Exactly Certified Dispersive Functionals from Automated Search: An Application to Gravitational Positivity Bounds

Matthew Ruckman
Unaffiliated
mruckman1@gmail.com ORCID: 0009-0002-1723-3823

July 2026

Abstract

Dual dispersive positivity bounds on gravitational effective field theories are normally obtained from hand-designed functionals whose optima are decided by floating-point semidefinite programming up to solver tolerance. We present a pipeline in which the functional design is performed by an evolutionary search (structural mutations proposed by a large language model) whose *sole selection pressure is an exact-arithmetic certificate*: a candidate counts only if per-spin positivity of its smeared spectral density is proven in exact rational arithmetic on the mass continuum, for every audited spin. A controlled ablation finds no search advantage for the language model over conventional evolutionary or random search on this task, so we treat it as the search's mutation operator. The identical machinery reproduces the known scalar EFT-hedron with exact *all-spin* certificates. For the R^4 coefficient g_0 of $D = 10$ maximal supergravity we obtain exact certificates for spin-truncated dual functionals—3.0136 ($J \leq 40$) and 3.0708 ($J \leq 120$) on the reference domain of Ref. [4], and 3.0401 ($J \leq 120$) in the extended-u domain that reference introduced but left uncomputed—which become bounds on g_0 only conditional on tail positivity beyond the audited spins (measured on extensive scans; not proven). An intermediate $J \leq 40$ truncation certifies 2.96514, numerically below the reference's 3.000; we show by direct measurement that this is truncation slack and does not survive tail hardening. Systematic attempts to certify below 3.000 along four independent routes all terminate at the same tangency wall near 3.0; we present that convergence, together with the failure modes an exact standard exposes (tolerance-decided constraints at the 10^{-10} level; grid scans passing functionals whose exact density is wrong by seven orders of magnitude), as the principal measured content of this methods-and-rigor study.

1 Introduction

Positivity bounds turn the consistency of a quantum field theory—unitarity, analyticity/causality, and boundedness in the Regge limit—into inequalities on the Wilson coefficients of its low-energy effective field theory (EFT) [1, 3, 5]. A coefficient outside the allowed region cannot descend from any consistent ultraviolet completion. For gravitational EFTs the program is historically framed around a sharp physical question: how much room does consistency leave for a quantum gravity that is *not* string theory? In the $D = 10$ maximally supersymmetric graviton sector the leading higher-derivative coefficient g_0 (the R^4 coupling) is bounded, while the type II superstring realizes a specific value. We stress at the outset that a *positivity-only* upper bound on g_0 is a *loose* probe of this question: the current literature closes most of the gap between the plain-positivity ceiling and the string value by incorporating the upper bound from unitarity of the partial-wave imaginary parts [16] and low-spin dominance [17], and the tree-level primal bootstrap finds the string strictly inside a small allowed region [18]. Our contribution to this sector is therefore the *rigor* of the positivity bound—replacing tolerance-decided inequalities with exact certificates (all-spin for the scalar validation, exact for the audited spins in the gravitational sector)—not a step toward the physically-relevant, near-string answer, which those complementary methods already constrain far more tightly. Throughout, g_0 is the standard dual-side coefficient in the maximal-supergravity amplitude $8\pi G/(stu) + g_0 + g_2(s^2{+}t^2{+}u^2) + \dots$, the same object bounded by the reference and by the crossing-symmetric dispersion literature [4, 20, 21], so our comparison to the reference is a genuine same-object comparison.

The problem is attacked from two sides. The *primal* S-matrix bootstrap constructs explicit unitary, crossing-symmetric amplitudes and maps the allowed region from the inside; for exactly this coefficient in $D = 10$ it finds a minimal allowed value—in its normalization $\alpha \gtrsim 0.14$, with the type II string sitting essentially at that edge, $\alpha_{\text{string}} \approx 0.1389$—very close to but not equal to the string value [6], with extensions to other dimensions and to M-theory [7] and bounds on the *next-to-leading* Wilson coefficient in seven dimensions [8]. The *dual* approach constructs dispersive functionals whose non-negativity proves *upper* bounds from the outside, in a different normalization (the reference's g_0, with the string at $2\zeta(3) \approx 2.4041$) [4]. Our work is on the dual side: an exact-certified pipeline that bounds the *same physical coefficient*—the leading R^4 correction to $D = 10$ maximal supergravity—from the outside, with g_0 defined on the supersymmetry-reduced amplitude as above. We stress that this is a companion in *target and method*, not in directly comparable numbers: the fixed normalization map between the primal α and the dual g_0 is not worked out here—nor, as far as we could determine, anywhere in the published literature—so we do not place the two results on a common axis (Sec. 8).

Two features of dual practice motivate the method. First, the functionals—linear combinations of sum rules, smearing profiles, and null constraints—are designed by hand. Second, the numbers are typically decided by floating-point semidefinite programming, up to a solver tolerance. We address both: an evolutionary search designs the functionals (its structural mutations proposed by a large language model), and the acceptance criterion is an exact-arithmetic proof. The generative model proposes candidates; it is never an oracle for correctness. We emphasize that we do *not* claim the language model outperforms a conventional evolutionary or greedy search on this problem: a controlled ablation (Sec. 7) finds no such advantage—conventional random-restart and genetic-algorithm searches, given the identical evaluator and budget, match the model at the narrower scope and, as the best of twelve runs, reproduce its headline value to within 10^{-3} at the $J \leq 40$ scope—and our wall-convergence experiments (Sec. 6) indicate that, within this basis class, the certified frontier is set by structural tangencies no mutation operator moved. The model should therefore be read as the mutation operator of the search, and the contribution as the certified pipeline around it.

Our contributions are (i) a *method*—certified dual functional design, validated by reproducing the scalar EFT-hedron with exact all-spin certificates; (ii) a *demonstration*—exact certificates for spin-truncated dual functionals targeting the $D = 10$ coefficient: 3.0136 ($J \leq 40$) and 3.0708 ($J \leq 120$) on the reference's own domain, and 3.0401 ($J \leq 120$, with strict impact-parameter margins) in the extended-u domain the reference introduced but left uncomputed—bounds on g_0 only conditional on tail positivity beyond the audited spins—together with the diagnosis that an intermediate sub-3.000 truncation value (2.96514 at $J \leq 40$) is truncation slack, priced by direct measurement (Sec. 5.2); and (iii) a set of *rigor findings* a tolerance-free treatment surfaces—that the reference-basis fine constraints are, in our reconstruction, tolerance-decided at the reference's own spin cutoff, that a "tangency tax" is the true barrier to tighter dual bounds, and that in the extended domain a grid scan can pass a functional whose exact spectral density violates positivity by seven orders of magnitude (the corner-window failure mode, Sec. 6).

We are careful about what is not claimed. The numbers are modest and axiom-conditional where the domain is extended; none is claimed as an improvement on the reference's 3.000, and the one value below it is diagnosed as a truncation artifact. Whether the certified-design method is new we do not assert beyond the literature comparison summarized in Sec. 9; and a controlled ablation (Sec. 7) finds *no* search advantage for the language-model proposer over a conventional evolutionary or random-restart search on this task, so we present the language model as the search's mutation operator, not as a demonstrated improvement over automated alternatives. The normalization reconciliation with the primal program's coefficient remains open.

2 Setup

2.1 Dispersive sum rules and functionals

For a crossing-symmetric $2 \to 2$ amplitude, fixed-transfer dispersion relations give sum rules $C_{k,u}$ equating a low-energy (EFT) side—a polynomial in the Wilson coefficients—to a high-energy (spectral) side: a sum over states weighted by non-negative spectral densities and Gegenbauer partial waves P_J. At $u < 0$ the Gegenbauer factors oscillate in sign, so an individual sum rule's spectral side is *not* manifestly non-negative. A dispersive functional is a linear combination of such sum rules across k, across smearing kernels, and including *null constraints* that vanish identically on the EFT side, chosen so that the combination's action on *every* state—each spin J and each mass $m \geq M$—is non-negative; that per-state positivity is what turns unitarity into a bound. The design problem is which sum rules, which smearing profiles $f(p)$, and which nulls to combine; this is what we automate.

For the graviton the structure is specific and was verified against the reference before any search [4]: the graviton pole hits exactly one sum-rule cell, ($k = 2$, Laurent u^{-1}), with coefficient $-8\pi G$; the contact terms are dispersively invisible; and because that cell is off the Taylor lattice, C_2-family functionals must be smeared, $f(p)$ integrated over $p \in (0, p_{\max}]$ with the improved kernel C_2^{imp} whose EFT side is exactly $8\pi G/(-u) + 2g_2 - g_3 u$. The coefficient g_0 itself enters through a second smearing family, the improved C_0 tower, whose EFT side carries g_0 with unit weight; a profile p^i on $(0, p_{\max}]$ contributes $p_{\max}^{i+1}/(i+1)$ to the functional's g_0-coefficient γ_0, while the C_2^{imp} profile sets the graviton-pole coefficient γ_G. Every certificate in this paper is an instance of one displayed identity: with the functional rescaled so $\gamma_0 = -1$ exactly and the g_2, g_3 content projected out exactly by null rows, the EFT side is $\gamma_G\, 8\pi G - g_0$, so per-state non-negativity of the spectral side proves $g_0 \leq \gamma_G\, 8\pi G/M^6$ with γ_G the exact rational quoted (App. A). An $m \to \infty$ closure imposes the impact-parameter positivity $\hat{f}(b) \geq 0$. These identities are checked symbolically (Sec. 3, App. B).

2.2 The target and its axioms

Maximal supersymmetry relates the coefficients, reducing the graviton EFT to a small independent set of which g_0 is the simplest. Our certificates carry two assumptions beyond the unitarity/analyticity core, stated wherever a number is quoted:

(A1) the Regge axiom $s^2\, \tilde{\mathcal{M}}(s, u) \to 0$ as $|s| \to \infty$ at fixed $u < 0$, stated on the supersymmetry-*reduced* amplitude $\tilde{\mathcal{M}}$ (the graviton amplitude with the maximal-susy prefactor stripped; it is $\tilde{\mathcal{M}}$ whose low-energy expansion is $8\pi G/(stu) + g_0 + \cdots$ above), not on the full amplitude;

a strengthening of the classical-Regge-growth bounds [10, 11] in the lineage of the causality-vs-higher-derivative program [2, 9], and present in *every* $D = 10$ g_0 bound, ours and the reference's alike;

(A2) the extended-u meromorphy-and-discreteness axiom, needed only for the extended-u bounds ($p_{\max} > 1$), which smear over a larger analytic region.

Axiom (A2) requires comment. Meromorphy is not an idiosyncratic assumption we invented: it is the field's standard *weak-coupling* axiom—the amplitude is meromorphic precisely when the UV completion is weakly coupled in $M_{\rm pl}$, i.e. tree-level—stated as such in the reference itself (weak coupling is defined via a tree-level family in its §2.1, and meromorphy is invoked explicitly in its App. B [4]) and in the perturbative-bounds literature [16, 6]. This legitimizes (A2) but also relocates the extended-u result: it is a *tree-level* (weakly coupled) bound, exactly the class in which the *primal* tree-level bootstrap already maps the allowed region for this coefficient [18]. Two honest consequences follow. First, (A2) is genuinely load-bearing and stronger than the core axioms, so the extended-u number is not a fully non-perturbative swampland exclusion; the minimal-assumption, non-perturbative value we certify is the $p_{\max}$=1 value (Sec. 5). Second, the string lives in exactly this weak-coupling class and sits essentially at the *primal* lower edge of the allowed region [6, 18], while remaining $\sim$20% below every plain-positivity ceiling here ($2\zeta(3) \approx 2.404$ against 3.0–3.04)—so a plain-positivity ceiling in this class measures the slack positivity alone leaves, slack that the unitarity upper bound, low-spin dominance [16, 17], and the tree-level primal map [18] already largely close, rather than physical "room for non-string quantum gravity." We therefore treat the extended-u result as *conditional on (A2)* and make no universal-swampland claim for it. The domain itself is the reference's own construction: Appendix B of [4] introduces precisely this extended-u region under the meromorphy assumption, computes only a $D = 6$ scalar example there, and explicitly leaves the gravitational case to future work—that gravitational case is what we compute. The enlarged-domain crossing-symmetric program [21] has since recomputed the supergravity bound via crossing-symmetric dispersion relations, finding $g_0 M^6/8\pi G \leq 2.97$ on the smearing region $p \leq M$—equivalent to the standard $u \in (-M^2, 0]$ domain—and presenting it explicitly as matching the fixed-u optimum 2.969 of [18] rather than improving it; its new bounds are for $D = 4$ graviton amplitudes, and no extended-u ($|u| > M^2$, meromorphy-based) g_0 bound is computed there (Sec. 9).

The reference targets we validate against are: the scalar single-null optimum $\tilde{g}_3 \geq -(9/2 + (7/4)\sqrt{61/5}) = -10.612487218800544$ and the box $0 \leq \tilde{g}_4 \leq 1/2$ [3]; and the $D = 10$ value $0 \leq g_0 \leq 3.000\, 8\pi G/M^6$ [4].

3 Method

The pipeline has three layers: an evolutionary search, an exact-arithmetic certifier that is the sole source of selection pressure, and an assertion battery that gates the physics kernels.

3.1 Evolutionary search

A candidate is a column structure: which sum-rule and null families to include and at which smearing powers, with the domain endpoint $p_{\max}$ and a solver policy (coefficient cap, robustness margin, refinement depth). The search is a fork of an evolutionary code [24, 25] (island archive, fitness-plus-novelty parent selection, distributed evaluation); a large language model proposes the structural mutations, and an adversarial judge from a different model family pre-filters proposals but is never the source of truth and fails open. The fitness is *certified-or-nothing*: a monotone function of the certified bound, calibrated so any certificate beats any refusal, with a discount for narrower spin scope. The search cannot game the fitness—every point is a theorem or it is nothing. As noted above, we make no claim that the language model beats a conventional search on this task; the controlled ablation of Sec. 7 finds it does not.

3.2 The exact-arithmetic certifier

A floating-point linear program (HiGHS [41]) shapes a candidate functional by minimizing the g_0-ceiling subject to positivity on a grid of (m^2, J), the normalization, and—for the susy sector—exact null projections enforcing $g_2 = g_3 = 0$. This float solution certifies nothing. It is converted to exact rationals by a zero-tolerance binary-mantissa conversion, and a positivity reserve (a Chebyshev-center maximum-margin functional of the basis) is blended in at a small δ to purchase audit margin, because the LP optima are generically *tangent* to zero (Sec. 6).

The certificate is an exact continuum audit. For each spin $J \leq j_{\text{audit}}$ the smeared spectral density $E(w)$, $w = 1/m$, is proven non-negative on the full mass interval $w \in (0, 1]$—not on a grid. For integer-power smears E reduces to a rational function of w plus arctan/log transcendentals; these are sandwiched by alternating-series rational bounds and decided by Sturm sequences or by Bernstein-coefficient certificates [22] (the latter pure-rational and faster). The higher-spin tail $J > j_{\text{audit}}$ is treated separately: for the scalar case a complete (all-spin) tail proof is available and used (Sec. 4); for the gravitational case the $m \to \infty$ impact-parameter (Bessel) closure is imposed only as shaping constraints on a finite grid—it is never audited and does *not* constitute a proven all-spin lemma—so the gravitational certificates are exact for $J \leq j_{\text{audit}}$ and the tail is uncontrolled beyond it (App. C, Sec. 8). A `valid` verdict without an auditable exact certificate is a hard error; interval arithmetic was profiled and found unusable here (the dependency problem: components varying by $\sim 10^{24}$ around a $\sim 10^{-24}$ sum), which is why the exact continuum proof is used. The certification primitives (Sturm, Bernstein positivity, and the tradition of computer-verified exact-arithmetic proof [22, 23]) are standard; the contribution is their assembly into a selection oracle for functional design.

3.3 The assertion battery and the designed asymmetry

Every physics kernel is derived symbolically [42] and checked against closed forms at import—before any solver runs. The battery (21 checks, all passing) verifies the graviton pole split and its cancellation at higher k, the dispersive invisibility of the contact terms, the exact C_2/C_4 EFT sides term by term, the improved-kernel resummation $C_2^{\text{imp}}[J{=}0] = 2/m^4 + 3p^2/m^6$ and its regularity at $u = -m^2$, the general-D partial-wave reduction to the validated $D = 4$ case, the no-gravity anchor $\kappa(4)$ equal to the scalar α^*, and a single-heavy-scalar end-to-end model (App. B). If any check fails the verifier refuses to build; this has caught real derivation errors.

The pipeline is architected so that the fallible layers (the float LP, the rationalization) cannot emit a bound without the exact audit passing: dropping spins *weakens* the positivity requirement, so a naive spin-truncation could in principle produce a spurious bound (an explicit counterexample motivates the a-posteriori exact audit); rationalization rounds in the weakening direction; the reserve purchases margin rather than assuming it; and tangency is treated as physics rather than forced to a spurious positive. A spurious *exact* certificate is therefore excluded by construction, not as an empirical finding. Consistent with this, across the program every observed failure was a refusal or an infeasibility, and the caught bugs (e.g. a wrong-by-m^2 kernel, a tolerance-rounding corruption) produced refusals or weaker-but-true bounds, never a wrong theorem.

Table 1: Certified dual-functional values (units $M = 1$); each is an exact-rational certificate for the audited spins in its scope column, and upper-bounds the physical coefficient only conditional on positivity beyond those spins (Sec. 8(i)). "+Regge" denotes axiom (A1); "+mero." denotes axiom (A2). The reference column quotes CMRS [4]; the sharpest published standard-domain dual bound is 2.969 [18].

Quantity	this work	ref.	scope
D=7 ray (slope +0.130)	2.9388	2.7585	$J \leq 40$
D=7 hard edge (max-g_3)	23.1723	18.0717	$J \leq 40$
D=10 g_0, $p_{\max}$=1	3.0136	3.000	$J \leq 40$, +Regge
D=10 g_0, extended-u	2.96514	—	$J \leq 40$, +Regge+mero.
D=10 g_0, $p_{\max}$=1 tail-hardened	3.0708	—	$J \leq$ **120**, +Regge
D=10 g_0, ext.-u tail-hardened	**3.0401**	—	$J \leq$ **120**, +Regge+mero.

4 Validation: the scalar EFT-hedron

Before any gravitational claim, the identical certifier reproduces the known scalar EFT-hedron in $D = 4$ with exact *all-spin* certificates: the analytic single-null optimum $\tilde{g}_3 \geq -10.612487219$ (matching the closed form to 10^{-8}, certificate margin on the sound side); a convergence ladder to -10.346862 that reproduces the reference's Table-2 value exactly at low order; and the box $0 \leq \tilde{g}_4 \leq 1/2$ with both endpoints exact (22/22 checks in the scalar validation suite, distinct from the 21-check physics-kernel battery of Sec. 3). Here the tail *is* proven for all spins (a $\deg_X \leq 2$ tail argument with an explicit discriminant/Cauchy-strip bound), which is why the scalar reproduction is genuinely all-spin. The battery includes an audit-guard regression: the audit *refuses* a known-spurious spin-truncated functional and *accepts* the true one. This demonstrates the certifier reproduces a known answer and rejects a known-wrong one, which is what licenses using it on the gravitational sector.

5 Results

Table 1 lists the certified bounds; each is an exact-rational certificate for its stated audited spin range (the exact rationals are in App. A). We reiterate the tail status: for the scalar case the certificate is genuinely all-spin; the gravitational cases are exact for their audited spins but none is an all-spin bound—the $J \leq 40$ functionals are measurably negative on higher-spin massive states, and even the $J \leq 120$ tail-hardened certificate rests on scans beyond its audited scope (Sec. 8(i), App. C).

The two $D = 7$ entries require their own setup, stated here once: they are rays in the (g_2, g_3) plane of the *non-supersymmetric* graviton-plus-scalar EFT in $D = 7$, the reference's own worked example. Each certifies an inequality $g_2 + \alpha\, g_3 M^2 + c\, 8\pi G/M^2 \geq 0$ along a fixed slope α (smaller c is stronger): the "ray" row is the reference's interior ray (our slope +0.1298 against its +0.1299), and the "hard edge" is the max-g_3 extremal ray at slope $\approx -1/3$. These two campaigns validate the gravitational machinery on the reference's $D = 7$ example before any $D = 10$ claim; both constants land above (weaker than) the reference's hand-tuned values, the expected direction of the exact-certification cost (Sec. 6).

Table 2: The three campaigns. "cert." counts certified evaluations. Best c is the widest-scope certified bound.

campaign	evals	cert.	best c	spend
D=7 hard edge	223	81	23.1723	\$20.3
D=10 $p_{\max}$=1	176	48	3.0136	\$11.2
D=10 extended-u	462	175	2.96514	\$44.6

5.1 Campaigns

Three search campaigns were run, all with the same configuration: generator `anthropic/claude-sonnet-5` (proposes the structural mutations; sampling temperatures $\{0.8, 1.0\}$), adversarial judge `google/gemini-3.5-flas` (temperature 0.6; a distinct model family, gating proposals but failing open), two evolution islands, archive size 60; total generative spend $\approx$ \$76. Table 2 summarizes them. Each certified value was re-audited outside the search harness with the canonical audit and reproduces its exact rational bit-for-bit.

The $D = 7$ hard-edge campaign is a genuine descent, $36.68 \to 23.17$ over ~ 15 improving certified steps, closing $\sim$71% of the gap from the prior certified bound 35.41 toward 18.0717—but it does *not* cross 18.0717; the reference value rests on the tolerance-decided standard (Sec. 6) while ours is exact for $J \leq 40$. The $D = 10$ $p_{\max}$=1 campaign converged: the floor 3.0136 was re-derived by six independent certified evaluations; it is 0.45% *above* the reference 3.000. The two numbers carry different computational standards: the reference's 3.000 is a floating-point SDPB optimum at a finite spin cutoff ($J_{\max} = 42$, with stability under the cutoff checked numerically), not an exact certificate, while 3.0136 is exact-rational for $J \leq 40$; a gap of this size and sign is the expected cost of the exact standard (the tangency tax, Sec. 6), not a discrepancy.

The extended-u campaign produced the sharpest truncation result. Its certified trajectory is 3.0495 (at $J \leq 36$) $\to$ 3.0400 ($J \leq 38$) $\to$ 3.0023 $\to$ 2.9928 $\to$ 2.96514 (at $J \leq 40$), then convergence. The winning functional added two column families—the C_2^{imp}-susy and E_1-null towers—that every prior champion had left empty, the structural move prescribed by the diagnostic of Sec. 6. It is an iteration-zero canonical certificate (no refinement, hence reproduction-stable), numerically below the reference 3.000—under axiom (A2) and at $J \leq 40$ truncation (Sec. 8(i)).

5.2 The tail-hardened certificate

To measure what the truncation is worth, we re-solved the champion basis with tail-hardened shaping: direct per-state positivity rows at $J = 42$–176 evaluated with cancellation-free large-J kernels (grids scaled with J^2 to track the eikonal dip; see Sec. 6), the impact-parameter grid extended to $b = 160$ with the asymptotic large-b sign row, and the exact audit run *inside* the refinement loop for all $J \leq 60$. The result is an iteration-zero canonical certificate

$$g_0 \;\leq\; 3.114449\, 8\pi G/M^6,$$

with exact-rational per-spin certificates at *every* even spin $J \leq 120$ (the certificate audits $J \leq 60$; a standalone process re-audited those and extended the exact audit through $J = 120$, every spin passing in seconds—proving genuine margin is fast, only refusals grind). A stable grid scan finds the functional positive at every spin out to $J = 240$, with the minimum margin occurring on the eikonal dip track $b \approx 14$ and *growing* monotonically with J (from $+1.7$ at $J = 44$ to $+16$ at $J = 240$); the scan is a measurement, not a proof. This already quantifies the truncation slack: the sub-3.000 value 2.96514 does not survive tail hardening.

Measuring this functional's impact-parameter transform on the continuum then exposed a structural weakness and, in the process, an implementation bug worth disclosing. The Bessel shaping rows (and the first continuum measurements) had silently used a transform truncated at $p = 1$ rather than the full smearing endpoint $p_{\max}$—a shaping-layer inconsistency that cannot affect any certificate (those rows are never audited), but which distorted early $\hat{f}$ diagnostics; it was found, fixed, and all $\hat{f}$ values quoted in this paper are under the corrected full-endpoint transform. Under that transform the first-pass functional's $\hat{f}$ is non-negative at all sampled points (min $+2.8 \times 10^{-9}$), but its leading large-b coefficient sits exactly on its sign boundary ($a_2 = 0$): the functional has *no positive asymptotic floor*, so positivity beyond any sampled range is uncontrolled—the tangency tax in impact-parameter space (App. C). A second re-solve therefore demands *strict margins*—$\hat{f}(b) \geq 10^{-3}\,\|\mathrm{row}\|_1$ on a densified b-grid extended to $b = 320$, and $a_2 \geq 10^{-4}$, restoring a manifest large-b floor. The result, again at iteration zero, is the sector's current best certificate,

$$g_0 \;\leq\; 3.040145\, 8\pi G/M^6,$$

with the same exact per-spin scope (every even $J \leq 120$, all audits passing, re-verified standalone) and grid scans positive to $J = 240$ (margins $+1.2$ to $+8.5$). Its impact-parameter transform, measured with the corrected full-endpoint transform on a dense continuum grid, is positive at every sampled point except a single residual dip of -1.1×10^{-8} at $b \approx 368$, beyond the margined grid—the recurring endpoint oscillation; a further re-solve with endpoint control ($|f(p_{\max})|$ bounded, a_2 floor raised) produces a functional whose transform is positive at *all* sampled points with its $16a_2/b^3$ asymptote approached monotonically, at the cost of a higher certified value, 3.1023 (App. D).

We are careful about what these numbers do and do not imply. The hardened values are single, unoptimized certificates, not optima (which is why the margin re-solve can land *below* the first pass' 3.1144). Monotonicity in the spin constraints bounds the all-spin optimum of this basis below by the *true* $J \leq 40$ optimum—a quantity we do not know: 2.96514 upper-bounds it (it is achieved), but does not lower-bound it, because certifiability requires audit margin, and the raw LP optima below 2.96514 are not continuum-feasible (Sec. 6). No unconditional numeric lower bound on the all-spin optimum therefore follows from our data. What can be said: the $J \leq 120$-truncated optimum is ≤ 3.0401 (achieved), and if the margin-hardened functional is in fact positive on all spins—which its measured tail supports but does not prove—the all-spin optimum is ≤ 3.0401 as well.

The same apparatus applied at $p_{\max} = 1$—the minimal-axiom, non-perturbative setting, using the basis that certifies 3.0136 at $J \leq 40$ (with the coefficient cap lowered to the audit-validated regime)—yields $g_0 \leq 3.070806\, 8\pi G/M^6$ at the same exact scope (every even $J \leq 120$; scans positive to $J = 240$). Two conclusions follow. First, the minimal-axiom truncation tax is also real but modest: $3.0136 \to 3.0708$. Second, the extended-u domain remains genuinely tighter *after* hardening ($3.0401 < 3.0708$), so axiom (A2) still buys about 0.03—but both hardened values now sit *above* the reference's 3.000, completing the truncation-slack diagnosis.

6 Findings

A tolerance-free treatment surfaces structure a tolerance-based one does not. Several of these findings corrected our own prior understanding by direct measurement.

The reference's fine constraints are tolerance-decided. In our reconstruction of the standard basis, the finely sampled $J = 42$ constraint set is marginally infeasible: re-evaluating reference-style optimal functionals on it, we find violations at the $\sim 1.8 \times 10^{-10}$ level, within solver tolerance (a measurement in our pipeline's reconstruction, not a claim about the reference's own code or data). We note that $J = 42$ is also the reference's own spin cutoff, so its quoted optimum is computed exactly at this knife-edge. This threatens *no* physical conclusion—the bound is $O(1)$ and the violation is 10^{-10}, and the reference numbers are not thereby wrong. What it illustrates is narrower and, we think, worth stating: a tolerance-decided inequality is not a theorem, and turning the audited-spin constraints into exact theorems (as our certification does for $J \leq 40$) is a legitimate mathematical value even when no physical conclusion was ever in doubt. Our certificates scope to $J \leq 40$ exactly rather than absorb the 10^{-10} violation in tolerance—though, as Sec. 8(i) stresses, our own high-spin tail beyond $J \leq 40$ is uncontrolled and by a far larger margin, so this is a statement about the audited constraints, not a claim that our functionals are closer to all-spin than the reference's.

Tangency is the binding constraint. The LP-optimal functionals touch zero on continuum families of states (near-threshold high-spin rows are large-b Bessel rows in disguise). A functional that saturates positivity carries no audit margin, so it cannot be certified as a strict bound. This "tangency tax"—the gap between the raw LP optimum and the certifiable value—is the actual barrier to tighter dual bounds, not the audit resolution and not the search.

A corrected misdiagnosis (a basis limit, not an audit limit). Earlier notes attributed the failure to certify the tight extended-u functional at $J = 38, 40$ to audit resolution. Direct measurement of $\min_w E(w)$ on $(0, 1]$ refuted this: for the $p_{\max}$=9/8 functional at value 3.0495 we find $\min E = +0.084, -0.028, -0.133$ at $J = 36, 38, 40$. The functional *genuinely violates* positivity at $J \geq 38$; the audit was refusing correctly. No audit improvement can certify a negative functional; the path was a richer basis, which the subsequent campaign supplied. (A scale-immune audit variant was built, proven sound, then reverted once the measurement showed it targeted the wrong bottleneck, keeping the certifier provably identical to the validated original.)

The bound 2.96514 is a measured floor. We probed how to push lower and measured each lever closed. A sweep of the domain endpoint (Table 3) shows $p_{\max}$=9/8 is the unique certifiable sweet spot: lower $p_{\max}$ reaches a lower *raw* optimum (2.826 at $p_{\max}$=1) but those functionals are tangent and uncertifiable, and higher $p_{\max}$ is infeasible. The same richer basis at $p_{\max}$=1 does not certify below the standard floor—so axiom (A2) is load-bearing—and more integer-power nulls are numerically intractable in the current shaping with no signal of payoff. The conclusion is measured: 2.96514 is the lowest value our sweeps could *certify* in the current integer-power basis *at the $J \leq 40$ truncation*; no certifiable functional below it was found across the domain sweep and campaign convergence. We stress the logical status: this is an empirical floor of the certifiable search, an *upper* bound on the true $J \leq 40$ optimum of the basis (which sits below it by the tangency tax, at an unknown value); it therefore does not supply a numeric lower bound on the all-spin optimum (Sec. 5.2).

The truncation tax, measured; and a corner-window failure mode. The tail-hardened re-solves (Sec. 5.2) price the spin truncation directly: enforcing positivity through $J \leq 120$ (with clean scans to 240) moves the certified extended-u value from 2.96514 to 3.0401 (first pass 3.1144), and the minimal-axiom $p_{\max}$=1 value from 3.0136 to 3.0708, so the sub-3.000 number was truncation slack, as the monotonicity argument of Sec. 8(i) predicted. Getting there surfaced two structures worth recording. First, the binding high-spin constraint travels: the minimum of E_J sits on the eikonal track $b = 2J/m \approx 14$, i.e. at $m^2 \sim (J/7)^2$, so any fixed mass grid is eventually blind to it (an intermediate solve was exactly positive through $J \leq 60$ yet failed at $J = 68$–70 just past its

Table 3: Domain sweep for the champion basis at $J \leq 40$. Raw optimum is the LP value; "certifies" is whether the exact audit accepts it. At $p_{\max}$=9/8 the sweep's own re-solve certifies 2.9684; the campaign's converged optimum in the same basis is 2.96514 (Table 2).

$p_{\max}$	raw LP optimum	certifies?
1	2.826	no (tangent)
33/32	2.956	no (tangent)
17/16	2.894	no
9/8	2.968	**yes** $\to$ 2.9684
5/4, 3/2	—	infeasible

grid edge). Second, and specific to the extended-u domain: for $p_{\max} > 1$ the $p = m$ kernel corner lies *inside* the smearing range whenever $m^2 < p_{\max}^2$, and there the functional can develop violent, narrow spikes—an intermediate functional passed every grid scan (margin +0.6) while its exact density at $J = 48$, $m^2 \approx 1.203$ plunged to -1.4911×10^7, verified in exact rational arithmetic. Corner-masked grid rows and scans cannot see this structure; only the continuum audit can, and it refused correctly. Both failure modes are cured by J^2-scaled grids, dense corner-window rows, and running the exact audit inside the refinement loop.

Four routes, one wall. After the tail-hardened certificates we attacked the sub-3.000 region along four independent routes, each under the same exact standard. (i) *Margin architecture and domain sweeps*: under tail-hardened shaping the raw LP optima reach 2.72–2.90 across $p_{\max} \in \{1, \ldots, 9/8\}$, but every certification attempt in this route's sweep—the stall points land between 2.93 and 3.13 depending on $p_{\max}$ and margin architecture (App. D)—was refused, with exact post-mortems showing genuine $O(1)$ violations in narrow sub-grid windows; the margin-priced LP optimum on the reference's own domain stalls at 3.0007–3.002 before the audit refuses it—numerically the reference's own 3.000, recovered from the exact side as the tangency point, though *not* a certificate. (ii) *The partial-wave unitarity ceiling* ($\rho_J \leq 2$, the non-projective upgrade): not licensed for the susy-reduced density (the reference defines the ceiling for component amplitudes and never uses it), and, even imposed as if licensed, a penalty formulation uses *zero* negativity at the optimum—the $D = 10$ measure ($n_J^{(10)} \sim J^7$) makes any negative spectral weight dearer than positivity. (iii) *Basis enlargement*: two new validated column families (an E_2 null tower from the second derivative of the master crossing identity, and a numerically stable X_8 crossing-null tower; both defined and validated in App. D) move the raw optimum from 2.90 to 2.63 and produce a near-miss at 2.91 blocked by a single near-threshold audit—yet every certified value in the enlarged basis lands at or above the old record. (iv) *Closure-structured re-solves* (endpoint control for the impact-parameter transform) certify cleanly but above the record. Four independent attacks terminate at the same place: the certified frontier of this functional class bottoms out at ≈ 3.0–3.04, with the binding structure always the near-threshold/corner-window tangencies. We regard this convergence—not any single number—as the measured content: under exact certification, ≈ 3.0 is where this functional class stops, from every direction we know how to approach it.

7 Ablation: does the language model help?

The pipeline uses a language model to propose structural mutations. A fair question is whether that proposer outperforms a conventional search on this task. We ran a controlled ablation: 12 independent conventional runs—six random-restart hill-climbs and six genetic-algorithm runs (population 10, tournament selection, tower-wise crossover)—against the *identical* evaluator, seed design, design-space ranges, per-evaluation budget (462, matching the language-model campaign), and per-evaluation wall-clock limit. The only substitution is the mutation operator: structural edits drawn from the documented design-space ranges in place of the model's proposals. (Two design subtleties matter and were controlled: the mutation operator must be able to populate the C_2^{imp}-susy and E_1 towers from empty—an early operator could not, and was corrected—and the comparison must be made at fixed audited scope, because the fitness's spin-scope discount is weak enough that free search drifts to the easier $J \leq 36$ truncation.)

At the narrower scope the conventional search matches the model. With j_{audit} free, both drift to $J \leq 36$; the conventional best there is 2.948, at or below the model campaign's own narrow-scope values (its $J \leq 38$ point 2.9556, which we do not promote for the same reason). No advantage for the model.

At the headline scope the best conventional run reproduces the model's value. Pinned to $j_{\mathrm{audit}} = 40$—the scope of the reported 2.96514, where the extended-u basis *refuses* until additional columns are built—the best of the twelve conventional runs certifies $g_0 \leq 2.96597\, 8\pi G/M^6$, a gap of 8×10^{-4} to the model's 2.96514: an effective tie. That certificate uses a (partial) C_2^{imp}-susy tower, $k \in \{3, 4\}$, the same column family the model's champion added. The reproduction is not uniform: only two of the twelve seeds broke 3.00 and only one reached the model's value, the rest plateauing near 3.01–3.03—so the model found the frontier in a single campaign whereas conventional search matched it as the best of twelve. A direct probe explains the spread: the C_2^{imp}-susy and E_1 towers are *jointly* beneficial but *individually* neutral-or-broken (adding the full susy tower alone renders the LP infeasible, as does truncating it or removing it from the champion), a narrow fitness valley a greedy or genetic search crosses only occasionally. The most one could read from this is a mild sample-efficiency difference—not a structural capability the conventional searches lack, and not a difference in the certified value reached.

We therefore make *no* claim that the language model outperforms a conventional evolutionary or greedy search on this problem: on the evidence here it does not. The model functions as the search's mutation operator, and the contribution of this work is the exact-certified pipeline around it, not the proposer. This is a single-task, single-basis result and we do not generalize it to a claim about language-model search overall; but for the present task it is what licenses the neutral framing of the title. Consistent with the wall-convergence findings (Sec. 6), the certified frontier appears set by structural tangencies that neither proposer moves.

8 Limitations

We state the boundary precisely. *(i) Tail / spin scope—the binding limitation.* A dual functional proves a universal (all-spin) bound only if its smeared spectral density $E_J(w)$ is non-negative on *every* massive state—every spin J and every mass $m^2 \geq M^2$. Because unitarity leaves each state's spectral weight free and non-negative, a single state with $E_J < 0$ admits a spectrum concentrating weight there that violates the bound; this is precisely the spin-truncation counterexample our audit-guard encodes (Sec. 4). Our exact audit establishes $E_J(w) \geq 0$ on the mass continuum only for the audited spins $J \leq 40$. For $J > 40$ the functional is *not* audited: the impact-parameter positivity $\hat{f}(b) \geq 0$—which is the large-(J, m) asymptotic *form* of the same per-state condition, not an aggregate substitute for it—is imposed only as constraints on a finite grid of b in the (unsound) shaping stage, and is neither proven in the continuum nor sufficient at finite J (it fixes only the leading eikonal term; $O(1/J)$ corrections can still turn E_J negative). Direct measurement of the certified champion functional confirms the gap is real and not benign: it is barely positive at the audit edge ($+0.16$ at $J = 40$), turns negative at $J = 42$, and then *diverges* monotonically to large negative values (App. C). Validity therefore does *not* rest on any *aggregate* tail non-negativity—that characterization would be false, and it is not the reference's mechanism either. The gravitational certificates are exact statements about their audited spin range, *not* all-spin bounds; we do not claim $g_0 \leq 2.96514$ for the full spectrum—nor does 2.96514 lower-bound the all-spin optimum, since it upper-bounds (does not equal) the true truncated optimum (Sec. 5.2).

The tail-hardened certificates (Sec. 5.2) measure the cost of enforcing positivity far past the old scope, reaching exact scope $J \leq 120$ at 3.0401 (extended-u, with strict impact-parameter margins) and 3.0708 ($p_{\max}$=1)—single unoptimized certificates, not measured floors. Even that is not all-spin: its tail beyond $J = 120$ rests on grid scans (with growing margins) rather than proof, and the corner-window structure of Sec. 6 means scans alone are not trustworthy there. The one genuinely all-spin certificate here is the scalar reproduction, which carries a complete finite-degree tail proof (Sec. 4). *(ii) Axioms.* Every $D = 10$ bound assumes (A1); the extended-u bounds additionally assume (A2), the standard weak-coupling (meromorphy) axiom (Sec. 2). (A2) is load-bearing and stronger than the core consistency axioms, so the extended-u number is a *tree-level* bound, not a non-perturbative swampland exclusion. The precise statement is that 2.96514 is below the reference value in the extended-u domain under a stronger computational standard and this additional axiom—not a strict improvement on the reference's own domain, where our exact optimum 3.0136 is 0.45% above. *(iii) Reproduction of a specific value.* An opt-in fast-refuse audit path used for throughput (sound; refuse-only) can reorder the refinement feedback, so refining certificates are trajectory-dependent; the campaign's lowest value (2.9556 at the narrower $J \leq 38$) is such a point and is not promoted. Only the iteration-zero, trajectory-independent $J \leq 40$ value 2.96514 is reported. *(iv) Numerics.* The dense LPs are occasionally non-deterministic under memory pressure and can land on different degenerate vertices; only re-audited bit-for-bit values are promoted. *(v) Novelty and normalization.* Our g_0 is the standard dual-side coefficient, in the same normalization as the reference and the crossing-symmetric dispersion literature [4, 21], so the comparison to the reference's 3.000 is same-object. One reconciliation remains open and we make neither "first" nor "record" claim: the fixed normalization constant relating our g_0 to the primal coefficient α [6], which would place our dual upper bound and the primal lower bound on one axis (a genuine two-sided pincer rather than a companion). The second check we had left open—the supergravity sections of the enlarged-domain crossing-symmetric program [21]—is now done (Sec. 9): its supergravity bound, $g_0 M^6/8\pi G \leq 2.97$, is computed on the standard domain and presented as matching [18], and no extended-domain g_0 appears there, so the extended-u domain remains uncomputed in the literature apart from this work. The sharpest published $D = 10$ maximal-supersymmetric *dual* upper bound on the standard domain is 2.969 [18], a minor improvement on the reference 3.000 (the sharper unitarity-upper and low-spin-dominance results [16, 17] are in four-dimensional general gravity, a different sector).

9 Related work

This work sits in the swampland/positivity program on the *dual* side, and its natural companion is the *primal* S-matrix bootstrap that maps the same $D = 10$ maximal-supergravity region from the inside [6], with its extensions to nine, ten, and eleven dimensions and to M-theory [7]. This exact coefficient is the object of an active 2024–2025 program on the uniqueness of the type II string as a UV completion of maximal supergravity: Bossard–Loty refine the next-to-leading coefficient toward that uniqueness question [8], and Albert–Knop–Rastelli's tree-level bootstrap places the string strictly inside a small allowed region and finds an extremal single-Regge-trajectory solution [18] (with a tree-level heterotic continuation [19])—the same tree-level (weak-coupling) class as our extended-u result. The dual/dispersive lineage includes the original causality argument [1, 2], the two-sided scalar bounds we reproduce [3], the EFT-hedron geometry [5], the gravitational bounds that are our direct reference [4], the crossing-symmetric dispersion relations [20] and their enlarged-domain revisit of the same amplitude [21], the Regge-boundedness analyses [9], and the causality and graviton-partial-wave constraints in general dimension [12, 13] that probe exactly the analyticity assumptions our axiom (A2) touches. Dispersive bounds on scalar EFTs coupled to gravity across dimensions [14, 15] are the closest analogues to our $D = 7$ ray entries. Adjacent frontier work isolates string values in related maximal-supersymmetry settings [37, 39, 38]. Crucially, the field has moved *beyond* plain positivity: incorporating the upper bound from unitarity of the partial-wave imaginary parts yields sharp individual-coefficient bounds [16], and low-spin dominance pins the gravitational Wilson coefficients into small islands much more constrained than dispersive bounds indicate [17]. We report a positivity-only bound deliberately—the object of this work is the *exact certification* of the positivity method, and no certificate in this paper incorporates the unitarity upper bound (the one experiment with it, a diagnostic penalty scout, is reported as wall route (ii) in Sec. 6 and App. D); doing so within the exact-certified pipeline is a natural extension (Sec. 10).

Two comparisons frame where our numbers sit. (a) Priority: the sharpest published $D = 10$ maximal-supersymmetric *dual* upper bound on the standard domain is $g_0 M^6/8\pi G \leq 2.969$, obtained by Albert–Knop–Rastelli alongside their primal map (their Eq. (50), a minor improvement on the reference 3.000) [18]; the crossing-symmetric revisit [21] recomputes the same bound as 2.97 on the smearing region $p \leq M$—shown there to be equivalent to the standard $u \in (-M^2, 0]$ domain—and presents it explicitly as matching [18], with its new results confined to $D = 4$ graviton MHV Wilson coefficients and a graviton–massive-spin-4 coupling. (b) Domain: the extended-u domain is the reference's own Appendix-B construction, applied there only to g_2, g_3 in $D = 6$ and left uncomputed for maximal supergravity; neither [18] nor [21] computes any g_0 bound over an extended domain $|u| > M^2$, so our extended-u numbers have no published counterpart to compare against. The fixed constant relating our g_0 to the primal α [6] remains genuinely open (we found no published closed-form conversion); deriving it would place the primal lower bound and the dual upper bounds on one axis, and is a concrete, self-contained next contribution.

On the methods side, machine learning has been applied to the S-matrix and conformal bootstrap: reinforcement-learning navigation of bootstrap constraints [26], transformer models for amplitude coefficients [27], machine-learned string-model bootstrap [28], and—closest to our setting—optimizers (including neural ones) driving the S-matrix bootstrap directly [29, 30, 31, 32]; we do not compare against these as baselines here. Closest in spirit on the rigor side is the a-posteriori verification of numerical semidefinite-program output in sphere-packing bounds [36], atop arbitrary-precision solvers [33, 40]; our pipeline differs in making the exact certificate the *selection pressure* of the search rather than a post hoc check. The exact certification itself builds on standard real-algebraic tools (Sturm sequences, Bernstein/SOS positivity certificates) and the tradition of computer-verified exact-arithmetic proof [22, 23]; framing the contribution as "certified functional design" rather than a new proof technique is the honest positioning.

10 Discussion and outlook

The results support a methods-and-rigor reading. Automated search under exact certification can design dual functionals that reproduce a known EFT-hedron and reach the hand-designed frontier, can explore a domain the reference left uncomputed, and turns tolerance-decided inequalities into theorems for the audited spins. The physics number is a demonstration, honestly modest and axiom-conditional; it does not improve the reference on the reference's own domain.

Pushing g_0 below 2.96514 was measured to require a substantial new construction, not tuning. Higher crossing nulls were subsequently built, validated, and run (X_8 and the E_2 tower, wall route (iii), App. D): they move the raw LP optimum from 2.90 to 2.63 but not the certified frontier, so basis reach is not the binding constraint—the tangency tax is. What remains untried is half-integer/algebraic smearing bases, requiring the exact integrator and audit extended to new transcendentals, with payoff capped by the same tax. The most important technical debt, however, is the all-spin tail (Sec. 8(i), App. C). Our gravitational certificates are exact only for the audited spins $J \leq 40$; direct measurement shows the certified functionals are negative on higher-spin massive states, so they are not all-spin bounds. Promoting them requires two ingredients this paper does not supply: (E1) a continuum proof that $\hat{f}(b) \geq 0$ for the certified profile (not a grid), and (E2) an explicit bound on the finite-(J, m) corrections to the eikonal limit showing they cannot drive E_J negative for any $J > j_{\text{audit}}$. We caution, from our own experience in this program, against reading the all-spin upgrade as a small number of outstanding steps: each time one ingredient was completed, measurement revealed another. A closure-structured re-solve achieved a textbook (E1) profile—$\hat{f}(b)$ positive at every sampled continuum point with its $16a_2/b^3$ asymptote approached monotonically—and its tail *still* failed, from a direction the apparatus had no constraints for: a migrating negative channel at intermediate impact parameter beyond the shaped spin range, confirmed in exact arithmetic ($E = -4.13$ at $(J, m^2) = (204,\ 21.2)$, deepening to -15.6 at $(240,\ 22.9)$ as the channel drifts; at *fixed* m^2 the sign flips back positive, so this is not a fixed-mass mechanism but precisely the finite-(J, m) eikonal-correction regime, where $\hat{f}$ is still small and the corrections dominate). The honest current scoping is therefore a research program with at least three coupled sub-problems: (E1$'$) rigorous continuum control of $\hat{f}$ (enclosures, modulus of continuity, a_2-dominance—machinery in place); (E2) explicit finite-(J, m) eikonal-correction bounds [9]—now demonstrated, with exact data points, to be the binding ingredient rather than a technical afterthought; and corner-window control beyond the audited spins. Any eventual all-spin optimum for this basis is bounded below by the (unknown) true truncated optimum and above by 3.0401 if the margin-hardened tail closes; the payoff of the closure program is rigor, not a lower headline. Until then the gravitational sector yields exact *truncated* certificates (now through $J \leq 120$) and the all-spin dual bound in this framework is open. A complementary, cheaper direction is to read out the extremal spectrum [34, 35]: the certified functional is tangent at specific (m, J), the masses and spins where a bound-saturating theory would place states.

We close on what the number does and does not target. Driving the plain-positivity ceiling toward the string value is not, by itself, the physically decisive question: the unitarity upper bound and low-spin dominance [16, 17], and the tree-level primal map [18], already place the leading coefficient near the string value in small islands. The present work tightens the *rigor* of the positivity ceiling and maps the current method's dual frontier; the extension that would target the physically-relevant window directly—incorporating the unitarity upper bound on the partial-wave imaginary parts into the exact-certified pipeline, so that the certificate constrains the coefficient from both sides—is the natural and, we think, most valuable next step.

A Exact rational certificates

The $J \leq 40$-truncation extended-u bound $g_0 \leq 2.96514\, 8\pi G/M^6$ has exact value N/D with

$$N = 4571717227703018539853043210372667360202506092897741895 9,$$

$$D = 154182106265449737162862948428769350121875381855570825216.$$

The margin-hardened extended-u bound $g_0 \leq 3.0401\, 8\pi G/M^6$ (Sec. 5.2; exact for every even spin $J \leq 120$) has exact value N/D with

$$N = 6086429680222317487582684788994075304730520431387 9717,$$

$$D = 2002019554617274314476013143913749595566150178099 2000,$$

and the minimal-axiom $p_{\max}$=1 tail-hardened bound $g_0 \leq 3.0708\, 8\pi G/M^6$ (also exact for every even $J \leq 120$) has exact value N/D with

$$N = 636312169578577262763924918987810831311629033308 21699,$$

$$D = 207213381443121567334714526744173441138728933590 37302.$$

The $p_{\max}$=1 bound 3.0136, the first-pass hardened bound 3.1144, the $D = 7$ ray and hard-edge slopes and constants, and the full functional coefficients (39 exact rationals per hardened functional) are provided with the accompanying code, together with standalone re-audit scripts that reproduce each value bit-for-bit.

B Kernel identities (assertion battery)

The load-bearing kernel identities—the $(-1, -1, +1)$ graviton pole split and its cancellation at $k \geq 4$; the dispersive invisibility of the contact terms; the term-by-term C_2 and C_4 EFT sides; the improved-kernel resummation $C_2^{\mathrm{imp}}[J{=}0] = 2/m^4 + 3p^2/m^6$ and its regularity at $u = -m^2$; the general-D partial-wave reduction to $D = 4$; and $\kappa(4) = \alpha^*$—are each encoded as an executable exact-symbolic check (21 in total, all passing) that the verifier runs at import and refuses to build on failure. The check source is provided with the code so the identities can be audited directly rather than trusted.

C Tail status

For the audited spins $J \leq j_{\mathrm{audit}} = 40$ the smeared spectral density $E_J(w)$—the certified functional evaluated on a *single* massive state of spin J and mass $m = 1/w$, $w \in (0, 1]$—is proven non-negative on the full mass continuum. This per-state non-negativity, spin by spin, *is* the certificate: a universal bound requires it for every spin, because unitarity lets a spectrum place all of its weight on any one state (Sec. 8(i)).

Past the audited scope the certified functionals are *not* non-negative. We measure $\min_{w\in(0,1]} E_J(w)$ directly for the $J \leq 40$ extended-u champion functional ($g_0 \leq 2.96514$; its exact-rational coefficients are provided with the code). At the audit edge it is barely positive—+0.16 at $J = 40$—and turns negative immediately beyond it, *diverging* monotonically: -0.38, -1.22, -2.10, -3.02, -3.99 at $J = 42, 44, 46, 48, 50$, with no recovery (the numerical Gegenbauer weight degrades past $J \approx 50$, but the trend is already unambiguous). The onset tracks the audit scope exactly: a functional in the same basis carrying more audit margin turns negative one step later ($J = 44$), and a lower-basis sibling certified only to $J \leq 36$ one step earlier ($J = 38$). A single unitary state placed at any of these (J, m) points makes the functional's spectral side negative, so beyond the audited spins the inequality is not a theorem for that spectrum.

We emphasize that validity does *not* rest on any *aggregate* tail non-negativity: such a characterization would be false, because the negativity does not cancel in aggregate—it *accumulates* (the per-spin minimum diverges rather than approaching zero from below), and a single admissible state suffices to violate the bound. The impact-parameter positivity $\hat{f}(b) \geq 0$ is the large-(J, m) asymptotic *form* of the per-state condition $E_J \geq 0$, pointwise in b—not a weaker aggregate—and it is imposed here only on a finite grid of b in the unsound shaping stage, never audited. It is necessary but not sufficient at finite J: it fixes only the leading eikonal term, and the $O(1/J)$ corrections seen above dominate. An all-spin upgrade therefore requires two ingredients this paper does not supply: (E1) a continuum proof that $\hat{f}(b) \geq 0$ for the certified f (not a grid), and (E2) an explicit bound on the finite-(J, m) eikonal corrections showing they cannot drive E_J negative for any $J > j_{\text{audit}}$. This is *harder* than proving $\hat{f}(b) \geq 0$ for the present functionals—for which the measurement shows the statement is *false*—so it is a genuinely new construction. The scalar sector, by contrast, has a complete finite-degree tail argument (a $\deg_X \leq 2$ discriminant/Cauchy-strip bound closing all spins) and is genuinely all-spin (Sec. 4).

The tail-hardened certificates (Sec. 5.2) are the current best tail status in the gravitational sector. Their measured impact-parameter profiles, all under the corrected full-endpoint transform, are: the first-pass functional (3.1144, exact at every even $J \leq 120$, scans positive to $J = 240$ with the minimum margin on the $b \approx 14$ eikonal track growing from $+1.7$ to $+16$; each audited spin passes in seconds, confirming that the historical "audit walls" were genuine violations, never audit weakness) has $\hat{f}$ non-negative at all 619 sampled points but $a_2 = 0$ exactly—no positive asymptotic floor, so its large-b positivity is uncontrolled beyond the sampled range. The margin re-solve (3.040145, strict $\hat{f} \geq 10^{-3}\|\text{row}\|_1$ on a densified grid, $a_2 \geq 10^{-4}$; exact at every even $J \leq 120$, standalone re-verified, scans positive to $J = 240$) has a positive asymptotic floor and $\hat{f}$ positive at every sampled point *except* a single residual dip of -1.1×10^{-8} at $b \approx 368$, beyond its margined grid—the recurring endpoint oscillation—so, strictly, the all-spin closure fails for it as well. An endpoint-controlled re-solve ($|f(p_{\max})| \leq 0.05$, $a_2 \geq 10^{-2}$; certified 3.1023 at the same exact scope, App. D) achieves $\hat{f}$ positive at all sampled points with its $16a_2/b^3$ asymptote approached monotonically—the profile the (E1′) rigor step needs—yet its tail scan fails from $J \approx 204$ in a migrating channel at intermediate impact parameter (exactly confirmed; Sec. 10), demonstrating that (E1′) alone is insufficient. Two cautions temper all non-audited ranges: the corner window $m^2 < p^2_{\max}$ (Sec. 6) is invisible to grid scans, and measurements are not proofs. What remains for a genuinely all-spin certificate: (E1′) rigorous enclosures of $\hat{f}$ at grid points plus an explicit modulus of continuity and an $a_2 > 0$ asymptotic dominance argument—certifiable machinery, applicable to a functional combining the margin and endpoint structures; and (E2) the finite-(J, m) eikonal-correction bound, which the migrating-channel measurement shows is the binding ingredient.

D The wall experiments

This appendix documents the four routes of Sec. 6 ("Four routes, one wall") at the level of methods and numbers, so the convergence claim is checkable rather than narrative. All experiments use the tail-hardened shaping of Sec. 5.2 and the unmodified exact audit; every driver script and captured output is in the repository (`results/positivity/drivers/`).

Common definitions. The susy-sector null towers derive from one master crossing identity: for the reduced amplitude under axiom (A1), $\langle K(s, u) - K(u, s)\rangle = 0$ with $K = m^2(2m^2 + u)\, P_J(1 + 2u/m^2)/\big((m^2 - s)(m^2 + s + u)\big)$, averaged against the heavy spectral measure. Its s-derivatives at $s = 0$ give the null family E_n: E_0 reproduces the improved C_0 kernel minus its EFT constant, and E_1, E_2 are the first- and second-derivative kernels (both matches verified exactly; `scripts/verify_e2_family.py`). X_k denotes the standard crossing-null tower at Taylor order k;

X_8 is its $k = 8$ member, evaluated in a numerically stable factored form (`scripts/verify_x8_stable.py`). Rows are smeared with integer powers p^i on $(0, p_{\max}]$ as in Sec. 2.

Route (i): margin architecture and domain sweeps. A sweep over $p_{\max} \in \{1, 33/32, 17/16, 9/8\}$ and margin architectures (reserve depth, grid densification, corner-window rows) re-solved the tail-hardened LP and attempted certification at each stall. Raw LP optima reach 2.72–2.90; every certification attempt was refused, the certifiable stalls landing at 2.93–3.13 depending on $p_{\max}$ and margin settings. Exact post-mortems of the refusals show genuine $O(1)$ violations in narrow sub-grid windows, not audit weakness. On the reference's own domain the margin-priced LP stalls at 3.0007–3.002 and is refused: numerically the reference's 3.000, recovered as the tangency point from the exact side—*not* a certificate. The grid-blindness this route must defeat is exactly documented: an intermediate $p_{\max}$=9/8 functional passed every grid scan (margin +0.6) while its exact density at $J = 48$ plunges to -1.4911×10^7 at $w = 114/125$ in the corner window (artifact `extu_tailhard_v2_j48_exact.json`). (Drivers: `floor_sweep.py`, `targeted_floor.py`, `j48_exact_artifact.py`; raw data `floor_sweep_raw.json`.)

Route (ii): the partial-wave unitarity ceiling. The non-projective upgrade $\rho_J \leq 2$ is defined for component amplitudes; the reference neither defines nor uses it for the susy-reduced density, so imposing it there is not licensed. We nevertheless measured what it could buy if it were: a penalty formulation (negative spectral weight permitted at per-unit cost—the LP relaxation of the two-sided constraint) re-solved the tail-hardened LP at both $p_{\max}$ endpoints. The optimum uses *zero* negativity (penalty $\sim 2 \times 10^{-15}$, solver noise): under the $D = 10$ partial-wave measure $n_J^{(10)} = \frac{512\pi^4}{3}(J+1)_6(2J+7) \sim J^7$, negative weight anywhere is dearer than the positivity it would relax. The ceiling is thus doubly closed for this problem: unlicensed, and worthless even if licensed. (Driver: `pen_scout.py`; output `pen_scout_out.txt`.)

Route (iii): basis enlargement. Adding the E_2 and X_8 towers (wired as ordinary columns after symbolic validation) moves the raw LP optimum from 2.90 to 2.63—the enlarged basis is genuinely stronger at the LP level—and produces the sweep's best near-miss, 2.9103, blocked by a single near-threshold audit failure. Every *certified* value obtained in the enlarged basis lands at or above the prior record (e.g. a threshold-sliced re-solve certifies 3.1737); the added LP reach is absorbed entirely by the same near-threshold tangencies. An evolutionary campaign over the enlarged design space produced no certified improvement either. (Drivers: the `round6`–`round9c` series.)

Route (iv): closure-structured re-solves. Demanding the structures an all-spin closure needs (Sec. 10): the endpoint-controlled re-solve ($|f(p_{\max})| \leq 0.05$, $a_2 \geq 10^{-2}$) certifies 3.1023 at exact scope $J \leq 120$ with $\hat{f}$ positive at all sampled points and a monotone $16a_2/b^3$ asymptote—the cleanest (E1′) profile we obtained—and its tail *still* fails, from $J \approx 204$: a negative channel at intermediate impact parameter that migrates with J, confirmed in exact arithmetic ($E = -4.13$ at $(J, m^2) = (204, 21.2)$, -15.6 at $(240, 22.9)$; at fixed m^2 the sign flips back positive as J grows, ruling out any fixed-mass mechanism and identifying the finite-(J, m) eikonal-correction regime). (Drivers: `round4d_soft.py`, `confirm_fixedm.py`; output `confirm_fixedm_out.txt`.)

Table 4: The four wall routes: what each reached and where it stopped. "Raw" is the LP optimum before rationalization and audit; only "certified" values are theorems (for their audited scope).

route	raw LP	certified	binding structure
(i) margin/domain sweeps	2.72–2.90	refused (stalls 2.93–3.13)	near-threshold tangency; corner spikes (exact -1.49×10^7 at $J{=}48$)
(ii) unitarity ceiling (scout)	—	—	zero slack used; $n_J^{(10)} \sim J^7$ makes negativity dearer than positivity
(iii) E_2/X_8 enlargement	$2.90 \to 2.63$	≥ 3.17; near-miss 2.9103 refused	single near-threshold audit window
(iv) closure re-solves	—	3.1023 ($J \leq 120$)	migrating channel from $J \approx 204$; finite-(J, m) eikonal corrections

Reproducibility and data availability

Every reported number is regenerable from the accompanying public repository, `https://github.com/mruckman1/gravitational-positivity-bounds`. Its layout: `qgse/verifiers/`—the verifier and certifier source, including the 21-check assertion battery, the soundness gate to run first; `results/positivity/artifacts/`—functional configurations and exact rational coefficients as JSON (the $J \leq 40$ champion, the margin- and $p_{\max}{=}1$ tail-hardened functionals with their tail scans and impact-parameter measurements, and the exact corner-window artifact of App. D); `results/positivity/drivers/`—standalone driver scripts reproducing every experiment of Secs. 5.2 and 6 and App. D, with captured outputs; `results/positivity/ablation/`—the Sec. 7 ablation (corrected harness, valley probe, per-seed run logs, and the pre-registered decision rule); `results/positivity/grav_rays.jsonl`—the append-only ledger of certified values; and `scripts/`—standalone kernel-validation scripts `verify_x8_stable.py` and `verify_e2_family.py`. No result depends on a proprietary solver at the certificate level: the bound (for the audited spins) is proven by exact rational arithmetic independent of the floating-point program that shaped the candidate.

Acknowledgments

The evolutionary search builds on an open-source evolutionary code [24] and the semidefinite-bootstrap toolchain [40]. Generative proposals and the adversarial judge were provided by commercial large-language-model APIs. The author thanks colleagues in the positivity-bounds community whose published work made the validation targets and the honest positioning of this note possible.

References

[1] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, "Causality, analyticity and an IR obstruction to UV completion," JHEP **10**, 014 (2006), arXiv:hep-th/0602178.

[2] X. O. Camanho, J. D. Edelstein, J. Maldacena, and A. Zhiboedov, "Causality constraints on corrections to the graviton three-point coupling," JHEP **02**, 020 (2016), arXiv:1407.5597 [hep-th].

[3] S. Caron-Huot and V. Van Duong, "Extremal effective field theories," JHEP **05**, 280 (2021), arXiv:2011.02957 [hep-th].

[4] S. Caron-Huot, D. Mazáč, L. Rastelli, and D. Simmons-Duffin, "Sharp boundaries for the swampland," JHEP **07**, 110 (2021), arXiv:2102.08951 [hep-th].

[5] N. Arkani-Hamed, T.-C. Huang, and Y.-t. Huang, "The EFT-hedron," JHEP **05**, 259 (2021), arXiv:2012.15849 [hep-th].

[6] A. Guerrieri, J. Penedones, and P. Vieira, "Where is string theory in the space of scattering amplitudes?," Phys. Rev. Lett. **127**, 081601 (2021), arXiv:2102.02847 [hep-th].

[7] A. Guerrieri, H. Murali, J. Penedones, and P. Vieira, "Where is M-theory in the space of scattering amplitudes?," JHEP **06**, 064 (2023), arXiv:2212.00151 [hep-th].

[8] G. Bossard and A. Loty, "Bounds on the next-to-leading Wilson coefficient in maximal supergravity," arXiv:2504.17840 [hep-th] (2025).

[9] K. Häring and A. Zhiboedov, "Gravitational Regge bounds," SciPost Phys. **16**, 034 (2024), arXiv:2202.08280 [hep-th].

[10] S. Dutta Chowdhury, A. Gadde, T. Gopalka, I. Halder, L. Janagal, and S. Minwalla, "Classifying and constraining local four photon and four graviton S-matrices," JHEP **02**, 114 (2020), arXiv:1910.14392 [hep-th].

[11] D. Chandorkar, S. Dutta Chowdhury, S. Kundu, and S. Minwalla, "Bounds on Regge growth of flat space scattering from bounds on chaos," JHEP **05**, 143 (2021), arXiv:2102.03122 [hep-th].

[12] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez, and D. Simmons-Duffin, "Causality constraints on corrections to Einstein gravity," JHEP **05**, 122 (2023), arXiv:2201.06602 [hep-th].

[13] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez, and D. Simmons-Duffin, "Graviton partial waves and causality in higher dimensions," Phys. Rev. D **108**, 026007 (2023), arXiv:2205.01495 [hep-th].

[14] C.-H. Chang and J. Parra-Martinez, "Graviton loops and negativity," JHEP **08**, 175 (2025), arXiv:2501.17949 [hep-th].

[15] J. Berman and N. Geiser, "Analytic bootstrap bounds on masses and spins in gravitational and non-gravitational scalar theories," JHEP **06**, 168 (2025), arXiv:2412.17902 [hep-th].

[16] L.-Y. Chiang, Y.-t. Huang, W. Li, L. Rodina, and H.-C. Weng, "(Non)-projective bounds on gravitational EFT," arXiv:2201.07177 [hep-th] (2022).

[17] Z. Bern, D. Kosmopoulos, and A. Zhiboedov, "Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude," J. Phys. A **54**, 344002 (2021), arXiv:2103.12728 [hep-th].

[18] J. Albert, W. Knop, and L. Rastelli, "Where is tree-level string theory?," JHEP **02**, 157 (2025), arXiv:2406.12959 [hep-th].

[19] M. Boisvert, W. Knop, and L. Rastelli, "Where is tree-level heterotic string theory?," arXiv:2606.09980 [hep-th] (2026).

[20] A. Sinha and A. Zahed, "Crossing symmetric dispersion relations in quantum field theories," Phys. Rev. Lett. **126**, 181601 (2021), arXiv:2012.04877 [hep-th].

[21] C. Pasiecznik, "Bootstrapping gravity with crossing symmetric dispersion relations," arXiv:2506.09884 [hep-th] (2025).

[22] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in Real Algebraic Geometry*, 2nd ed., Algorithms and Computation in Mathematics **10**, Springer (2006).

[23] T. C. Hales *et al.*, "A formal proof of the Kepler conjecture," Forum Math. Pi **5**, e2 (2017), arXiv:1501.02155 [math.MG].

[24] Sakana AI, "ShinkaEvolve: an open-source evolutionary program-search framework," https://github.com/SakanaAI/ShinkaEvolve (2025), Apache-2.0.

[25] R. T. Lange, Y. Imajuku, and E. Cetin, "ShinkaEvolve: towards open-ended and sample-efficient program evolution," arXiv:2509.19349 [cs.NE] (2025).

[26] G. Kántor, V. Niarchos, and C. Papageorgakis, "Solving conformal field theories with artificial intelligence," Phys. Rev. Lett. **128**, 041601 (2022), arXiv:2108.08859 [hep-th].

[27] T. Cai, G. W. Merz, F. Charton, N. Nolte, M. Wilhelm, K. Cranmer, and L. J. Dixon, "Transforming the bootstrap: using transformers to compute scattering amplitudes in planar $\mathcal{N} = 4$ super-Yang–Mills theory," Mach. Learn. Sci. Tech. **5**, 035073 (2024), arXiv:2405.06107 [hep-th].

[28] F. Bhat, D. Chowdhury, A. P. Saha, and A. Sinha, "Bootstrapping string models with entanglement minimization and machine-learning," Phys. Rev. D **111**, 066013 (2025), arXiv:2409.18259 [hep-th].

[29] M. A. Gümüş, D. Leflot, P. Tourkine, and A. Zhiboedov, "The S-matrix bootstrap with neural optimizers. Part I. Zero double discontinuity," JHEP **07**, 210 (2025), arXiv:2412.09610 [hep-th].

[30] A. Dersy, M. D. Schwartz, and A. Zhiboedov, "Reconstructing S-matrix phases with machine learning," JHEP **05**, 200 (2024), arXiv:2308.09451 [hep-th].

[31] S. Mizera, "Scattering with neural operators," Phys. Rev. D **108**, L101701 (2023), arXiv:2308.14789 [hep-th].

[32] V. Niarchos and C. Papageorgakis, "Learning S-matrix phases with neural operators," Phys. Rev. D **110**, 045020 (2024), arXiv:2404.14551 [hep-th].

[33] D. Simmons-Duffin, "A semidefinite program solver for the conformal bootstrap," JHEP **06**, 174 (2015), arXiv:1502.02033 [hep-th].

[34] D. Poland and D. Simmons-Duffin, "Bounds on 4D conformal and superconformal field theories," JHEP **05**, 017 (2011), arXiv:1009.2087 [hep-th].

[35] S. El-Showk and M. F. Paulos, "Bootstrapping conformal field theories with the extremal functional method," Phys. Rev. Lett. **111**, 241601 (2013), arXiv:1211.2810 [hep-th].

[36] H. Cohn, D. de Laat, and A. Salmon, "Three-point bounds for sphere packing," arXiv:2206.15373 [math.MG] (2022).

[37] H. Elvang, A. Herderschee, and R. Morales, "String theory from maximal supersymmetry," JHEP **07**, 105 (2026), arXiv:2601.11705 [hep-th].

[38] J. Berman, H. Elvang, and A. Herderschee, "Flattening of the EFT-hedron: supersymmetric positivity bounds and the search for string theory," JHEP **03**, 021 (2024), arXiv:2310.10729 [hep-th].

[39] J. Berman and H. Elvang, "Corners and islands in the S-matrix bootstrap of the open superstring," JHEP **09**, 076 (2024), arXiv:2406.03543 [hep-th].

[40] W. Landry and D. Simmons-Duffin, "Scaling the semidefinite program solver SDPB," arXiv:1909.09745 [hep-th] (2019).

[41] Q. Huangfu and J. A. J. Hall, "Parallelizing the dual revised simplex method," Math. Program. Comput. **10**, 119 (2018).

[42] A. Meurer *et al.*, "SymPy: symbolic computing in Python," PeerJ Comput. Sci. **3**, e103 (2017).