The implementation stores candidates as byte-aligned streams and reports storage in bytes, including all codec metadata.

For each pair, we compare paired Taylor with the minimum of spectral, residual-interval, and residual/Jacobian bounds. An independent Taylor ablation isolates nonlinear pairing: it retains signed coordinate slopes but propagates the two remainder bounds separately. At the 512-cell limit, paired Taylor gives a smaller bound on all 288 pairs. Checkpoint-level median reductions range from 28.68% to 41.19%, with a median of 34.79% across the nine checkpoints (Figure 2). At $\varepsilon = 0.01$, paired Taylor certifies 63 pairs and residual/Jacobian certifies 36. Both prove 205 violations, leaving 20 and 47 pairs unresolved, respectively.

The tighter bounds also yield more certificates at both reported time limits (Table 1). All spatial methods use grids with 1, 2, 4, and 8 subdivisions per axis, two linear-algebra threads, and a 4 GiB memory limit. Times include model loading, decoding, and verification, with one run per pair. By 0.3 seconds, every method has obtained all certificates recorded in its final column. Independent Taylor certifies none, demonstrating the benefit of joint remainder propagation on these pairs under the tested subdivision schedule. The ordering of the bounds changes on larger networks. Among 12 additional pairs, paired Taylor remains tighter on all three from the 2,273-parameter network. The baseline is tighter on the other nine, spanning the 12,801-parameter network and two 198,657-parameter networks. Section 6.2 examines this change on pretrained fields.

Table 1: Certified pairs (of 288) at $\varepsilon = 0.01$. Final: Results at the 512-cell limit or one spectral bound. Time: Median total wall time, including prep.

Method	0.1 s	0.3 s	Final	Time (ms)
Spectral	0	0	0	2.02
Residual interval	17	20	20	108.49
Residual/Jacobian	17	36	36	142.83
Independent Taylor	0	0	0	54.17
Paired Taylor	**29**	**63**	**63**	102.12

6.1 Repair and storage and Geometry checks

We next test whether head repair reduces the storage needed for certification. Six methods evaluate nine references and 32 compressed bodies per reference, giving 1,728 candidate–method evaluations.

The sampled baselines fit the head by minimizing mean squared error (MSE) or sampled maximum error; interval-head repair bounds residuals and features separately. These alternatives test the value of optimizing a bound that retains shared coordinates. All methods use the same 512-cell certificate portfolio for final verification and retain the reference head if repair increases the bound (Appendix I).

After complete candidate search at $\varepsilon = 0.02$, paired-head repair reduces mean selected storage by 5.84% relative to adaptive sampled minimax and by 6.62% relative to interval-head repair. Against adaptive sampled minimax, it selects smaller streams on six checkpoints and ties on three; full results are given in Appendix I.

These storage gains assume complete candidate search, which takes a mean of 26.65 seconds per reference for paired-head repair and 18.61 seconds for adaptive sampled minimax. At $\varepsilon = 0.02$, faster methods select smaller streams by 10 seconds; all methods finish within 30 seconds.

The selection certificate and sampled fidelity can also favor different heads. On 18 equal-byte bodies, paired-head repair has larger sampled maximum error than adaptive sampled minimax in 11 cases, with median paired/adaptive ratios of 1.6 for maximum error and 2.3 for MSE. Recomputing bounds directly from these pairs favors paired-head repair over adaptive sampled minimax in all 18 residual/Jacobian comparisons. Spectral bounds, however, favor the unrepaired and interval-head candidates over paired-head repair on all 18. Repair therefore improves the selection certificate without uniformly improving sampled fidelity or other certificates (Appendix I).

To assess topology near the zero set, we test the additional cellwise conditions of Corollary 9 on the nine analytic fields at band half-widths $a \in \{0.05, 0.02, 0.01\}$. These checks certify topology preservation in 20 of 27 configurations, covering seven fields.

Local gradient gaps matter: 11 successful cellwise checks fail the global test on the same spatial cover. The global test compares the smallest reference gradient with the largest residual gradient, even when they occur in different regions. Even a successful topology certificate can give a loose distance bound. For torus seed 19, the Hausdorff bound $V = 25.34$ exceeds the domain diameter and adds no distance information. Torus seeds 7 and 31 remain unresolved (Appendix J).

6.2 External verifiers and Pretrained fields

We compare paired Taylor with GenBaB (Shi et al., 2025), CROWN (Zhang et al., 2018), and α-CROWN (Xu et al., 2021) to test how general-purpose verifiers handle compression residuals. The comparison uses nine analytic 10/8-bit pairs at $\varepsilon = 0.02$ and six pretrained 16-bit pairs at $\varepsilon = 0.01$. The CROWN variants bound one input box, whereas paired Taylor subdivides; GenBaB has separate time limits for its two signed queries (Appendix L).

On the analytic pairs, paired Taylor certifies two, proves one violation, and leaves six unresolved; the standard external verifiers certify none. To test whether exposing cancellation helps these verifiers, we rewrite the same function as a residual computation graph. CROWN and α-CROWN bounds then decrease by medians of 97.15% and 98.58% (Appendix M). These bounds remain above tolerance and larger than paired Taylor's, whose comparison also includes subdivision. On the pretrained

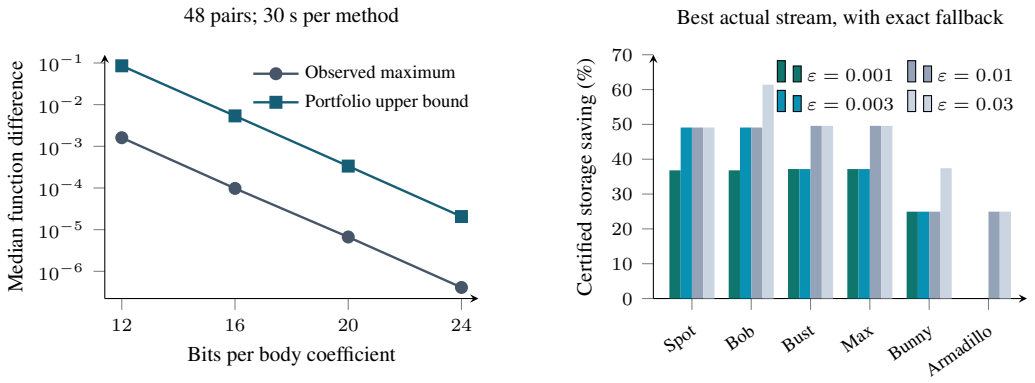


Figure 3: Certified compression on six pretrained SIRENs. Left: median portfolio bounds and median sampled maximum errors across 12 streams per precision. Bounds cover the full domain; maxima use 32,768 held-out queries. Right: per-shape byte savings at four tolerances relative to exact-copy storage, including all metadata. Spectral bounds certify every selected compressed stream.

pairs, the external methods certify none and paired Taylor certifies two. Spectral bounds certify four matching pairs in the separately timed pretrained benchmark (Appendix L).

To assess larger networks, we use six pretrained SIRENs from Buonomo et al. (2025), with 4,481 to 83,201 parameters and frequency 60 at every sine layer. Four body precisions under direct and DCT coding yield 48 compressed streams (Appendix K).

On all 48 pairs, the minimum of spectral and residual/Jacobian bounds is smaller than the paired Taylor bound. At $\varepsilon = 0.01$, spectral bounds certify 28 streams, residual/Jacobian certifies 23, and paired Taylor certifies 20. The portfolio certifies a compressed stream for every shape, with 41.21% mean per-shape storage savings relative to exact FP32 copies (range: 24.94%–49.58%). Pooling bytes across shapes gives a 30.01% reduction (Figure 3). These results motivate selecting among complementary bounds rather than relying on one verifier across all networks.

7 Discussion and Conclusion

CertINR treats continuous-domain error as a constraint in neural-field compression. Its verifier combines global spectral, local residual/Jacobian, and paired Taylor bounds, with the latter exploiting cancellation between reference and decoded networks at shared coordinates. The same certificates are integrated into compression: CertINR searches compressed candidates, optionally repairs the final affine layer, verifies the decoded representation, and selects the lowest-storage candidate meeting the target tolerance.

No single certificate dominates across regimes: paired Taylor is strongest on the analytic benchmarks, whereas spectral and residual/Jacobian bounds are stronger on larger pretrained models. At $\varepsilon = 0.02$, paired-head repair reduces mean certified storage by 5.84% relative to adaptive sampled minimax, and on six pretrained signed distance fields CertINR achieves 41.21% mean per-shape savings at $\varepsilon = 0.01$. These gains come with computational trade-offs, and minimizing a certificate need not improve sampled reconstruction error.

The guarantees concern fidelity to a fixed reference field, not the accuracy of that reference itself. Our implementation focuses on scalar-output sinusoidal networks. For signed distance fields, uniform error certificates preserve inside–outside decisions away from the zero set, while additional gradient and regularity conditions guarantee topology preservation and bound surface displacement. Reference-model and floating-point inference errors require separate control in downstream use. Extending verification to broader INR architectures and vector-valued fields, integrating stronger compression methods, and jointly optimizing the compressed body and output head are natural directions for future work.

AI-USE DISCLOSURE

We used OpenAI's ChatGPT for language editing, coding assistance, and identifying related works. The authors have fully reviewed and remain responsible for all content, mathematical statements, and experimental results.

ETHICS STATEMENT

The authors acknowledge and adhere to the ICLR Code of Ethics. Our research on compressing and verifying Implicit Neural Representations (INRs) utilizes standard, public 3D datasets and involves no human subjects or privacy-sensitive data. Furthermore, this work promotes AI safety by establishing provable error bounds and geometric guarantees, ensuring that compressed neural fields can be trusted in physical and digital downstream tasks, such as collision checking.

REPRODUCIBILITY STATEMENT

We state the assumptions and derive the verification, repair, and geometric guarantees in Appendices C–G, and specify certificate arithmetic in Appendix H. We detail the repair and geometry protocols in Appendices I and J, pretrained decoding in Appendix K, external-verifier comparisons in Appendices L and M, and analytic reference training in Appendix N. We provide a supplementary package containing the source code, experimental configurations, fixed reference checkpoints, compressed streams, numerical results, and scripts for generating figures and tables. The README and usage guide describe dependencies, validation checks, and commands for reproducing the experiments. To reduce the package size, we omit the large adaptive geometry cover archives while retaining their numerical summaries, shared geometric checks, and regeneration instructions. Upon acceptance, we will publicly release the code, configurations, and supporting results on GitHub, together with instructions for accessing the third-party datasets.

REFERENCES

Debangshu Banerjee, Changming Xu, and Gagandeep Singh. Input-relational verification of deep neural networks. *Proceedings of the ACM on Programming Languages*, 8(PLDI):147, 2024. doi: 10.1145/3656377.

Camille Buonomo, Julie Digne, and Raphaëlle Chaine. Volume preserving neural shape morphing. *Computer Graphics Forum*, 44(5):e70196, 2025. doi: 10.1111/cgf.70196.

David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Discrete & Computational Geometry*, 37(1):103–120, 2007. doi: 10.1007/s00454-006-1276-5.

Guillaume Coiffier and Louis Béthune. 1-Lipschitz neural distance fields. *Computer Graphics Forum*, 43(5):e15128, 2024. doi: 10.1111/cgf.15128.

Taha Entesari and Mahyar Fazlyab. Hierarchical end-to-end Taylor bounds for complete neural network verification. In *Proceedings of the 8th Annual Learning for Dynamics and Control Conference*, volume 331 of *Proceedings of Machine Learning Research*, pp. 1070–1087. PMLR, 2026. URL https://proceedings.mlr.press/v331/entesari26a.html.

Dhananjaya Jayasundara, Sudarshan Rajagopalan, Yasiru Ranasinghe, Trac D. Tran, and Vishal M. Patel. SINR: Sparsity driven compressed implicit neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3061–3070, 2025. doi: 10.1109/CVPR52734.2025.00291.

Anushrut Jignasu, Aditya Balu, Soumik Sarkar, Chinmay Hegde, Baskar Ganapathysubramanian, and Adarsh Krishnamurthy. SDFConnect: Neural implicit surface reconstruction of a sparse point cloud with topological constraints. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 5271–5279, 2024. doi: 10.1109/CVPRW63382.2024.00536.

Julian McGinnis, Suprosanna Shit, Florian A. Hölzl, Paul Friedrich, Paul Büschl, Vasiliki Sideri-Lampretsa, Mark Mühlau, Philippe C. Cattin, Bjoern Menze, Daniel Rueckert, and Benedikt Wiestler. Beyond uniformity: Regularizing implicit neural representations through a Lipschitz lens. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/655846cc914cb7ff977a1ada40866441-Abstract-Conference.html.

Jonathan B. Michaux, Yong Seok Kwon, Qingyi Chen, and Ram Vasudevan. Reachability-based trajectory design with neural implicit safety constraints. In *Robotics: Science and Systems*, 2023. doi: 10.15607/RSS.2023.XIX.062.

Sara Mohammadinejad, Brandon Paulsen, Jyotirmoy V. Deshmukh, and Chao Wang. DiffRNN: Differential verification of recurrent neural networks. In *Formal Modeling and Analysis of Timed Systems*, volume 12860 of *Lecture Notes in Computer Science*, pp. 117–134, 2021. doi: 10.1007/978-3-030-85037-1_8.

Brandon Paulsen, Jingbo Wang, and Chao Wang. ReluDiff: Differential verification of deep neural networks. In *International Conference on Software Engineering*, pp. 714–726, 2020a. doi: 10.1145/3377811.3380337.

Brandon Paulsen, Jingbo Wang, Jiawei Wang, and Chao Wang. NeuroDiff: Scalable differential verification of neural networks using fine-grained approximation. In *IEEE/ACM International Conference on Automated Software Engineering*, pp. 784–796, 2020b. doi: 10.1145/3324884.3416560.

Jonathan Richard Schwarz, Jihoon Tack, Yee Whye Teh, Jaeho Lee, and Jinwoo Shin. Modality-agnostic variational compression of implicit neural representations. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 30342–30364. PMLR, 2023. URL https://proceedings.mlr.press/v202/schwarz23a.html.

Nicholas Sharp and Alec Jacobson. Spelunking the deep: Guaranteed queries on general neural implicit surfaces via range analysis. *ACM Transactions on Graphics*, 41(4):107, 2022. doi: 10.1145/3528223.3530155.

Zhouxing Shi, Qirui Jin, Zico Kolter, Suman Jana, Cho-Jui Hsieh, and Huan Zhang. Neural network verification with branch-and-bound for general nonlinearities. In *Tools and Algorithms for the Construction and Analysis of Systems*, volume 15696 of *Lecture Notes in Computer Science*, pp. 315–335, 2025. doi: 10.1007/978-3-031-90643-5_17.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems*, volume 33, pp. 7462–7473, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/53c04118df112c13a8c34b38343b9c10-Abstract.html.

Xidan Song, Youcheng Sun, Mustafa A. Mustafa, and Lucas C. Cordeiro. QNNRepair: Quantized neural network repair. In *Software Engineering and Formal Methods*, volume 14323 of *Lecture Notes in Computer Science*, pp. 320–339, 2023. doi: 10.1007/978-3-031-47115-5_18.

Yannick Strümpler, Janis Postels, Ren Yang, Luc Van Gool, and Federico Tombari. Implicit neural representations for image compression. In *Computer Vision – ECCV 2022*, volume 13686 of *Lecture Notes in Computer Science*, pp. 74–91, 2022. doi: 10.1007/978-3-031-19809-0_5.

Tsui-Wei Weng, Pu Zhao, Sijia Liu, Pin-Yu Chen, Xue Lin, and Luca Daniel. Towards certificated model robustness against weight perturbations. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(4):6356–6363, 2020. doi: 10.1609/aaai.v34i04.6105.

Kaidi Xu, Huan Zhang, Shiqi Wang, Yihan Wang, Suman Jana, Xue Lin, and Cho-Jui Hsieh. Fast and Complete: Enabling complete neural network verification with rapid and massively parallel incomplete verifiers. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=nVZtXBI6LNn.

Huan Zhang, Tsui-Wei Weng, Pin-Yu Chen, Cho-Jui Hsieh, and Luca Daniel. Efficient neural network robustness certification with general activation functions. In *Advances in Neural Information Processing Systems*, volume 31, pp. 4939–4948, 2018. URL https://papers.nips.cc/paper_files/paper/2018/hash/d04863f100d59b3eb688a11f95b0ae60-Abstract.html.

Yedi Zhang, Fu Song, and Jun Sun. QEBVerif: Quantization error bound verification of neural networks. In *Computer Aided Verification*, volume 13965 of *Lecture Notes in Computer Science*, pp. 413–437, 2023. doi: 10.1007/978-3-031-37703-7_20.

Yedi Zhang, Guangke Chen, Fu Song, Jun Sun, and Jin Song Dong. Certified quantization strategy synthesis for neural networks. In *Formal Methods*, volume 14933 of *Lecture Notes in Computer Science*, pp. 343–362, 2025. doi: 10.1007/978-3-031-71162-6_18.

APPENDIX OVERVIEW

The appendices follow the same progression as the main text: error bounds, head repair, geometric guarantees, then numerical and experimental details. Table 2 locates each argument and its supporting experiments.

Table 2: Guide to the theoretical arguments, numerical guarantees, and experimental details.

Appendix	Contents
A	Notation, verification domain, and scope.
B	Why average error does not certify sign decisions; proof of Proposition 3.
C	Global spectral bound, exact residual and Jacobian recurrences, soundness, and refinement.
D	Paired Taylor remainder, certificate gap, sufficient partition size, smooth activations, and gradient and rounding bounds.
E	Joint paired-head repair objective and interval-head baseline.
F	Sign and clearance guarantees; global and cellwise conditions for topology preservation.
G	Full proof of the geometric guarantee and its distance interpretation.
H	Arithmetic model, supported networks, and numerical reproducibility.
I	Head-repair comparisons, storage, computation budgets, fidelity, and recomputed bounds.
J	Cellwise geometry experiments and refinement results.
K	Public pretrained models, frequency-aware decoding, bounds, and storage.
L	Official GenBaB, CROWN, and α-CROWN comparisons on analytic and pretrained pairs.
M	Residual computation-graph control on the nine analytic pairs.
N	Analytic reference training configurations and fit errors.

A NOTATION AND SCOPE

Table 3 summarizes the quantities used to compare the decoded and reference fields. For a fixed stream, we abbreviate f_θ by f, $f_{\mathcal{D}(s)}$ by $\widehat{f}$, and $g_s = \widehat{f} - f$ by g. Unless stated otherwise, all suprema and certificates concern the verification domain Ω.

Table 3: Notation for compression error, its certificates, head repair, and geometric interpretation.

Symbol	Meaning
$\theta, \widehat{\theta}, \mathcal{D}$	Reference and decoded parameters; stream decoder
$g = \widehat{f} - f$	Compression residual
$\rho = \|\widehat{\theta} - \theta\|_\infty$	Largest absolute change in effective weights or biases
$R_g(B), R_B$	Bounds on nonlinear error beyond the local affine model
δ, U, L_Ω	True maximum absolute error, certified upper bound, certified pointwise lower bound on that maximum
B, c, r_i	Spatial cell, center, and coordinate half-widths
$e_\ell, \Delta z_\ell$	Hidden activation and preactivation differences
J_ℓ, K_ℓ	Reference and residual spatial Jacobians
$\omega_\ell, A_\ell, a_\ell$	Sine frequency and effective weights/bias
d, ζ	Input dimension; local preactivation difference
v, w, κ_q	Reference/residual variation radii; difference curvature bound
ψ, H, J	Augmented decoded-body features, center, and Jacobian
Δw_{head}	Proposed output-head correction
$\Delta w_{\text{head}}^{\text{dec}}$	Correction represented by the serialized output head
ε, η, τ	Compression, reference-model, and inference arithmetic error budgets
a, m, κ	Field-value band half-width; reference-gradient lower bound; residual-gradient upper bound

Coordinate normalization changes derivative scales, so certificates and geometric bounds use the same coordinates. A field-value bound does not automatically bound spatial distance; Appendix F gives sufficient conditions for that conversion. Certificates apply only on the verified box or union of boxes. The geometric results concern scalar fields, but value certificates extend to vector outputs by combining componentwise bounds in the chosen output norm.

B AVERAGE DISTORTION DOES NOT CERTIFY SIGN PRESERVATION

Average distortion can hide a sign reversal confined to a small region. In Proposition 3, we fix the perturbation's amplitude and shrink its support. The sign reversal persists at a point with a fixed reference margin, even as the mean squared error tends to zero.

Proposition 3 (Average distortion does not certify a margin)**.** *For every $d \geq 1$, there are smooth scalar fields on $[-1,1]^d$ and perturbations with mean squared error tending to zero that reverse the sign at a point with a fixed positive reference margin.*

Proof of Proposition 3. We choose $f(x) = x_1$ and $x_0 = (a, 0, \ldots, 0)$ with $a = 1/4$, giving a reference margin a at x_0. A smooth nonnegative bump function ϕ supported inside the unit Euclidean ball, with $\phi(0) = 1$ and $\|\phi\|_\infty = 1$, fixes the change at this point while we vary the affected volume. For $0 < w < a/2$, define

$$g_w(x) = -2a\phi((x - x_0)/w), \qquad \widehat{f}_w = f + g_w.$$

The perturbation stays inside the domain, has uniform magnitude $2a$, and changes $f(x_0) = a$ to $\widehat{f}_w(x_0) = -a$. Its support volume, however, scales as w^d. Changing variables under the uniform probability measure on $[-1,1]^d$ gives

$$\mathbb{E}[g_w(x)^2] = 2^{-d}(2a)^2 w^d \int_{\mathbb{R}^d} \phi(u)^2 \, du \longrightarrow 0.$$

The perturbed field therefore retains the opposite sign as its mean squared error tends to zero. This construction concerns smooth scalar fields; a fixed neural architecture or quantizer need not realize them. Guarantees under such restrictions require additional assumptions. □

C RESIDUAL BOUNDS

We first bound compression error globally using matrix norms, then retain spatial dependence through residual and Jacobian recurrences. Their interval enclosures contain every possible value on a cell, establishing valid bounds and explaining when subdivision resolves a candidate.

C.1 GLOBAL SPECTRAL BOUND

Our global certificate tracks how each layer amplifies coefficient changes. Let H_ℓ and D_ℓ bound the Euclidean norms of the reference activation and activation residual at hidden layer ℓ. Matrix norms are induced spectral norms. The shared input gives $H_0 = \sup_{x \in \Omega} \|x\|_2$ and $D_0 = 0$. For a sine layer of width n_ℓ, we use $H_\ell = \sqrt{n_\ell}$ and combine the residual identity with the activation range:

$$D_\ell = \min\left\{2\sqrt{n_\ell}, \left\|\widehat{A}_\ell\right\|_2 D_{\ell-1} + \|\Delta A_\ell\|_2 H_{\ell-1} + \|\Delta a_\ell\|_2\right\}, \qquad 1 \leq \ell \leq L. \tag{11}$$

The recurrence in Equation (11) combines propagated residual error with new weight and bias errors, capped by the sine range. The affine output head has the same decomposition, giving the global field-error bound

$$U_{\text{global}} = \left\|\widehat{A}_{L+1}\right\|_2 D_L + \|\Delta A_{L+1}\|_2 \sqrt{n_L} + \|\Delta a_{L+1}\|_2.$$

We verify $\|e_\ell\|_2 \leq D_\ell$ at every input by induction, applying Equation (4), the triangle inequality, and the one-Lipschitz property of sine at each layer. The domain affects this bound through H_0, while frequency scaling enters through the effective weights and their perturbations. Numerical certification requires upper bounds on both the matrix norms and the propagated residual magnitudes.

Numerical norm bounds. A computed matrix factorization need not be exact, so we bound its reconstruction error and deviation from orthogonality. For $A \in \mathbb{R}^{p\times q}$, take factors $Q \in \mathbb{R}^{p\times k}$, $\Sigma \in \mathbb{R}^{k\times k}$, and $W \in \mathbb{R}^{k\times q}$ with $A \approx Q\Sigma W$. Using the Frobenius norm $\|\cdot\|_F$ and the identity on $\mathbb{R}^k$, we certify

$$e_Q \geq \|Q^\top Q - I\|_F, \qquad e_W \geq \|WW^\top - I\|_F, \qquad e_R \geq \|A - Q\Sigma W\|_F.$$

The matrix-product norm inequality and $\|M\|_2 \leq \|M\|_F$ then give

$$\|A\|_2 \leq \sqrt{1+e_Q}\,\|\Sigma\|_2\sqrt{1+e_W} + e_R.$$

When Σ is diagonal, $\|\Sigma\|_2 = \max_i |\Sigma_{ii}|$; otherwise, its norm also requires a certified upper bound. These error terms make the factorization bound valid. For an interval matrix with midpoint A_0 and entrywise half-width matrix R_A, every enclosed matrix differs from A_0 by at most $\|R_A\|_F$ in spectral norm. Adding $\|R_A\|_F$ to the certified midpoint norm covers this uncertainty.

C.2 Exact difference identities

The global recurrence summarizes each layer by a norm. To retain spatial dependence, we subtract the reference and decoded computations before bounding them. Their preactivations satisfy

$$\widehat{A}_\ell\widehat{h}_{\ell-1} - A_\ell h_{\ell-1} = \widehat{A}_\ell(\widehat{h}_{\ell-1} - h_{\ell-1}) + (\widehat{A}_\ell - A_\ell)h_{\ell-1},$$

which proves Equation (4) after adding the bias difference. The scalar identities

$$\sin(u+v) - \sin u = 2\sin(v/2)\cos(u+v/2),$$
$$\cos(u+v) - \cos u = -2\sin(u+v/2)\sin(v/2)$$

give the activation residual in Equation (5) and the derivative difference in Equation (12). Both retain a factor that vanishes with the preactivation perturbation.

For the spatial derivatives, we write $J_\ell = \partial h_\ell/\partial x$ and $K_\ell = \partial e_\ell/\partial x$ for the reference and residual Jacobians in $\mathbb{R}^{n_\ell\times d}$. The initial values $J_0 = I_d$ and $K_0 = 0$ reflect the shared input, and the decoded Jacobian is $\widehat{J}_\ell = J_\ell + K_\ell$. With diagonal activation derivatives $C_\ell = \mathrm{Diag}(\cos z_\ell)$ and $\widehat{C}_\ell = \mathrm{Diag}(\cos(z_\ell + \Delta z_\ell))$, the difference $\Delta C_\ell = \widehat{C}_\ell - C_\ell$ retains an explicit perturbation factor:

$$\Delta C_\ell = \mathrm{Diag}\left[-2\sin(z_\ell + \Delta z_\ell/2)\odot\sin(\Delta z_\ell/2)\right], \tag{12}$$

$$J_\ell = C_\ell A_\ell J_{\ell-1}, \qquad K_\ell = \widehat{C}_\ell\big(\widehat{A}_\ell K_{\ell-1} + \Delta A_\ell J_{\ell-1}\big) + \Delta C_\ell A_\ell J_{\ell-1}. \tag{13}$$

Subtracting the two Jacobian recurrences gives

$$\begin{aligned}\widehat{J}_\ell - J_\ell &= \widehat{C}_\ell\widehat{A}_\ell\widehat{J}_{\ell-1} - C_\ell A_\ell J_{\ell-1}\\ &= \widehat{C}_\ell\widehat{A}_\ell K_{\ell-1} + \widehat{C}_\ell\Delta A_\ell J_{\ell-1} + (\widehat{C}_\ell - C_\ell)A_\ell J_{\ell-1},\end{aligned}$$

which establishes Equation (13). The final affine layer has derivative $(\nabla_x g_s)^\top = \widehat{A}_{L+1}K_L + \Delta A_{L+1}J_L$. Differentiating the residual before bounding it preserves cancellation that separate bounds on the two gradients can lose.

C.3 Soundness and sensitivity

We evaluate the residual identities using *natural interval extensions*, replacing each operation with an interval that contains all results allowed by its input intervals. Rounding can give even a single-point, or *singleton*, evaluation positive width. Theorem 4 proves that these bounds are valid (sound) and shrink with the parameter perturbation.

Theorem 4 (Soundness and perturbation sensitivity)**.** *Evaluate the stated recurrences using bounded natural interval extensions, intersect trigonometric ranges with $[-1,1]$, and cap each sine magnitude by the magnitude of its argument. Then $\sup_{x\in B}|g_s(x)| \leq U_{\mathrm{RJ}}(B)$. In exact interval arithmetic, the residual activation and Jacobian enclosures are zero when the two checkpoints coincide.*

Fix the architecture, frequency multipliers, and compact input domain. Assume bounded reference parameters and decoded parameters in a sufficiently small neighborhood of the reference. Let $\rho = \|\widehat{\theta} - \theta\|_\infty$. In exact interval arithmetic, a finite constant C satisfies $\mathrm{mag}(G_B) \leq C\rho$ and $M_i(B) \leq C\rho$ uniformly over all subcells and coordinates. Consequently, the mean-value term exceeds $|g_s(c)|$ by at most $\chi_c + C\rho\sum_i r_i$. Here, $\chi_c = \mathrm{mag}(G_c) - |g_s(c)| \geq 0$ measures overestimation at the cell center. Exact singleton evaluation gives $\chi_c = 0$.

Proof of Theorem 4. For the fixed candidate $g = g_s$, we initialize the recurrences with the input cell, an exact zero residual, and the identity reference Jacobian. Each affine map, product, sine, and cosine uses its natural interval extension with the stated caps. Induction then encloses every layer value for every $x \in B$, including the output residual G_B and derivatives $\partial_i g(B)$.

The output enclosure justifies the first term of Equation (6). For the centered term, we integrate the enclosed derivative along the segment from c to x. Convexity keeps this segment inside B, giving

$$|g(x)| \le |g(c)| + \sum_i |x_i - c_i| \sup_{y \in B} |\partial_i g(y)| \le \mathrm{mag}(G_c) + \sum_i r_i M_i(B).$$

Both terms in Equation (6) therefore bound the same cell error, so their minimum is valid.

When $\Delta A = \Delta a = 0$, the initially zero residual stays zero: Equation (4) gives $\Delta z = 0$ at every layer, and $\sin(0) = 0$ makes $e = 0$ and $\Delta C = 0$ in Equations (5) and (12). Equation (13) then gives $K = 0$, so the output bounds vanish as well. Exact-zero checks preserve this identity in our implementation.

For perturbation sensitivity, let ΔZ_ℓ and E_ℓ enclose Δz_ℓ and e_ℓ, respectively, with magnitudes measured in any fixed finite-dimensional vector norm. The reference activations, phases, and Jacobians have finite enclosures on the root box and hence uniformly finite bounds on all subboxes. We measure ρ between the effective coefficients, including decoding roundoff.

We now use these uniform reference bounds to induct on the residual magnitude. If the incoming residual activation is $O(\rho)$, bounded decoded weights and $O(\rho)$ parameter differences give $\mathrm{mag}(\Delta Z_\ell) = O(\rho)$ in Equation (4). The bound $|\sin t| \le |t|$ then gives $\mathrm{mag}(E_\ell) = O(\rho)$ and $\mathrm{mag}(\Delta C_\ell) = O(\rho)$. The same argument in Equation (13), starting from $K_0 = 0$, gives $\mathrm{mag}(K_\ell) = O(\rho)$. There are only finitely many layers and coordinate derivatives, so one constant C bounds all output and derivative magnitudes. Substituting these derivative bounds and $\mathrm{mag}(G_c) = |g_s(c)| + \chi_c$ proves the stated excess bound for the mean-value term. Exact singleton evaluation gives $\chi_c = 0$, although a valid enclosure need not have zero width. □

C.4 Refinement and finite precision

A valid bound may still exceed tolerance when the true error does not. Subdivision succeeds under *strict slack*: the true maximum error is strictly below tolerance. Building on Theorem 4, Theorem 5 proves eventual certification in this case and shows when point evaluations establish a violation.

Theorem 5 (Termination with strict slack)**.** *Fix a pair of sine networks on a compact box. Assume exact arithmetic, or arithmetic and transcendental enclosures whose numerical excess tends uniformly to zero under a refining arithmetic schedule. Use fair bisection so that the maximum unresolved cell diameter tends to zero. If $\delta < \varepsilon$, this process terminates with a certificate. If $\delta > \varepsilon$, a dense sequence of point evaluations with enclosure widths tending to zero eventually proves a violation. At $\delta = \varepsilon$, finite termination is not guaranteed.*

Proof of Theorem 5. For the fixed network pair, the bounded interval Jacobians on the initial box give a uniform bound $M < \infty$ on every coordinate derivative of the residual on its subboxes. With exact singleton evaluations, the centered cell bound satisfies

$$U(B) \le |g(c)| + M \sum_i r_i \le \delta + M \sum_i r_i$$

in exact arithmetic. Thus $\delta < \varepsilon$ leaves enough slack to certify every cell with $M \sum_i r_i < \varepsilon - \delta$. Fair subdivision reaches such a partition after finitely many splits. With numerical arithmetic, we reserve part of the slack for enclosure width and tighten the numerical enclosures until the same condition holds.

When $\delta > \varepsilon$, continuity on the compact domain gives a relatively open neighborhood where $|g| > \varepsilon$. A dense sequence of point evaluations reaches this neighborhood, where sufficiently narrow intervals give an absolute-error lower bound greater than ε. Neither strict argument applies at equality. □

Certifying each box extends the result to a finite union. Numerical error and spatial variation both consume slack, so certification may require higher precision as well as smaller cells. A violation requires a point evaluation with a certified lower error bound above tolerance.

D PAIRED TAYLOR ANALYSIS

D.1 PROOF OF PROPOSITION 1

The derivative argument establishes eventual certification under strict slack; paired Taylor analysis now quantifies how the local remainder decreases. For Proposition 1, we use the fact that the activation difference's curvature with respect to the reference preactivation vanishes with the perturbation. This dependence lets us bound the remainder at each layer while retaining the parameter-change factor.

Proof of Proposition 1. For one activation, write $q(z, \zeta) = \sin(z + \zeta) - \sin z$. Its center derivatives are $q_z = \Delta c_c$ and $q_\zeta = \widehat{c}_c$, while its second derivatives satisfy

$$q_{zz} = -\sin(z + \zeta) + \sin z, \qquad q_{z\zeta} = q_{\zeta\zeta} = -\sin(z + \zeta).$$

The segment from (z_c, ζ_c) to a feasible pair $(z(x), \zeta(x))$ satisfies $|z - z_c| \le v$, $|\zeta - \zeta_c| \le w$, and $|\zeta| \le |\zeta_c| + w$. Consequently, $\kappa_q = \min\{2, |\zeta_c| + w\}$ from Section 4.3 bounds $|q_{zz}|$ componentwise, as well as the paired sine and cosine differences used below. The other two derivative magnitudes are at most one. The second-order integral Taylor formula therefore bounds the nonlinear remainder by

$$\tfrac{1}{2}\kappa_q v^2 + vw + \tfrac{1}{2}w^2.$$

The preactivation remainders also contribute through the linear term. Substituting $z - z_c = L_z y + \xi_z$ and $\zeta - \zeta_c = L_\zeta y + \xi_\zeta$ adds at most $|\Delta c_c| R_z + |\widehat{c}_c| R_\zeta$, establishing Equation (7). Applying $|\sin''| \le 1$ directly gives the corresponding reference formula.

Both remainder radii start at zero at the input. Across an affine map, they satisfy

$$R_z = |A| R_h, \qquad R_\zeta = |\widehat{A}| R_e + |\Delta A| R_h.$$

With a fixed finite architecture and bounded parameters, the reference Jacobians stay bounded, while residual center values and Jacobians are $O(\rho)$. The recurrences then preserve $R_h = O(\|r\|_1^2)$, $R_e = O(\rho\|r\|_1^2)$, $v = O(\|r\|_1)$, and $w = O(\rho\|r\|_1)$ by induction. Because $\kappa_q = O(\rho)$, every term of Equation (7) has the required local order in radius and perturbation. The final affine layer preserves this order. □

This scaling describes how the bound shrinks, but its constants determine the required cell size. Absolute-weight matrices propagate remainder radii through affine layers, while nonlinear terms depend on the reference variation v and residual variation w. Frequencies enter both the effective weights and spatial Jacobians, so amplification across layers can leave loose bounds even for small perturbations. To limit this growth, we cap R_h by $2 + |L_h| r$ and, using $|e(x)| \le \kappa_q$, cap R_e by $\kappa_q + |e(c)| + |L_e| r$. These caps preserve validity and the local order. Arithmetic enclosure widths impose a numerical floor; identical checkpoints have exact zero error.

D.2 PROOF OF THEOREM 2 AND REFINEMENT CONSEQUENCES

An affine function attains its exact maximum absolute value at a box corner, so the remainder determines how far a Taylor certificate can exceed the true error. In proving Theorem 2, we also identify a corner whose residual nearly attains that certificate. Let $B = \{c + y : |y_i| \le r_i\}$, with exact real coefficients $a \in \mathbb{R}$, $b \in \mathbb{R}^d$, and $R_B \ge 0$ satisfying

$$g(c + y) = a + b^\top y + e_B(y), \qquad |e_B(y)| \le R_B \quad (c + y \in B).$$

Write $A_B = |a| + \sum_i |b_i| r_i$ for the affine maximum, $U_B = A_B + R_B$ for the cell certificate, and $\delta_B = \sup_{x \in B} |g(x)|$ for the true cell error. A corner $x_B^\star$ of B then satisfies

$$0 \le U_B - \delta_B \le U_B - |g(x_B^\star)| \le 2R_B. \tag{14}$$

Equation (14) controls both certificate overestimation and how closely the corner error approaches the certificate. Taking the largest cell bound extends this result to any finite box partition $\mathcal{P}$ of Ω:

$$0 \le \max_{B \in \mathcal{P}} U_B - \sup_{x \in \Omega} |g(x)| \le 2 \max_{B \in \mathcal{P}} R_B.$$

Proof of Theorem 2 and its corner-witness form. Choose $\varsigma \in \{-1, 1\}$ with $\varsigma a = |a|$, allowing either sign when $a = 0$, and set $y_i^\star = \varsigma \operatorname{sign}(b_i) r_i$, with either endpoint when $b_i = 0$. At the corner $x_B^\star = c + y^\star$, the affine part attains $|a + b^\top y^\star| = A_B$. We apply the triangle and reverse-triangle inequalities to obtain $\delta_B \le A_B + R_B$ and $|g(c + y^\star)| \ge A_B - R_B$. Together they prove Equation (14), even when $A_B < R_B$. For a partition, every cell satisfies $U_B \le \delta_B + 2R_B \le \sup_\Omega |g| + 2 \max_{B'} R_{B'}$, so taking the maximum proves the domain result. □

Evaluating the corner in Equation (14) gives a certified lower bound on the maximum error, which we compare with the cell certificate. Together with the perturbation-sensitive remainder, this gap estimate yields Corollary 6: a bound on the number N of cells sufficient for certification under strict slack.

Corollary 6 (Perturbation-sensitive subdivision with strict slack)**.** *Let $\Omega = \prod_{i=1}^d [c_i - s_i, c_i + s_i]$ with $s_i > 0$. We fix an architecture, a bounded parameter neighborhood, and $0 < \rho \le \rho_0$. Assume common constants $C > 0$ and $r_0 > 0$ such that the exact-arithmetic affine models of g satisfy*

$$R_B \le C\rho \|r_B\|_1^2$$

on every box $B \subset \Omega$ whose half-width vector r_B satisfies $\|r_B\|_1 \le r_0$. We write $\delta = \sup_\Omega |g|$ and require strict slack $\sigma = \varepsilon - \delta > 0$. We choose the coordinate half-width threshold

$$h_\star = \min\left\{ \frac{r_0}{d}, \sqrt{\frac{\sigma}{2C\rho d^2}} \right\}. \tag{15}$$

A regular partition with every coordinate half-width at most $h_\star$ certifies $\delta \le \varepsilon$ using the affine-model upper bounds. Such a partition exists with

$$N \le \prod_{i=1}^d \left\lceil \frac{s_i}{h_\star} \right\rceil \le \prod_{i=1}^d \left(1 + \frac{s_i}{h_\star} \right).$$

For fixed domain, dimension, architecture, C, and r_0, this is $N = O\big(1 + (\rho/\sigma)^{d/2}\big)$.

Proof of Corollary 6. We partition coordinate i into $\lceil s_i/h_\star \rceil$ equal intervals. The resulting half-widths satisfy $\|r_B\|_1 \le d h_\star \le r_0$, so every remainder obeys $R_B \le C\rho d^2 h_\star^2$. The choice of $h_\star$ then leaves enough slack for Theorem 2 to give

$$\max_B U_B \le \delta + 2C\rho d^2 h_\star^2 \le \delta + \sigma = \varepsilon.$$

The construction gives the stated cell count. With the listed quantities fixed, $h_\star^{-1}$ is at most a constant times $1 + \sqrt{\rho/\sigma}$, which yields the claimed order. □

The two terms in Equation (15) keep the radius within the remainder estimate's valid range and the error within the available slack. Corollary 6 gives a sufficient cell count; the algorithm need not know δ or σ and can refine until its bounds meet tolerance. Exact identity handles $\rho = 0$, whereas zero slack gives no finite partition guarantee. Constructing each cell model incurs a separate cost.

Finite-precision bounds. Rounding and interval widths consume some of the slack in Theorem 2. Suppose the certified numerical bounds satisfy $U_B^{\mathrm{num}} \le U_B + \nu$ uniformly, for a justified arithmetic allowance $0 \le \nu < \sigma$. Replacing σ by $\sigma - \nu$ in Equation (15) reserves this allowance and leaves enough slack for spatial error. If the computed corner witness also underestimates $|g(x_B^\star)|$ by at most ν_L, the numerical upper–lower gap is at most $2R_B + \nu + \nu_L$. Arbitrarily small slack may require higher precision and tighter function approximations, as well as finer cells.

D.3 Smooth activations

The sine construction retains cancellation by bounding the change in activation curvature between the two networks. This mechanism extends to activations with three continuous derivatives: the first two derivative bounds control the affine and quadratic terms, while the third controls the curvature change. Proposition 7 formalizes this extension.

Proposition 7 (Paired remainder for a smooth activation)**.** *Let ϕ be C^3 on an open interval containing a compact interval I. On one cell, take $z_c = z(c)$ and $\zeta_c = \zeta(c)$. Suppose the affine models satisfy*

$$z = z_c + L_z y + \xi_z, \qquad \zeta = \zeta_c + L_\zeta y + \xi_\zeta, \qquad |\xi_z| \le R_z, \quad |\xi_\zeta| \le R_\zeta.$$

We set $v = |L_z|r + R_z$ and $w = |L_\zeta|r + R_\zeta$, so $M_\zeta = |\zeta_c| + w$ bounds the preactivation difference. Require I to contain the reference and decoded preactivations at the center and everywhere on the cell. The smallest interval containing both certified ranges is a sufficient computable choice. For $j = 1, 2, 3$, choose $L_j \ge \sup_{u \in I} |\phi^{(j)}(u)|$. We use these derivative bounds to define

$$H_{zz} = \min\{2L_2, L_3 M_\zeta\}, \qquad \Delta\phi'_c = \phi'(z_c + \zeta_c) - \phi'(z_c).$$

For $q(z, \zeta) = \phi(z + \zeta) - \phi(z)$, a valid affine model has center $q(z_c, \zeta_c)$, linear coefficient

$$L_e = \phi'(z_c + \zeta_c) L_\zeta + \Delta\phi'_c L_z,$$

and remainder radius

$$R_e = |\phi'(z_c + \zeta_c)| R_\zeta + |\Delta\phi'_c| R_z + \tfrac{1}{2} H_{zz} v^2 + L_2 v w + \tfrac{1}{2} L_2 w^2. \tag{16}$$

The reference activation model has coefficient $L_h = \phi'(z_c) L_z$ and valid remainder radius

$$R_h = |\phi'(z_c)| R_z + \tfrac{1}{2} L_2 v^2.$$

The statement applies componentwise to elementwise activations, with separate valid derivative bounds for each component if desired.

Proof of Proposition 7. We control the remainder through paired curvature along the Taylor expansion segment. Differentiating the activation difference gives

$$q_z = \phi'(z + \zeta) - \phi'(z), \qquad q_\zeta = \phi'(z + \zeta),$$
$$q_{zz} = \phi''(z + \zeta) - \phi''(z), \quad q_{z\zeta} = q_{\zeta\zeta} = \phi''(z + \zeta).$$

Consider the segment from (z_c, ζ_c) to a fixed feasible (z, ζ). Because I is an interval, it contains the reference and decoded preactivations along this segment, together with the intervals between them. Using $|\zeta| \le M_\zeta$, the triangle inequality and mean-value theorem give $|q_{zz}| \le H_{zz}$, while $|q_{z\zeta}|, |q_{\zeta\zeta}| \le L_2$. Taylor's theorem bounds the nonlinear remainder by $H_{zz} v^2/2 + L_2 v w + L_2 w^2/2$. The first two terms of Equation (16) account for the preactivation remainders when we substitute the affine models into the linear term. The same argument applied directly to $\phi(z)$ gives the reference formula. □

Network perturbation sensitivity. To extend Equation (16) through the network, fix a finite architecture, frequency multipliers, a compact input domain, and a compact convex parameter neighborhood. Assume each activation's reachable preactivations lie in a compact interval I within its C^3 domain. Continuity and compactness bound the derivatives uniformly. With $\rho = \|\widehat{\theta} - \theta\|_\infty$ measured after decoding, including frequency scaling and reconstruction rounding, smooth parameter dependence makes residual center values and Jacobians $O(\rho)$ uniformly in the cell center.

The input remainder radii are zero, and the affine updates $R_z = |A| R_h$ and $R_\zeta = |\widehat{A}| R_e + |\Delta A| R_h$ preserve the perturbation factor. Under the inductive hypotheses $R_h = O(\|r\|_1^2)$ and $R_e = O(\rho \|r\|_1^2)$, we obtain

$$v = O(\|r\|_1), \quad w = O(\rho \|r\|_1), \quad H_{zz} = O(\rho), \quad |\Delta\phi'_c| \le L_2 |\zeta_c| = O(\rho).$$

Substitution into Equation (16) preserves $R_e = O(\rho \|r\|_1^2)$; the reference recurrence likewise preserves $R_h = O(\|r\|_1^2)$, and both orders survive the final affine layer. Their constants depend on the architecture and derivative bounds and are uniform over sufficiently small radii and perturbations in the specified neighborhood. This establishes the exact-arithmetic hypothesis of Corollary 6. General activations need valid range and derivative bounds in place of the sine-specific caps. The argument also covers C^2 activations with a Lipschitz second derivative, with L_3 interpreted as its Lipschitz constant.

D.4 Gradient bounds

The geometric conditions in Theorem 8 also require gradient bounds throughout a band around the surface. We derive them from the local models by tracking changes from the center Jacobians. Let nonnegative matrices E_J and E_K bound the reference and residual Jacobian deviations componentwise. Across an affine layer, we propagate them as

$$E_z = |A|E_J, \qquad E_\zeta = |\widehat{A}|E_K + |\Delta A|E_J.$$

For the activation step, C_{abs} and $\widehat{C}_{\mathrm{abs}}$ bound the absolute reference and decoded cosines componentwise. Valid choices are

$$C_{\mathrm{abs}} = \min\{1, |\cos z_c| + v\}, \qquad \widehat{C}_{\mathrm{abs}} = \min\{1, |\widehat{c}_c| + v + w\}.$$

The cosine difference has magnitude at most κ_q. Together with the cosine variation bounds, this gives the deviations after sine activation, denoted by the superscript $+$:

$$\begin{aligned} E_J^+ &:= C_{\mathrm{abs}}E_z + \min\{2, v\}|L_z|, \\ E_K^+ &:= \widehat{C}_{\mathrm{abs}}E_\zeta + \kappa_q E_z + \min\{2, v+w\}|L_\zeta| + \min\{4, \kappa_q v + w\}|L_z|. \end{aligned} \tag{17}$$

Activation vectors in Equation (17) scale the corresponding Jacobian rows. Its final term follows by differentiating $\cos(z+\zeta) - \cos z$ with respect to (z, ζ): the derivative magnitudes on the segment are at most κ_q and one, so the mean-value theorem bounds the change by $\kappa_q v + w$. Starting from zero deviations, we propagate these bounds to the output. Adding them to the center-Jacobian intervals gives enclosures valid throughout the cell.

D.5 Rounding error

The recurrences describe exact evaluation of the decoded network; computing their bounds must also account for rounding. Our batched implementation assumes finite IEEE-754 binary64 arithmetic, round-to-nearest elementary operations, gradual underflow, and the standard dot-product error model for its Basic Linear Algebra Subprograms (BLAS) products. For unit roundoff $u = 2^{-53}$ and dot-product length n, define $\gamma_{2n} = 2nu/(1-2nu)$. The assumption $4nu < 1$ ensures $0 < \gamma_{2n} < 1$ and gives the dot-product error bound

$$|\,\mathrm{fl}(a^\top x) - a^\top x| \le \gamma_{2n}|a|^\top|x| + n\,\mathrm{tiny},$$

where fl denotes the computed floating-point result and tiny is the smallest positive normal binary64 value. The absolute allowance covers gradual underflow in the relevant operations. Dividing the computed positive dot product plus this allowance by a downward bound on $1 - \gamma_{2n}$ gives an upper bound on the exact positive product. Interval matrix products then combine the propagated radii with this error bound on the product of interval centers. We expand elementary interval operations by one representable value in the required direction, apply this expansion at every operation of a multistep nonnegative bound, and reject nonfinite enclosures.

To bound sine and cosine, we enclose π using dyadic arithmetic and Machin's identity, $\pi = 16\arctan(1/5) - 4\arctan(1/239)$, then reduce point arguments to a chosen quadrant using this interval. We evaluate the degree-17 Taylor polynomials there using interval Horner arithmetic. An enclosed reduced argument of magnitude at most one admits the remainder bound $1/18!$ for either function; otherwise, we use $[-1, 1]$. For a general argument interval, global one-Lipschitz continuity allows us to add its radius and intersect with the trigonometric range. Certificates use these explicit enclosures; training and empirical metrics use the system sine routine. The fixed precision and polynomial degree can leave a numerical floor; the termination result in Theorem 5 permits further arithmetic and approximation refinement when needed. Appendix H specifies how this arithmetic model enters the certificate computation.

E Compression and head repair

The local residual and feature models also guide head repair. Our joint objective retains their shared spatial dependence, while the interval-head baseline bounds them separately. Both proposals undergo the same verification after serialization.

E.1 JOINT REPAIR OBJECTIVE

With the decoded hidden body fixed, an output-head correction can cancel part of the residual through the same local models used for verification. The augmented features $\psi(x) = (\widehat{h}_L(x)^\top, 1)^\top$ have $m = n_L + 1$ entries, and the correction adds $\Delta w_{\text{head}}^\top \psi$ to g. Their shared-coordinate affine models combine into center $a + H^\top \Delta w_{\text{head}}$, slopes $b + J^\top \Delta w_{\text{head}}$, and remainder $e_g + \Delta w_{\text{head}}^\top e_\psi$. On a box, the maximum absolute affine value equals the absolute center plus radius-weighted absolute slopes. Bounding the remainder by the triangle inequality gives Equation (10). The auxiliary variables in Equation (18) express these bounds as linear constraints, so one correction minimizes the largest cell bound.

With *paired-head repair*, we optimize the center-error and slope terms jointly through Δw_{head}. One head must serve every spatial cell, so we index the local coefficients by cell B and minimize the largest cell bound through the following sparse linear program (LP):

$$\begin{aligned} \min_{\Delta w_{\text{head}}, p, q, z, t} \quad & t \quad \text{s.t. } p \geq \pm \Delta w_{\text{head}}, \;\; q_B \geq \pm(a_B + H_B^\top \Delta w_{\text{head}}), \\ & z_{Bi} \geq \pm\big(b_{Bi} + (J_B^\top \Delta w_{\text{head}})_i\big), \;\; q_B + r_B^\top z_B + R_{g,B} + R_{\psi,B}^\top p \leq t, \end{aligned} \tag{18}$$

for every cell $B \in \mathcal{P}$, with componentwise inequalities. Equation (18) expresses the absolute-value bound using auxiliary variables: $p \in \mathbb{R}^m$ bounds the absolute head correction, while $q_B \in \mathbb{R}$ and $z_B \in \mathbb{R}^d$ bound the corrected center error and slope magnitudes. The final constraint combines these terms with the remainder contribution $R_{\psi,B}^\top p$, making t a common upper bound over all cells. Minimizing t therefore minimizes the largest cell bound. The following construction accounts for coefficient uncertainty gives the interval-arithmetic details used to account for uncertainty in the repair coefficients.

The LP must account for uncertainty in its center and Jacobian coefficients. We enclose these coefficients by intervals with midpoints $\bar{a}, \bar{b}, \bar{H}, \bar{J}$ and outward radii s_a, s_b, s_H, s_J, absorbing their uncertainty into the remainder radii. For coordinate index $i = 1, \ldots, d$ and augmented-feature index $j = 1, \ldots, m$, the resulting radii are

$$\overline{R}_g = R_g + s_a + \sum_i r_i s_{b_i}, \qquad \overline{R}_{\psi,j} = R_{\psi,j} + s_{H_j} + \sum_i r_i s_{J_{ji}}.$$

These enlarged radii allow midpoint coefficients in Equation (18). Acceptance concerns the stored head, so we use the original coefficient intervals and remainder radii to certify $\Delta w_{\text{head}}^{\text{dec}}$. Interval subtraction encloses this decoded correction; propagating it through the coefficients and adding $R_\psi^\top |\Delta w_{\text{head}}^{\text{dec}}|$ bounds its residual contribution. We serialize the body from the reference coefficients using the candidate's codec settings and replace only the head, preserving both the body and byte count without requantization. The LP proposes coefficients; outward-rounded verification certifies the decoded head, accounting for parameter rounding and numerical error in the bound.

The zero correction is feasible, so exact optimization cannot increase the unrepaired joint-model bound. Another portfolio bound may be tighter, however; we therefore retain the unrepaired stream whenever its portfolio bound is smaller. Every repair method follows this rule, including least-squares and minimax heads fitted at 512 centers. Adaptive sampled minimax adds the 64 queries with the largest empirical residuals from each of three independent 4,096-point search pools and resolves its LP. The common final verifier isolates the repair objective: samples guide the proposal, while certification controls error throughout the domain.

E.2 INTERVAL-HEAD BASELINE

The interval-head baseline isolates the benefit of shared coordinates by bounding residuals and features independently on each cell. We fix the quantized body, retain an FP32 affine head, and enclose the residual on B by $[\ell_B, u_B]$. The augmented decoded-body features $\psi(x)$ lie in $[\mu_B - s_B, \mu_B + s_B]$, where $\mu_B, s_B \in \mathbb{R}^m$ describe feature ranges and c_B, r_B describe the coordinate cell. A correction adds $\Delta w_{\text{head}}^\top \psi(x)$ to the residual, yielding the interval objective

$$\begin{aligned} \min_{\Delta w_{\text{head}}, p, t} \quad & t \\ \text{s.t.} \quad & u_B + \mu_B^\top \Delta w_{\text{head}} + s_B^\top p \leq t, \quad -\ell_B - \mu_B^\top \Delta w_{\text{head}} + s_B^\top p \leq t \quad (\forall B), \\ & p \geq \Delta w_{\text{head}}, \quad p \geq -\Delta w_{\text{head}}, \quad p \geq 0, \quad t \geq 0. \end{aligned} \tag{19}$$

Algorithm 1 Certified compression

Input: Reference checkpoint θ, domain Ω, error tolerance ε, verification budget, finite codec options.

1. Retain a lossless copy of the reference checkpoint and metadata as a zero-error fallback.
2. Enumerate body precision and transform options; optionally solve Equation (18) for each body, retaining its reference-head candidate.
3. For each candidate s, decode the serialized bytes to $\widehat{\theta} = \mathcal{D}(s)$ and measure $R(s)$.
4. Compute the certificate portfolio and refine cells within the verification budget.
5. If $\max_{B \in \mathcal{P}} U(B) \le \varepsilon$, retain s when it uses fewer bits than the smallest certified stream found so far.
6. Return the smallest retained stream and its domain and error certificate.

In Equation (19), $p \in \mathbb{R}^m$ bounds the absolute correction componentwise, so the feature radii bound its cellwise contribution and t bounds the largest repaired error. With exact interval coefficients, the zero correction attains the unrepaired upper bound; exact optimization cannot increase it. After FP32 serialization and decoding, we intersect recomputed residual bounds with the unrepaired enclosure plus the verified correction. The same reference-head retention rule as for paired-head repair isolates the local objectives. Both methods optimize only the head, avoiding interval differentiation over all quantized parameters. The encoder then compares body precisions and transforms to select the smallest certified representation.

E.3 Certified compression procedure

The pseudocode for CertINR's compression procedure is provided in Algorithm 1.

F Geometric implications

The acceptance criterion bounds field-value error relative to the reference network. A separate reference-model error bound lets us translate it into physical clearance. Gradient bounds on a band inside the domain then yield topology and surface-displacement guarantees between the reference and decoded zero sets.

Sign and clearance. The bound $\|g_s\|_\infty \le \varepsilon$ preserves the sign at every point with $|f_\theta(x)| > \varepsilon$. Physical clearance also depends on how accurately the reference models the scene. We let s_{true} denote the physical signed distance function, positive in free space. Given a separate model-error certificate $\|s_{\text{true}} - f_\theta\|_\infty \le \eta$, we obtain

$$s_{\text{true}}(x) \ge f_{\mathcal{D}(s)}(x) - \varepsilon - \eta. \tag{20}$$

Equation (20) accounts for both compression and reference-model error. A finite-precision query introduces an additional error: if it returns $\widetilde{f}(x)$ with $|\widetilde{f}(x) - f_{\mathcal{D}(s)}(x)| \le \tau$, the clearance lower bound becomes $\widetilde{f}(x) - \varepsilon - \eta - \tau$. Clearance follows when this bound is positive on the swept set: all points occupied by the robot's full body during the motion. The model-error certificate η must link the reference INR to the physical scene.

To preserve topology, we control the deformation of the whole zero set. Theorem 8 uses a gradient gap on an interior band to keep the gradient nonzero during this deformation. We then transport the reference surface to the decoded surface and bound its displacement.

Theorem 8 (Preservation of a regular zero set)**.** *Let Ω be compact, and let f and $\widehat{f} = f + g$ be C^2 on a neighborhood of Ω. Define $K_a = \{x \in \Omega : |f(x)| \le a\}$. Assume this band is nonempty and lies in* $\operatorname{int}\Omega$*, and suppose*

$$\|g\|_{\infty,\Omega} \le \varepsilon < a, \qquad \inf_{x \in K_a} \|\nabla f(x)\|_2 \ge m > 0, \qquad \sup_{x \in K_a} \|\nabla g(x)\|_2 \le \kappa < m. \tag{21}$$

Set $S = f^{-1}(0) \cap \Omega$ and $\widehat{S} = \widehat{f}^{-1}(0) \cap \Omega$. If S is nonempty, the two zero sets are ambient isotopic by an isotopy supported in the interior of Ω. Their Hausdorff distance satisfies $d_H(S, \widehat{S}) \le \varepsilon/(m - \kappa)$.

Proof. We construct the ambient isotopy and bound its displacement in Appendix G. □

The interior-band and value-error conditions keep the zeros of $f + tg$, for $0 \le t \le 1$, away from the domain boundary. The gradient gap keeps its gradient nonzero on the band, allowing the surface to deform without splitting or merging. We check Equation (21) using the value certificate and gradient enclosures on the retained band. The residual Jacobian supplies the upper bound κ, while derivative intervals I_i enclosing $\partial_i f$ give the reference bound $\|\nabla f\|_2 \ge [\sum_i \operatorname{dist}(0, I_i)^2]^{1/2}$ on each cell; $\operatorname{dist}(0, I_i)$ is the interval's smallest absolute value. Cells whose reference range intersects $[-a, a]$ form the retained cover. We refine unresolved bounds and check boundary exclusion; training gradients toward unit norm (eikonal training) does not replace this uniform check.

Our interpolation uses the classical regular-level-set homotopy $F_t = f + tg$ and bounds its transport velocity. The gradient perturbation matters because a uniform field perturbation such as $\widehat{f}(x) = x - \varepsilon \sin(kx)$ can introduce extra zeros when $k\varepsilon > 1$. Uniform value control also gives a different form of stability. Points with field values at most a threshold form a sublevel set, and persistence diagrams record the thresholds where connected components and holes appear or disappear. For tame fields on a common compact triangulable domain, classical persistence stability yields $d_B(\mathrm{Dgm}(f), \mathrm{Dgm}(\widehat{f})) \le \varepsilon$ (Cohen-Steiner et al., 2007), where Dgm denotes a sublevel-set persistence diagram and d_B is the bottleneck distance between diagrams. Tameness requires finite-dimensional homology and finitely many homological changes. This persistence bound tracks changes across thresholds; Theorem 8 instead preserves the zero set up to ambient isotopy under its gradient assumptions.

The gradient condition in Theorem 8 compares the smallest reference gradient with the largest residual gradient, even if they occur in different regions. Corollary 9 instead compares gradient bounds cellwise, retaining the theorem's value-error and boundary hypotheses.

Corollary 9 (Cellwise regular-band check)**.** *Retain the domain, smoothness, nonemptiness, interior-band, and value-error hypotheses of Theorem 8. Let finitely many cells cover K_a. On every retained cell B, certify*

$$\|\nabla f\|_2 \ge m_B, \qquad \|\nabla g\|_2 \le \kappa_B, \qquad |g| \le U_B, \qquad \gamma_B = m_B - \kappa_B > 0.$$

Then the reference and decoded zero sets are ambient isotopic, and

$$d_H(S, \widehat{S}) \le \max_B \frac{\min\{\varepsilon, U_B\}}{\gamma_B} \le \frac{\varepsilon}{\min_B \gamma_B}.$$

Proof of Corollary 9. For $F_t = f + tg$, the reverse triangle inequality gives $\|\nabla F_t(x)\|_2 \ge \gamma_B$ on each retained cell. The finite cover therefore supplies a uniform positive gradient gap on K_a. Every zero during the interpolation satisfies $|f(x)| = t|g(x)| \le \varepsilon < a$. We choose $\varepsilon < b < a$ to leave room for a cutoff around these zeros. Since $K_b = \{x \in \Omega : |f(x)| \le b\}$ is a compact subset of the open set $\{x \in \operatorname{int}\Omega : |f(x)| < a\}$, we can choose a smooth spatial cutoff equal to one on K_b and supported in this open set. Multiplying $-g\nabla F_t / \|\nabla F_t\|_2^2$ by this cutoff and extending the product by zero outside the band gives the same interior ambient isotopy as in Theorem 8. We bound its speed along each zero trajectory by the stated maximum over cells and integrate from $t = 0$ to one to obtain the two-sided Hausdorff bound. □

The maximum $V = \max_B \min\{\varepsilon, U_B\}/\gamma_B$ in Corollary 9 bounds both transport speed and Hausdorff distance. Local gradient comparisons can certify a cover that fails the global check. Our geometry study also retains certified cells, refines unresolved cells, and narrows the band. Appendix J separates the effect of the local condition on a fixed cover from this combined refinement procedure.

G Proof of Theorem 8

To prove Theorem 8, we construct a velocity field that follows the zeros as we interpolate between fields. The gradient gap in Equation (21) bounds this velocity; the interior-band condition keeps the deformation away from the domain boundary.

Proof of Theorem 8. We interpolate between the fields by defining $F_t = f + tg$ for $t \in [0, 1]$. Every zero of F_t in Ω satisfies $|f(x)| = |tg(x)| \le \varepsilon < a$ and hence lies inside K_a, where

$$\|\nabla F_t(x)\|_2 \ge \|\nabla f(x)\|_2 - t\|\nabla g(x)\|_2 \ge m - \kappa > 0.$$

A smooth cutoff restricts the deformation to the band: we choose χ supported in the interior of K_a and equal to one on a neighborhood of $K_\varepsilon = \{x \in \Omega : |f(x)| \le \varepsilon\}$. Such a choice is possible because K_ε is a compact subset of $\{x \in \operatorname{int}\Omega : |f(x)| < a\}$. On $\{x \in \operatorname{int}\Omega : |f(x)| < a\}$, we define

$$V_t(x) = -\chi(x)\frac{g(x)\nabla F_t(x)}{\|\nabla F_t(x)\|_2^2}, \tag{22}$$

and extend it by zero outside the band. The time-dependent vector field in Equation (22) is C^1 in space and compactly supported in $\operatorname{int}\Omega$, so it generates ambient diffeomorphisms for $0 \le t \le 1$. Its numerator offsets the change in field value caused by interpolation. Along a trajectory starting on S, the cutoff initially equals one, so

$$\frac{d}{dt}F_t(x(t)) = g(x(t)) + \langle \nabla F_t(x(t)), V_t(x(t))\rangle = 0.$$

As long as the trajectory stays where $\chi = 1$, this identity keeps $F_t(x(t)) = 0$ and hence $|f(x(t))| \le \varepsilon$. Since $\chi = 1$ on a neighborhood of the compact set K_ε, a first-exit argument extends this property to all $t \in [0, 1]$. Applying the argument backward from each endpoint zero shows that every decoded zero comes from a reference zero. The flow therefore gives the required ambient isotopy while fixing a neighborhood of the domain boundary.

The velocity bound in Equation (22) also controls the distance traveled along each zero-level trajectory:

$$\|V_t(x(t))\|_2 = \frac{|g(x(t))|}{\|\nabla F_t(x(t))\|_2} \le \frac{\varepsilon}{m - \kappa}.$$

Integrating over a unit time interval bounds the distance between each matched pair of points. Because the flow is bijective, we obtain both directed Hausdorff bounds and hence the result. □

Why derivatives matter. We illustrate the need for derivative control with $f(x) = x$ and $\widehat{f}(x) = x - \varepsilon\sin(kx)$, whose field error is at most ε. When $k\varepsilon > 1$, the derivative of $\widehat{f}$ at zero is negative. Thus $\widehat{f}(x) < 0$ for sufficiently small positive x, whereas $\widehat{f}(\pi/k) = \pi/k > 0$. An additional positive zero lies in $(0, \pi/k)$, and oddness gives an additional negative zero. On a domain containing these points, arbitrarily small uniform error can therefore change the zero set when derivative control is absent.

Distance interpretation. If the reference equals the signed distance to a surface Γ, every decoded zero $x \in \Omega$ satisfies $\operatorname{dist}(x, \Gamma) = |f(x)| \le \varepsilon$. Here, Γ is the full surface. The bound alone does not control distance to the cropped set $\Gamma \cap \Omega$ or the reverse direction. Suitable domain coverage and sign crossings can provide those conclusions. Theorem 8 gives a two-sided bound without assuming an exact signed distance field. Its smoothness and nonzero-gradient conditions exclude medial-axis points, where distance functions can be nonsmooth.

H Certificate arithmetic

Numerical bounds must contain the exact quantities they estimate. We treat effective reference and decoded coefficients as exact real constants and round bounding operations outward, so implementation and analysis concern the same two fields. Under IEEE-754 binary64 round-to-nearest arithmetic, gradual underflow, and the matrix-product error model in Appendix D.5, interval operations include a γ_{2n} relative dot-product error bound and an absolute underflow allowance. An enclosure of Machin's identity bounds error during trigonometric range reduction; finite arguments beyond the 10^{12} cutoff receive the full range $[-1, 1]$. These bounds remain valid despite cancellation, underflow, or large phases.

The arithmetic model covers finite, real, scalar-output sine networks with matching architectures and parameter entries represented with at most 64 bits each. Each layer's output dimension must match the next layer's input dimension. The input box must have finite ordered endpoints in the network's input dimension, with point and face boxes permitted. When the complete effective coefficient arrays and architectures agree, we use the exact zero certificate; this identity check includes frequency scaling. These certificates concern the real-valued fields; Section 3 distinguishes them from floating-point inference.

Table 4: Mean minimum certified bytes across nine references after complete candidate search. All methods share the body pool, FP32 head, certificate portfolio, and reference-head retention rule. Adaptive minimax abbreviates adaptive sampled minimax. No selected stream uses the exact-copy fallback.

Method	$\varepsilon = .005$	.01	.02	.04
No repair	580.8	526.6	498.6	444.3
MSE	580.8	526.6	491.0	444.3
Sampled minimax	580.8	526.6	485.2	444.3
Adaptive minimax	573.2	526.6	479.4	442.6
Interval-head repair	571.4	519.0	483.4	444.3
Paired-head repair	**543.0**	**517.2**	**451.4**	**439.0**

I HEAD-REPAIR EXPERIMENTS

We test whether using the residual and decoded features at shared coordinates during head repair reduces certified storage. All methods use the same candidate pool, final verifier, and rule for retaining the reference head. We compare storage after a complete search, then examine encoding time, sampled errors, and independently recomputed bounds.

Experimental setup. We evaluate nine fixed 3–16–16–1 reference networks on CPU: sphere, torus, and rounded box, each with seeds 7, 19, and 31. Every method uses the same $\{6, 8, 10, 12\}^2$ body precisions, direct and layerwise DCT coding, an FP32 head, and an 8^3 spatial partition. The 288 reference–body combinations yield 1,728 candidate–method evaluations across six methods. At each tolerance, we average storage over the nine references, selecting the smallest certified stream for each.

The six methods are no repair, least-squares MSE repair, sampled minimax, adaptive sampled minimax, interval-head repair, and paired-head repair. Sampled methods start with the 512 cell centers. Adaptive sampled minimax draws 4,096 uniform coordinates per round, adds the 64 with the largest current residual, and solves its LP again for three rounds. It therefore searches 12,288 additional coordinates and ends with 704 optimization coordinates and four LP attempts per body. We then verify each proposed head over the continuous domain. Head optimization uses seed 1294; the separate error evaluation uses 32,768 held-out coordinates with seed 20260922.

All methods use the same certificate portfolio: recomputed paired Taylor, residual/Jacobian, and spectral bounds, together with bounds combining the unrepaired residual and head correction. We store and decode each proposed head before verification, retaining the unrepaired stream if its bound is smaller. After searching the pool, we select the smallest certified stream. The exact-copy fallback occupies 1,431 bytes and remains available throughout each search, although no complete search selects it at the four tested tolerances.

Repair outcomes and storage After complete candidate search at $\varepsilon = 0.02$, paired-head repair reduces mean selected storage by 5.84% relative to adaptive sampled minimax and by 6.62% relative to interval-head repair (Table 4). Against adaptive sampled minimax, it selects smaller streams on six checkpoints and ties on three. Mean storage savings relative to this baseline range from 0.80% to 5.84% across the four tolerances.

Both certificate-based repair methods reduce the bound used for selection on all 288 bodies. Sampled methods often retain the reference head because their proposals do not improve this bound: MSE reverts on 200 bodies, sampled minimax on 188, and adaptive sampled minimax on 169.

Paired outcomes. Table 5 counts the checkpoints on which paired-head repair uses fewer, equal, or more bytes than each baseline. Against adaptive sampled minimax, it reduces mean storage by 5.27%, 1.77%, 5.84%, and 0.80% at tolerances 0.005, 0.01, 0.02, and 0.04. The corresponding win/tie/loss counts are 5/4/0, 3/6/0, 6/3/0, and 2/7/0, showing how these savings vary with tolerance.

Table 5: Storage wins/ties/losses for paired-head repair against each baseline over nine checkpoints. A win means fewer certified bytes at the same tolerance. Adaptive minimax abbreviates adaptive sampled minimax.

Baseline	.005	.01	.02	.04
No repair	6/3/0	3/6/0	8/1/0	3/6/0
MSE	6/3/0	3/6/0	7/2/0	3/6/0
Sampled minimax	6/3/0	3/6/0	7/2/0	3/6/0
Adaptive minimax	5/4/0	3/6/0	6/3/0	2/7/0
Interval-head repair	5/4/0	1/8/0	5/4/0	3/6/0

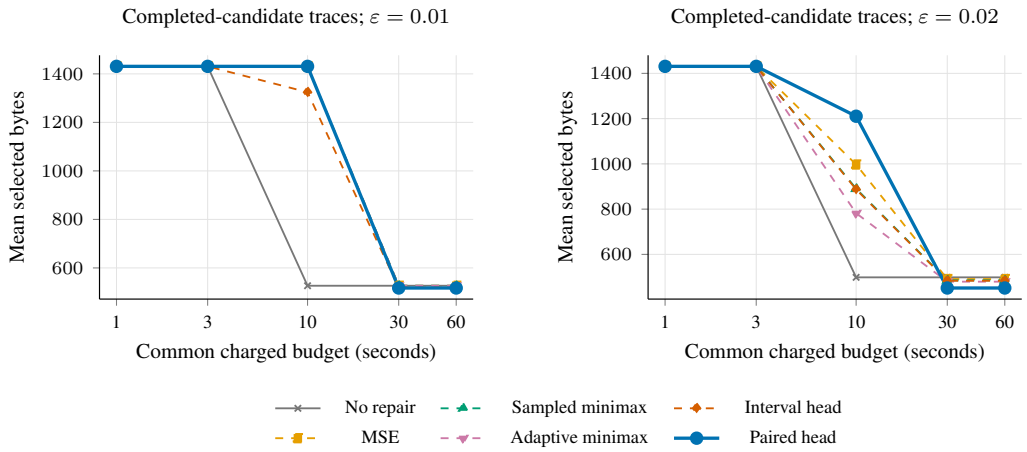


Figure 4: Mean smallest certified byte count available by each encoding deadline across nine references at $\varepsilon = 0.01$ and 0.02. Markers show evaluated deadlines; connecting lines aid comparison. The exact-copy fallback uses 1,431 bytes. Complete search favors paired-head repair; shorter deadlines favor faster methods. Times include preparation, head optimization, serialization, and verification.

Computation budgets. Storage at a fixed encoding deadline also depends on how quickly each method evaluates candidates. We process bodies by increasing byte count, break ties by transform and precision tuple, and rotate method order across bodies. Each method uses two BLAS threads. Timings include checkpoint and candidate-pool setup, shared per-body preparation, head optimization, serialization, and final verification. We charge shared preparation to each method to measure a standalone pipeline; this also charges no repair for feature-model preparation that an optimized implementation could omit. Held-out evaluation and result export are excluded.

Searching the full pool takes mean times of 9.40 seconds without repair, 16.25 for MSE, 16.65 for sampled minimax, 18.61 for adaptive sampled minimax, 16.71 for interval-head repair, and 26.65 for paired-head repair. Paired-head repair takes approximately 43.2% longer than the adaptive baseline. At a 10-second deadline and $\varepsilon = 0.02$, the mean selected sizes are 1,210.6 bytes for paired-head repair, 782.1 for adaptive sampled minimax, and 498.6 without repair. By 30 seconds, every method has evaluated all 32 candidates and reached its smallest certified stream in the pool. Thus the optimization cost outweighs the complete-search storage savings at this shorter deadline. Figure 4 uses the exact-copy fallback whenever no compressed candidate has been certified.

Head selection. A proposed head is retained only when its certificate is no larger than the unrepaired candidate's. MSE, sampled minimax, and adaptive sampled minimax retain the reference head on 200, 188, and 169 of the 288 bodies, respectively. Both certificate-based methods reduce the selection bound on every body. Because these decisions use the full portfolio, including bounds combining the residual and correction, we next check sampled errors and independently recomputed bounds.

Table 6: Compression errors on 32,768 held-out coordinates, comparing paired-head repair with adaptive sampled minimax on 18 equal-byte bodies from nine references. Smaller/equal/larger counts paired-head errors relative to the baseline. Ratios are paired/adaptive; values above one favor the baseline. These are empirical errors relative to the reference field.

Empirical metric	Smaller/equal/larger	Median paired/baseline
Held-out maximum	7/0/11	1.595
Held-out MSE	7/0/11	2.294

Table 7: Independently recomputed bounds for paired-head repair versus adaptive sampled minimax on 18 bodies from nine references. Smaller/equal/larger compares their upper bounds. Relative reduction is the baseline bound minus the paired-head bound, divided by the baseline bound; a positive median favors paired-head repair. Independent Taylor is our separate-remainder ablation.

Verifier	Smaller/equal/larger	Median reduction (%)
Paired Taylor	9/0/9	0.85
Residual/Jacobian	18/0/0	21.75
Independent Taylor	9/0/9	0.08
Spectral	11/0/7	12.24

Storage and fidelity. We compare sampled errors using the `direct_10-8` and `dct_10-8` bodies for every reference, giving 18 equal-byte comparisons across nine checkpoints. On both held-out maximum error and MSE, paired-head repair has larger errors than adaptive sampled minimax on 11 bodies and smaller errors on seven (Table 6). The median paired/adaptive ratios are 1.595 for the sampled maximum and 2.294 for MSE. Sampled maxima need not find the true worst-case error, but both metrics show that a smaller selection certificate can accompany larger sampled errors.

Rechecking repaired heads. We recompute bounds directly from each reference–decoded pair, excluding the bound combining the unrepaired residual and head correction. On the `direct_10-8` and `dct_10-8` bodies across nine references and six repair methods, we evaluate paired Taylor, residual/Jacobian, independent Taylor, and spectral bounds. Spatial methods use a new 16^3 partition; the spectral bound is global. The 108 method records contain 85 distinct reference–stream pairs because some methods retain the same unrepaired stream. We use the main study's checkpoints, streams, and outward-rounded arithmetic. Independent Taylor is our ablation that propagates the two remainders separately.

Under residual/Jacobian, the recomputed bounds favor paired-head repair over adaptive sampled minimax on all 18 bodies, with a median reduction of 21.75% (Table 7). Against the same baseline, paired Taylor gives nine smaller and nine larger bounds, with a median reduction of 0.85%; independent Taylor also gives nine smaller and nine larger bounds. Spectral bounds favor paired-head repair on 11 bodies and the adaptive baseline on seven, but favor both unrepaired and interval-head candidates over paired-head repair on all 18. Thus some benefits extend beyond the optimized LP bound, while others depend on the verifier. This comparison fixes two body configurations per reference; the complete search above measures storage savings.

J GEOMETRY EXPERIMENTS

We test whether gradient bounds certify topology preservation for the analytic fields. Comparing cellwise and global gradient gaps on the same cells isolates the effect of the condition. We then examine how refining selected cells and narrowing the surface band affect certification.

Models and refinement. We use the nine fixed 353-parameter reference networks on CPU: sphere, rounded box, and torus, with seeds 7, 19, and 31. Direct 12-bit coding of both hidden layers, with the reference FP32 head, gives the unrepaired pairs. The uniform-refinement comparison below uses these same pairs and their nine global value-error bounds.

Table 8: Cellwise topology certification on nine reference networks evaluated on CPU, with direct 12-bit bodies and FP32 heads. The band is $|f_\theta| \le a$. We retain certified cells and refine unresolved cells up to resolution 256. C certifies isotopy of the reference and decoded zero sets; U means unresolved. Best V is the smallest transport-based Hausdorff bound across the three bands; values above the domain diameter add no distance information.

Shape	Seed	$a = .05$	$a = .02$	$a = .01$	Best V
box	19	C	C	C	2.31
box	31	C	C	C	0.7683
box	7	C	C	C	3.478
sphere	19	C	C	C	0.002728
sphere	31	C	C	C	0.004623
sphere	7	C	C	C	0.02625
torus	19	U	C	C	25.34
torus	31	U	U	U	–
torus	7	U	U	U	–

We test band half-widths $a \in \{0.05, 0.02, 0.01\}$ from a 32^3 partition of $[-1, 1]^3$. A cell is excluded only if its reference-value interval misses $[-a, a]$. For each remaining cell, outward-rounded bounds give m_B, κ_B, and U_B in Corollary 9. We retain cells with a certified positive gap $m_B - \kappa_B > 0$ and bisect each unresolved cell along all axes into eight children. Certified and excluded cells need no further refinement. We stop at equivalent resolution 256, 500,000 evaluated volume cells, or 120 seconds per band, returning UNKNOWN if any band cell remains unresolved.

Acceptance also requires the value and boundary conditions in Corollary 9. We first certify global value error on 8^3 cells and a lower bound on the boundary margin, $\inf_{x \in \partial\Omega} |f_\theta(x)|$, using 32^2 cells on each of the six cube faces. Certified opposite signs among $\{(0, 0, 0), (1, 1, 1), (0.55, 0, 0)\}$ establish a nonempty reference zero set by continuity along a segment in the cube. We require $\varepsilon < a$, a boundary margin above a, and positive gradient gaps throughout the band, thereby checking all hypotheses for each accepted attempt.

Topology and distance. Table 8 reports six certified references out of nine at $a = 0.05$ and seven at each narrower band: 20 successful attempts out of 27. Torus seed 19 accounts for the additional reference. All seven unresolved attempts reach resolution 256 with uncertified gradient gaps; none reaches the time or cell limit. At the narrowest band, torus seeds 7 and 31 still have 3,285 and 2,946 unresolved cells, respectively, preventing topology certification with the current bounds.

On 11 of the 20 certified covers, $\min_B m_B - \max_B \kappa_B \le 0$ even though every cell has a positive local gap. Using the same cells and derivative bounds shows that the cellwise condition can succeed when the global condition fails. The refinement comparison changes several choices together: refining every retained band cell at $a = 0.05$ up to resolution 128 certifies six references. The increase to seven therefore combines the cellwise condition with selective refinement, narrower bands, and a higher resolution limit.

A certified positive gradient gap can still give a large displacement bound. Because we stop refining certified cells, the procedure targets topology rather than a specified Hausdorff distance. For the transport-speed and Hausdorff upper bound V in Corollary 9, the smallest certified values are 0.002728 and 0.004623 for sphere seeds 19 and 31. Torus seed 19 gives $V = 31.17$ at $a = 0.02$ and $V = 25.34$ at $a = 0.01$, both above the domain diameter $\sqrt{12}$. Its certificate establishes isotopy but adds no distance information. Several box bounds are also large, and uniform refinement gives smaller valid bounds in some cases. All bounds compare the decoded and reference neural zero sets.

Refinement cost. Running all attempts sequentially with two BLAS threads takes 328.40 seconds, including shared reference checks. Across 27 band attempts, we evaluate 2,104,208 volume cells; the largest attempt uses 273,272 cells and 40.64 seconds. These timings use the interval arithmetic in Appendix H.

Table 9: Reference fit on 32,768 surface samples per shape, weighted by triangle area. An exact signed distance field f would be zero on the source mesh. Normal cosine measures gradient alignment with mesh normals. These empirical diagnostics use the released normalized units and measure reference fit, separately from compression error.

Shape	Mean $\lvert f\rvert$	Maximum $\lvert f\rvert$	Median $\lVert\nabla f\rVert$	Mean normal cosine
Spot	0.00250	0.01011	1.0051	0.9993
Bob	0.00279	0.00685	1.0042	0.9998
Bust	0.00106	0.01661	0.9984	0.9924
Max	0.00290	0.00796	1.0009	0.9990
Bunny	0.00264	0.00697	1.0006	0.9980
Armadillo	0.00312	0.00962	1.0000	0.9905

K PRETRAINED SHAPE EXPERIMENTS

We next test value-error certification and storage on released pretrained fields with larger architectures and higher sine frequencies. Every verifier receives the same reference–decoded pair. This study does not include the additional topology checks of Appendix J.

Pretrained models. We use six static neural signed distance fields released by Buonomo et al. (2025), whose source meshes appear in Figure 5. The released morphing configurations define three pairs: Spot/Bob, Bust/Max, and Bunny/Armadillo, with 4,481, 17,153, and 83,201 parameters per model, respectively. We selected these pairs to cover different architecture sizes before inspecting certification results, and use their checkpoints without fine-tuning or retraining.

Frequency and parameters. The released models store binary32 parameters and use frequency 60 in every sine layer. We absorb this multiplier into the effective coefficients used by the verifier. The products are exactly representable in binary64, so this conversion preserves the reference function in real arithmetic.

Frequency-aware decoding. The codec stores unscaled parameters with a 12-byte frequency header in each direct/DCT stream. We quantize all hidden layers at the same precision and retain the reference FP32 head. The decoder reconstructs the parameters and applies the frequency in binary64; its resulting coefficients include inverse-DCT and multiplication rounding. The exact-copy candidate stores the original FP32 parameters and frequency, reproducing the reference effective coefficients. Its sizes are 17,955, 68,643, and 332,851 bytes for the three architecture pairs. As in Section 3, the certificate concerns the decoded real-valued function; floating-point inference error requires separate control.

Domain and reference fit. All mesh coordinates lie in $[-1, 1]^3$, our verification domain; we retain the released coordinates without rescaling. Table 9 evaluates reference fit on points sampled from the mesh triangles. The surface residual is the reference field value at a mesh point, where an exact signed distance field would be zero. The normal cosine measures alignment between its gradient and the mesh normal. These diagnostics measure agreement with the source geometry; compression certification bounds the reference–decoded difference throughout the cube.

Experimental setup. Direct and layerwise orthonormal DCT coding at body precisions $\{12, 16, 20, 24\}$ yield 48 reference–stream pairs. We chose this grid from the known frequencies and architecture sizes before computing certificates. Each pair receives spectral, residual/Jacobian, and paired Taylor bounds; the spatial methods share a refinement schedule. We compare the paired Taylor bound with the minimum of the spectral and residual/Jacobian bounds. Each method receives 30 seconds, giving the portfolio up to 90 seconds per pair, and spatial refinement stops at 512 cells. Retaining reference heads isolates verification from head optimization.

Compression results. Spectral bounds certify compressed streams for all six models at tolerance 0.01 (Table 11). They account for every portfolio acceptance at the tested tolerances; spatial methods

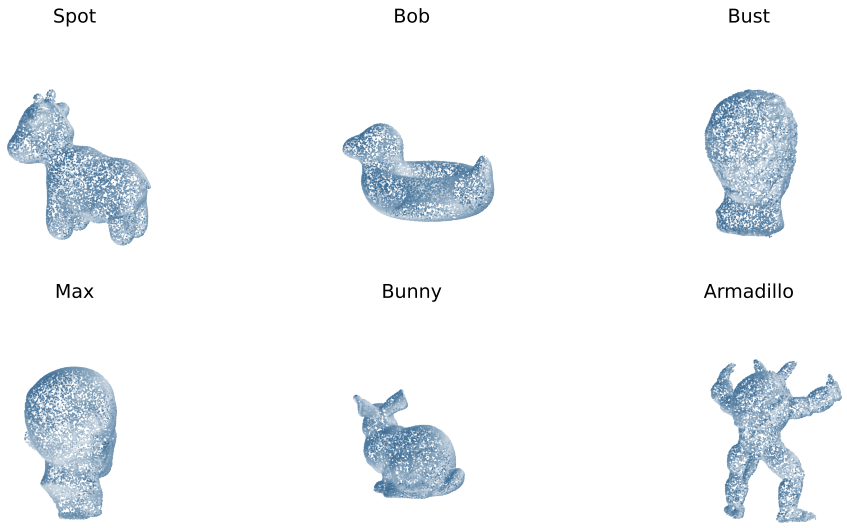


Figure 5: The six source meshes, visualized by samples from their triangles. Their pretrained fields provide the fixed reference networks for compression. Source: Buonomo et al. (2025).

Table 10: Pretrained outcomes out of 48 streams: certified / violated / unknown. Violated requires a certified pointwise error above tolerance; unknown means neither certification nor a proved violation. The portfolio combines the three methods under separate time limits.

Method	$\varepsilon = .001$	$\varepsilon = .003$	$\varepsilon = .01$	$\varepsilon = .03$
Spectral	17 / 0 / 31	22 / 0 / 26	28 / 0 / 20	31 / 0 / 17
Residual/Jacobian	16 / 12 / 20	20 / 3 / 25	23 / 0 / 25	25 / 0 / 23
Paired Taylor	13 / 12 / 23	18 / 3 / 27	20 / 0 / 28	25 / 0 / 23
Portfolio	17 / 12 / 19	22 / 3 / 23	28 / 0 / 20	31 / 0 / 17

additionally prove violations at selected points (Table 10). On all 48 pairs, the minimum of the spectral and residual/Jacobian bounds is smaller than the paired Taylor bound. This reversal from the shallow analytic results supports retaining several bounds in the portfolio. The recurrence in Appendix D explains how architecture can affect the remainder, but this comparison does not isolate any single architectural factor (Section 7).

Computational cost. Across 144 method–pair evaluations, median total times are 0.0288 s for spectral, 1.0476 s for residual/Jacobian, and 0.7706 s for paired Taylor; maxima are 0.1535, 3.2886, and 2.5119 s, respectively. Each timing includes preparation and the completed refinement schedule, with one run per method–pair evaluation.

L Official verifier comparisons

We compare paired Taylor with official general-purpose verifiers on the same compression residuals. The standard graph passes shared coordinates through both networks and subtracts their outputs. We test this graph on analytic and pretrained pairs, then rewrite the analytic graphs to expose layerwise cancellation and measure its effect on the external bounds.

Benchmarks and budgets. We evaluate GenBaB (Shi et al., 2025), CROWN (Zhang et al., 2018), and α-CROWN (Xu et al., 2021) on two fixed benchmarks. The nine analytic 353-parameter references use unrepaired direct 10/8-bit bodies and FP32 heads at $\varepsilon = 0.02$. The six pretrained fields

Table 11: Selected streams at $\varepsilon = 0.01$, including all metadata; byte-count ties use a fixed ordering. Spectral bounds certify every stream. Mean per-shape savings are 41.21%; total storage decreases from 838,898 to 587,154 bytes, a 30.01% reduction.

Shape	Body bits	Exact bytes	Selected bytes	Saving (%)	Upper bound
Spot	16	17,955	9,139	49.10	0.001969
Bob	16	17,955	9,139	49.10	0.001595
Bust	16	68,643	34,611	49.58	0.005979
Max	16	68,643	34,611	49.58	0.004517
Bunny	24	332,851	249,827	24.94	0.001061
Armadillo	24	332,851	249,827	24.94	0.005682

in Appendix K use direct 16-bit bodies and FP32 heads at $\varepsilon = 0.01$. All methods receive the same function and tolerance for each pair.

Each method uses two CPU threads, a 90-second process limit including startup, and a 4 GiB limit on periodically sampled resident memory. GenBaB searches each signed query with its default branching heuristic, uniform split points, and a nominal 30-second computation limit. CROWN and α-CROWN bound one input box within 30 seconds; α-CROWN uses 20 optimization iterations. Paired Taylor receives 30 seconds for grids with 1, 2, 4, and 8 subdivisions per axis, up to 512 cells. GenBaB thus receives two signed-query allowances per pair; the other methods bound both signs in one computation. We count certificates only within the query or pair limit and report computation time separately from process time including startup. Section 6.2 summarizes results under these different time allowances and subdivision schedules.

Matched functions. Each standard graph passes the same three-dimensional input through the reference and decoded sine networks, returning $g = f_{\mathcal{D}(s)} - f_\theta$ on $[-1, 1]^3$. We store their actual coefficients as binary64 constants, including frequency scaling for pretrained models, so every verifier receives the same function.

Certificate targets. We ask the external verifiers whether $g \geq \varepsilon$ or $g \leq -\varepsilon$ occurs, using their VNNLIB input format. Two UNSAT (unsatisfiable) reports would rule out both events and establish $|g| < \varepsilon$ under the solvers' numerical conventions. Our target permits equality, $|g| \leq \varepsilon$, so we confirm violations only when outward-rounded point evaluation proves an absolute residual above tolerance. External solvers retain their float64 arithmetic; our verifier uses outward-rounded bounds.

Graph representation. The residual-graph comparison in Appendix M exposes layerwise cancellation while keeping the external algorithm fixed. Both CROWN representations use one box, whereas paired Taylor subdivides; comparisons between these methods therefore combine representation and subdivision effects. Related domain-dependent bounds include neural-implicit affine arithmetic (Sharp & Jacobson, 2022) and HiTaB (Entesari & Fazlyab, 2026). Our independent Taylor method is an ablation that propagates each network's remainder separately.

Analytic results. Paired Taylor certifies two of nine analytic pairs, proves one violation, and leaves six unresolved. Neither CROWN variant certifies a standard-graph pair; GenBaB identifies one violation and leaves eight unresolved (Table 12). The residual-graph comparison below tests whether exposing layerwise cancellation reduces these external bounds.

We confirm the GenBaB counterexample for torus seed 7 with the residual enclosure

$$[0.020936832830204456,\ 0.02093683283021751],$$

whose lower endpoint exceeds 0.02, proving a violation. Median computation times per pair are 0.0211 s for CROWN, 1.0935 s for α-CROWN, and 0.1194 s for paired Taylor. GenBaB's two signed queries sum to a median of 61.63 s; its deadline checks occur between iterations and can exceed the nominal 60-second pair allowance.

Pretrained results. Paired Taylor certifies Spot and Bob among the six pretrained pairs, leaving four unresolved. Standard-graph external methods leave all six unresolved; none of these methods

Table 12: Nine analytic direct 10/8-bit pairs at $\varepsilon = 0.02$. Certified counts certificates completed within the computation time limit; Unknown counts unresolved pairs. Standard external graphs have two sine-network branches sharing one input; Table 14 tests residual graphs. Paired Taylor certificates and confirmed violations use outward-rounded arithmetic.

Method	Certified	Violated	Unknown	Fail/missing
GenBaB	0	1	8	0
CROWN	0	0	9	0
α-CROWN	0	0	9	0
Paired Taylor	2	1	6	0

Table 13: Six pretrained direct 16-bit pairs at $\varepsilon = 0.01$. External rows report float64 solver outcomes; paired Taylor uses outward-rounded bounds. Rows below the rule show matching functions from the separately timed pretrained benchmark. All rows use identical references and decoded streams.

Method	Evaluation	Certified	Violated	Unknown
GenBaB	Official comparison	0	0	6
CROWN	Official comparison	0	0	6
α-CROWN	Official comparison	0	0	6
Paired Taylor	Matched paired run	2	0	4
Spectral	Pretrained benchmark	4	0	2
Residual/Jacobian	Pretrained benchmark	3	0	3
Portfolio	Pretrained benchmark	4	0	2

proves a violation at $\varepsilon = 0.01$. Median computation times are 0.0287 s for CROWN, 1.2582 s for α-CROWN, and 0.8019 s for paired Taylor. GenBaB's two signed queries sum to a median of 64.05 s, again reflecting its deadline checks.

Table 13 also includes matching direct 16-bit pairs from the separate pretrained benchmark: spectral certifies four and residual/Jacobian certifies three, giving four portfolio certificates. Both studies use the same reference fields, decoded streams, and tolerances, but the timings above come from runs matched to the external comparison. Paired Taylor certifies the same two pairs in both studies (Appendix K).

M RESIDUAL-GRAPH COMPARISON

Nearly identical networks can vary substantially across the domain while retaining a small difference, which separate bounds on their branches can miss. We rewrite their difference as an exact residual graph and compare both graphs using the same CROWN and α-CROWN implementations on all nine analytic direct-10/8 pairs. This isolates the effect of graph representation within each external method.

The residual graph carries the reference activation h together with the difference $e = \widehat{h} - h$, starting from $h = x$ and $e = 0$. We propagate them through each hidden layer using the effective coefficients of Equation (3):

$$z = Ah + a, \qquad \zeta = \widehat{A}e + \Delta A h + \Delta a, \tag{23}$$

$$e^{+} = 2\sin(\zeta/2) \odot \cos(z + \zeta/2), \qquad h^{+} = \sin(z), \tag{24}$$

where $\Delta A = \widehat{A} - A$ and $\Delta a = \widehat{a} - a$; the superscript plus denotes the following layer. Equation (23) forms the exact preactivation difference, and the sine-difference identity in Equation (24) preserves $e^{+} = \widehat{h}^{+} - h^{+}$. This relation holds by induction throughout the network, so the final affine head gives

$$g = \widehat{A}_{L+1}e + \Delta A_{L+1}h + \Delta a_{L+1}.$$

The rewritten graph equals the original shared-input difference in real arithmetic. The implemented graph preserves this equality when each binary64 parameter difference equals the exact difference

Table 14: Graph representation comparison on nine analytic pairs using one input box. Ranges span the smallest to largest upper bounds on absolute compression error. Median reduction is the per-pair percentage decrease from standard to residual graph. Certified counts pairs certified at $\varepsilon = 0.02$ within 30 seconds. Displayed ranges are rounded; decisions use unrounded endpoints.

Method	Standard range	Residual range	Median reduction	Certified
CROWN	4.654–6.789	0.0983–0.2273	97.15%	0/9
α-CROWN	3.597–5.962	0.0380–0.0800	98.58%	0/9

of its source constants; 2 and $1/2$ are exactly representable. All nine pairs satisfy this condition. Floating-point outputs can still differ because the graphs order their operations differently.

Both representations use the full box $[-1, 1]^3$ at tolerance 0.02 under the limits in Appendix L: two CPU threads, 30 seconds of computation time, a 90-second process limit including startup, and a 4 GiB limit on sampled resident memory. Only α-CROWN uses 20 optimization iterations with learning rate 0.1. Both graphs use one global box and the external solvers' float64 bounds.

Effect of representation. The residual graph reduces both external bounds on every analytic pair, with median reductions of 97.15% for CROWN and 98.58% for α-CROWN (Table 14). Although still above $\varepsilon = 0.02$, these bounds show the effect of exposing cancellation. Paired Taylor gives smaller bounds on the same nine pairs, but also subdivides the domain, so that comparison includes both bounding and refinement effects.

Residual-graph computations take 0.0509 to 0.0873 s for CROWN and 3.332 to 4.082 s for α-CROWN. Paired Taylor gives upper bounds from 0.01420 to 0.03173, certifying two pairs, proving one violation, and leaving six unresolved.

N Analytic reference training

We train each analytic reference once, then fix its exported FP32 parameters throughout verification and repair. The errors below measure fit to the analytic target; compression certificates compare the decoded field with this fixed reference.

Data and initialization. Training targets are analytic signed distances to a sphere of radius 0.6, a torus with major radius 0.55 and minor radius 0.22, and a rounded box with half-extents $(0.5, 0.4, 0.3)$ and rounding radius 0.08. We sample coordinates uniformly from $[-1, 1]^3$. Initialization frequency ω_0 sets the first-layer scale: effective weights are uniform on $[-\omega_0/3, \omega_0/3]$. Later hidden layers with input width n use $[-\sqrt{6/n}, \sqrt{6/n}]$; output weights use one third of that scale, and biases are zero. We optimize and store effective coefficients, so $\omega_\ell = 1$ during evaluation and ω_0 controls only initialization.

Each update draws 512 independent coordinates and uses float64 Adam with $(\beta_1, \beta_2) = (0.9, 0.999)$ and denominator stabilizer 10^{-8}. The learning rate decreases linearly from approximately 10^{-3} to $2.5 \cdot 10^{-4}$. Small fields train for 6,000 steps; width-32 and width-64 fields for 4,000; and width-256 fields for 1,800. The nine small checkpoints use width 16, depth 2, and initialization frequency 3, with seeds 7, 19, and 31 for each shape. Four larger sphere checkpoints use seed 7 with width/hidden-depth/initialization-frequency settings $(32, 3, 3)$, $(64, 4, 3)$, $(256, 4, 3)$, and $(256, 4, 30)$. We export each checkpoint to binary32 before compression and evaluate its target fit (Table 15). The high-frequency model has a larger maximum fit error than its low-frequency counterpart.

Table 15: Analytic reference models. Test errors use 32,768 held-out uniformly sampled coordinates; maxima are empirical. Frequencies describe initialization of effective first-layer weights.

Shape	Seed	Parameters	ω_0	Test MSE	Test maximum
box	19	353	3	9.66e-04	0.152
box	31	353	3	1.56e-03	0.192
box	7	353	3	1.24e-03	0.204
sphere	19	353	3	2.98e-04	0.231
sphere	31	353	3	2.36e-04	0.222
sphere	7	353	3	2.71e-04	0.228
sphere	7	198,657	3	1.69e-05	0.092
sphere	7	198,657	30	1.28e-04	0.791
sphere	7	2,273	3	1.28e-04	0.189
sphere	7	12,801	3	1.55e-05	0.104
torus	19	353	3	1.07e-03	0.249
torus	31	353	3	2.22e-03	0.331
torus	7	353	3	1.43e-03	0.283