
IDPG: Infinitesimal Discrete Probability Geometry

A Nonstandard Axiomatic Framework with Loeb Measure and Information Structure

Version 3.9

AL_78 | Independent Research Group | August 2026

Important Note

This document represents a **conceptual and theoretical work**. Statements regarding continuum limits and full nonlinear geometric field mapping are **working hypotheses, not proven theorems**. The linearized field equations (Sec. 4.9) are postulated within the weak-field expansion; the discrete Bianchi identity (Corollary 1) and the $O(\varepsilon^2)$ error estimate (Theorem 2) are, however, derived rigorously from them within the hyperfinite framework. The author is an independent researcher, not affiliated with any institution, and does **not seek funding**. The connection to full geometric field theory remains an open problem. The framework is offered to **invite discussion and collaboration**. The discrete Bianchi identity (Corollary 1) is proven only for the flat hyperfinite torus; extension to curved backgrounds remains open.

Abstract

We present a closed axiomatic framework, Infinitesimal Discrete Probability Geometry (IDPG), built on nonstandard analysis and the Loeb measure construction. Within a hyperfinite probabilistic universe, IDPG introduces a discrete curvature deviation, constructed from local measure imbalance, and establishes its exact asymptotic order relative to the local entropy gradient at finite N ; we show that this order coincides with a naive $O(N^{-2\alpha})$ correspondence only at a single parameter value ($\alpha = 1$), and that the entropy gradient is otherwise asymptotically subdominant to the curvature deviation rather than predictive of it. IDPG does not claim that classical Ricci curvature is derived from Shannon entropy in general; it examines a specific, precisely-bounded relationship between two quantities defined within this hyperfinite measure geometry, jointly formulating probability, geometry, and information theory through a normalized infinitesimal measure at the hyperfinite level.

The main rigorous results are: (i) the exact order of the entropy-curvature deviation at finite N , and the parameter regime in which it is and is not informative (Theorem 1); (ii) a discrete Bianchi identity for the linearized field on the flat hyperfinite torus (Corollary 1, Sec. 5.2); (iii) $O(\varepsilon^2)$ convergence of the discrete geometric action for the linearized field to its continuum limit (Theorem 2, Lemma 2). A proposed graph generalization of (i) was tested

and its claimed identity withdrawn (Prop. 6); a valid graph-level generalization remains open. Classical probabilistic laws—normalization, independence, and the Borel–Cantelli lemma—are recovered as consistency checks of the measure layer. An information-theoretic layer is derived via the uniqueness theorem for entropy (Proposition 1). Computational verification in 1D and 2D and of the Bianchi identity in 2D (Sec. 8.1–8.3) supports the framework’s internal consistency within the tested parameter range; see Sec. 7 for what that range does not cover.

The extension of the variational principle to the full nonlinear metric, as well as its detailed identification with classical geometric field theory, is formulated as Open Problem 1 and left for future work. IDPG is compatible with standard quantum mechanical amplitudes as a special case (Sec. 4.6). It is offered as a theoretical contribution for mathematical discussion, with no claim of immediate empirical testability or algorithmic application. An explicit list of what is and is not established by this paper, separating the constructed measure-theoretic language from claims that were tested and did not survive, is given in the Conclusion.

Contents

1	Introduction	5
2	Related Work	6
2.1	Recent Developments in Nonstandard Analysis	6
2.2	Comparison with Discrete Geometric Formalisms	8
2.2.1	Regge Calculus [16]	8
2.2.2	Causal Dynamical Triangulations (CDT) [17]	8
2.2.3	Loop Quantum Gravity (LQG) [11]	8
2.2.4	Discrete Exterior Calculus (DEC) [18]	8
2.2.5	Summary Table	8
3	Measure Convention	10
4	Axiomatic System	10
4.1	Nonstandard Foundation	10
4.2	Loeb Construction	10
4.3	Infinitesimal Geometry	11
4.4	Information Theory	11
4.5	Independence Structure	11
4.6	Quantum Compatibility (Sketch)	11
4.7	Geometric Structure	11
4.8	Why Hyperfinite and Loeb, Beyond Large- N Notation	12
4.9	Linearized Geometric Structure	13
4.10	Axiomatic Minimality and Status	14
5	Consistency Checks and Main Theorems	15
5.1	Consistency Checks of the Measure Layer	15
5.2	Main Theorems	16
6	Extensions of IDPG	22
6.1	Discrete Curvature Structure	22
6.2	Thermodynamic Limit and Entropy Density	23
6.3	Rare-Event Geometry and Borel-Cantelli	23
6.4	Entropy-Geometry Duality	23
6.5	Unified Variational Functional	23
6.6	Geometric Variational Principle	24
7	Limitations and Open Questions	24

8	Computational Verification of the Main Results	28
8.1	Numerical Validation of the Entropy-Curvature Relation (1D Chain)	28
8.2	2D Numerical Validation of the Entropy-Curvature Relation	29
8.3	Numerical validation of the discrete Bianchi identity	34
9	Conclusion	37
A	Worked example: $K(i)$ and $\Delta H_{\text{loc}}(i)$ at $N = 4$	38
B	Variational derivation of the linearized field equations	38
C	Sketch of the nonlinear variational principle	39
D	Python implementation for the 1D entropy-curvature simulation	39

1. Introduction

Classical probability theory is formulated on σ -additive measure spaces, while nonstandard analysis provides an alternative foundation based on hyperfinite sets and infinitesimal quantities. In this work, we construct a framework that combines these perspectives into a single axiomatic system.

Scope and positioning. IDPG is a foundational mathematical framework situated at the intersection of nonstandard analysis, measure-theoretic probability, and discrete differential geometry. It is not intended as a machine-learning methodology paper; no novel training algorithm, benchmark result on standard datasets, or engineering application is claimed. IDPG is not proposed as an ML algorithm, a numerical simulation framework, or a phenomenological theory of gravity. Its purpose is foundational: to investigate whether probability, information, and geometry can be represented as mutually constrained manifestations of a single hyperfinite measure structure, rather than as three independently-founded formalisms that happen to be compared after the fact. The target contribution is a common axiomatic language in which geometric curvature and information-theoretic entropy are jointly generated from a single normalized measure structure. The numerical experiments in Sec. 8 are therefore not intended as empirical validation in the conventional experimental sense; their role is falsification-oriented consistency testing of consequences derived from the axiomatic system, including instances (Sec. 7) where such testing shows a derived consequence to be narrower, or an earlier claim to be wrong, than initially stated.

What IDPG contributes beyond a reformulation. A fair question, raised in review, is what IDPG offers beyond restating known asymptotics in unfamiliar notation. In one sentence: *IDPG does not claim that standard analysis is incapable of reproducing its individual asymptotic results; its claim is that probability, information, and discrete geometry are constructed as layers of a single hyperfinite object before the standard-part/continuum passage, and the mathematical results then determine which proposed correspondences survive, separate, or fail.* IDPG does not claim superiority over standard mathematical physics, nor that any single result here is unreachable by standard means (Sec. 4.8 addresses this directly). Its contribution is, specifically: (i) a common hyperfinite architecture in which probability, information, and discrete geometry are represented as internal objects on the same index set, prior to any continuum limit being taken; (ii) this architecture is what allows Theorem 1 to state precisely what survives the continuum limit (the curvature deviation $K(i)$, Lemma 1) and what does not (the entropy-gradient correspondence, Remark 11), within one internal calculation rather than two separately-founded ones compared after the fact; (iii) the nonlinear extension of this architecture—where a genuinely new geometric field theory, rather than a reformulation of the linearized case, would have to be built—is precisely the unresolved next step (Open Problem 1), not a claim already made here.

We consider a hyperfinite sample space equipped with an internal finitely additive probability measure. Through the Loeb construction [2], this internal measure is extended to a standard σ -additive probability space, allowing probabilistic systems to be interpreted as discretized geometric structures at an infinitesimal scale.

The central idea is that probability arises from normalized geometric volume in a hyperfinite universe, while information and entropy emerge as logarithmic invariants. A variational geometric structure is incorporated as a natural extension, yielding discrete field equations that admit a geometric interpretation in the continuum limit.

The resulting system provides a consistent foundation for studying rare-event asymptotics, entropy, and geometric variational principles in a common probabilistic-geometric setting—“unified” in the sense that all three layers share a single measure-theoretic origin, though the geometric layer’s linearized field structure (Axioms 18–20) is a postulated, weak-field extension of that origin rather than a derived consequence of it (Remark 4).

2. Related Work

Nonstandard analysis. Robinson [1] introduced hyperreal numbers and hyperfinite sets as a rigorous framework for infinitesimals.

Loeb measure theory. Loeb [2] constructed the bridge between internal finitely additive measures and classical σ -additive probability spaces via the standard-part map and Carathéodory extension.

Classical probability. Kolmogorov’s axiomatization [3] provides the baseline σ -additive framework. The present work formulates this structure within a hyperfinite setting.

Information theory. Shannon’s entropy [4] provides a logarithmic measure of uncertainty. In our framework, entropy is formulated within the measure structure via the uniqueness theorem for entropy functionals (Sec. 4.10).

Borel–Cantelli lemmas. The Borel–Cantelli lemmas describe asymptotic behavior of rare events; in IDPG they are consistent with the structure of infinite hyperfinite repetition [5].

Quantum foundations. The probabilistic interpretation of quantum mechanics was formalized by von Neumann [8] and Dirac [9]. In IDPG, the Born rule and unitary evolution are formulated within the geometric structure of the hyperfinite space.

Geometric field theory. The Einstein–Hilbert variational principle [10] provides a prototype for field equations derived from metric structure. IDPG provides a hyperfinite regularization of this variational structure, applicable to general metric fields, avoiding lattice artifacts through infinitesimal discretization [11].

2.1. Recent Developments in Nonstandard Analysis

Two recent works apply nonstandard analysis and Loeb measures to domains closely related to IDPG, though with different objectives:

Hyperfinite risk measures (Kania, 2026) [12]. Kania constructs hyperfinite probability spaces for coherent risk estimation in financial mathematics. The framework uses internal finitely additive measures and Loeb extension to represent extreme tail risks ($P \ll 10^{-300}$). While both Kania’s work and IDPG employ hyperfinite Loeb measures for rare-event analysis, the applications diverge: Kania targets *stochastic optimization* (risk-weighted expectations), whereas IDPG targets *geometric variational principles* (entropy-curvature duality, field equations). IDPG extends the hyperfinite formalism from static risk measures to dynamic geometric fields.

Hyperfinite Markov processes (Duanmu et al., 2021) [13]. Duanmu, Rosenthal, and Weiss establish ergodicity for Markov processes via nonstandard analysis, proving convergence of hyperfinite chains to continuous limits. Their framework provides a rigorous foundation for discrete-to-continuum transitions in stochastic dynamics. IDPG differs in scope: while Duanmu et al. focus on *temporal evolution* (Markov semigroups), IDPG focuses on *spatial geometry* (metric fields, curvature tensors, variational principles). The entropy-curvature correspondence (Theorem 1) has no analogue in their work, as it requires the geometric measure structure (Axioms 8–10) absent from Markov chains.

Remark 1 (Novelty of IDPG relative to existing nonstandard applications). *Existing nonstandard applications fall into two categories: (i) stochastic analysis (risk measures, Markov processes, stochastic calculus), where Loeb measures regularize path integrals; and (ii) foundational physics (Robinson’s framework, Loeb’s original construction), where nonstandard analysis provides alternative axiomatizations. IDPG bridges these categories by applying Loeb measures to geometric field theory: the hyperfinite sample space carries a metric structure (Axioms 14–17), and the measure itself generates curvature (Theorem 1). This geometric extension of the Loeb framework is novel.*

By contrast, the Lott–Villani–Sturm synthetic curvature [14, 15] characterises curvature through the convexity of entropy along W_2 -geodesics in the space of probability measures. IDPG inverts this logic: curvature is generated by the local entropy gradient on a single hyperfinite probability space, and no auxiliary Wasserstein geometry is required. The LVS framework is macroscopic (metric measure spaces), whereas IDPG is microscopic (infinitesimal grid). Consequently, the IDPG entropy-curvature correspondence is a pointwise identity at finite N , not a convexity inequality along optimal transport paths.

Comparison with Ollivier–Ricci and synthetic curvature. Ollivier’s discrete Ricci curvature [6] defines curvature on metric measure spaces via optimal transport between balls. In IDPG, $K(i)$ measures the deviation of the local geometric mean $\mu(B_r(i))/|B_r(i)|$ from the global mean; an earlier version of this paper claimed this coincides with the average Ollivier–Ricci curvature up to $O(\delta^2)$, but that specific quantitative claim was found to rest on an incorrect calculation and is withdrawn (Sec. 6.1). The two constructions differ structurally— $K(i)$ compares a local region to the global mean, Ollivier–Ricci compares neighboring local regions to each other—and no tested convention reproduces the claimed agreement.

However, IDPG differs in two respects: (i) the measure μ is derived from a nonstandard Loeb construction rather than imposed externally, and (ii) the entropy gradient ΔH_{loc} provides

an information-theoretic counterpart to curvature that has no direct analogue in Ollivier’s framework.

2.2. Comparison with Discrete Geometric Formalisms

The IDPG framework is not the first attempt to discretize geometric field structures. The following comparison is provided for contextual orientation; IDPG does not claim to reproduce or supersede the physical predictions of gravitational frameworks.

2.2.1. Regge Calculus [16]

Regge calculus replaces a smooth manifold with a simplicial complex (triangulation), assigning lengths to edges and computing curvature concentrated at hinges (deficit angles). A variational action is approximated by a sum over hinges. In IDPG, the discrete structure is a hyperfinite grid (not a simplicial complex), and curvature emerges from measure perturbations rather than deficit angles. The Loeb construction provides a σ -additive continuum limit that is not naturally present in Regge calculus, where the continuum limit requires careful refinement sequences.

2.2.2. Causal Dynamical Triangulations (CDT) [17]

CDT sums over causal triangulations, weighted by a discrete action, to define a non-perturbative path integral over geometric structures. IDPG does not perform a path integral; it constructs a single hyperfinite probabilistic universe with a variational principle. The entropy-curvature correspondence (Theorem 1) has no direct analogue in CDT, where entropy (if defined) is a thermodynamic property of horizons, not a local geometric generator.

2.2.3. Loop Quantum Gravity (LQG) [11]

LQG employs Ashtekar variables to construct a spin network basis with discrete spectra for geometric operators. IDPG does not quantize gravity; it regularizes geometry at a hyperfinite level and recovers classical structures in the Loeb limit. The spin network $\rightarrow$ spin foam dynamics of LQG has no counterpart here; IDPG’s variational principle (Open Problem 1) is classical, not quantum, though quantum amplitudes are compatible (Sec. 4.6).

2.2.4. Discrete Exterior Calculus (DEC) [18]

DEC defines differential forms, exterior derivatives, and Hodge stars on simplicial complexes, preserving Stokes’ theorem at the discrete level. IDPG uses centered finite differences on a hyperfinite torus (Axioms 18–20), which is computationally simpler but restricted to regular grids. DEC’s topological flexibility (arbitrary meshes) is traded for the analytic power of nonstandard analysis (infinitesimal ε , Loeb measure).

2.2.5. Summary Table

Key distinction: IDPG’s novelty lies in defining curvature directly from local measure perturbation within a single hyperfinite probabilistic structure, rather than imposing geometric

Feature	Regge	CDT	LQG	DEC	IDPG
Discrete structure	Simplicial complex	Triangulation	Spin network	Simplicial complex	Hyperfinite grid
Curvature source	Deficit angles	Regge action	Area/volume spectra	Discrete forms	Measure perturbations
Continuum limit	Refinement sequences	Ensemble average	Semiclassical states	Mesh refinement	Loeb standard-part map
Variational principle	Discrete EH action	Path integral	Hamiltonian constraint	None explicit	Entropy-geometric functional
Quantum compatibility	No	Intrinsic	Intrinsic	No	Compatible (Sec. 4.6)

Table 1: Comparison of key features across discrete geometric formalisms.

discretization ab initio; its relation to the local entropy gradient is examined precisely in Theorem 1, where the entropy gradient is shown to be asymptotically subdominant rather than a predictive proxy for curvature (Remark 11).

3. Measure Convention

Throughout, P denotes the normalized geometric measure defined in Axiom 10. We distinguish three related measures:

- P_H : internal finitely additive pre-measure on $\mathcal{A}_{int}$,
- P_L : Loeb measure (standard-part extension to $\mathcal{F}_L$),
- P : normalized geometric measure induced by μ .

Compatibility. μ is an internal measure; P is its normalized form after dividing by $\mu(\Omega_H)$, and P_L is its standard-part extension to the Loeb σ -algebra.

Notation glossary. The following symbols recur throughout the paper; each is defined precisely where first introduced (cross-referenced below) and is repeated here for quick reference.

Symbol	Meaning	First defined
N, ε	hyperfinite grid size; spacing $\varepsilon = 1/N$	Sec. 4
$P(i), \delta P(i)$	normalized measure; perturbation from uniform $P_0 = 1/N$	above; Thm. 1
$K(i)$	discrete curvature deviation (local vs. global mean)	Sec. 6.1
$H_{loc}(i), \Delta H_{loc}(i)$	local entropy; its discrete gradient	Thm. 1
α	perturbation-amplitude exponent, $\max_i \delta P(i) = O(N^{-1-\alpha})$	Thm. 1
r	ball radius defining $K(i)$ (fixed to $r = 1$ throughout)	Sec. 6.1
$G_{\alpha\beta}^{lin}(i)$	linearized discrete Einstein tensor	Sec. 4.9
$h_{\alpha\beta}(i)$	linear metric perturbation, $g_{\alpha\beta} = \delta_{\alpha\beta} + h_{\alpha\beta}$	Sec. 4.9
$st(\cdot)$	standard-part map (nonstandard analysis)	Sec. 4

4. Axiomatic System

4.1. Nonstandard Foundation

Axiom 1 (Hyperreal Structure). $\mathbb{R} \subset {}^*\mathbb{R}, \mathbb{N} \subset {}^*\mathbb{N}$.

Axiom 2 (Hyperfinite Universe). $\Omega_H = \{1, 2, \dots, N\}, N \in {}^*\mathbb{N} \setminus \mathbb{N}$.

Axiom 3 (Internal Algebra). $\mathcal{A}_{int} \subset \mathcal{P}(\Omega_H), \Omega_H \in \mathcal{A}_{int}$.

Axiom 4 (Internal Pre-Measure). $P_H : \mathcal{A}_{int} \rightarrow {}^*[0, 1], P_H(\Omega_H) = 1$.

Axiom 5 (Uniform Atomicity). $P_H(\{i\}) = 1/N$.

Axiom 6 (Finite Additivity). *For disjoint A, B , $P_H(A \cup B) = P_H(A) + P_H(B)$.*

4.2. Loeb Construction

Axiom 7 (Carathéodory–Loeb Extension). *There exists a unique countably additive measure P_L on $\mathcal{F}_L = \sigma(\mathcal{A}_{int})$ such that $P_L|_{\mathcal{A}_{int}} = st \circ P_H$.*

4.3. Infinitesimal Geometry

Axiom 8 (Infinitesimal Scale). $\exists \varepsilon \in {}^*\mathbb{R}^+ : 0 < \varepsilon < 1/n, \forall n \in \mathbb{N}$.

Axiom 9 (Geometric Measure). $\mu(\{i\}) = \varepsilon, \mu(\Omega_H) = N\varepsilon$.

Axiom 10 (Normalization). $P(A) = \mu(A)/\mu(\Omega_H)$.

4.4. Information Theory

Axiom 11 (Information Content). $I(E) = -\log P(E)$.

Axiom 12 (Entropy). $H(X) = -\sum_i P_i \log P_i$.

4.5. Independence Structure

Axiom 13 (Independence of Events). *For countable (E_k) , independence means $P_H(\bigcap_{k \in J} E_k) = \prod_{k \in J} P_H(E_k)$ for all finite J .*

4.6. Quantum Compatibility (Sketch)

Complex amplitudes $\psi(i)$ with $\sum_i |\psi(i)|^2 = 1$ are compatible with the framework. For constant phases the classical limit is recovered (Remark 2), while varying phases yield interference (Remark 2, equation below). No additional axioms are required. These observations are heuristic compatibility notes, not derived theorems; a full quantum extension remains beyond the scope of the present axiomatic layer.

Remark 2 (General quantum superposition). *The interference formula above extends to arbitrary superpositions $\psi = \sum_k c_k \psi_k$ with $\sum_k |c_k|^2 = 1$:*

$$P(C) = \sum_k |c_k|^2 P_k(C) + 2 \sum_{k < l} \Re \left(c_k c_l^* \sum_{i \in C} \psi_k(i) \psi_l^*(i) \right).$$

The classical limit is recovered when all relative phases vanish or when the supports of ψ_k are disjoint. This is a direct consequence of the measure structure and requires no additional axioms.

Remark 3 (Status of the quantum extension). *The compatibility of complex amplitudes $\psi(i)$ with the IDPG measure structure means that $|\psi(i)|^2$ can be identified with the normalized geometric density $P(i)$ of Axiom 10 without introducing additional axioms. This does not constitute a derivation of the Born rule from the Loeb measure; rather, it shows that the standard probabilistic interpretation of quantum mechanics is consistent with the hyperfinite framework. A full quantum extension—including operator spectra, unitary evolution as a geodesic flow on the hyperfinite state manifold, and the emergence of the Born rule from a variational principle on amplitudes—remains beyond the scope of the present axiomatic layer and is left as an open direction.*

4.7. Geometric Structure

Axiom 14 (Hyperfinite Metric). $g_{\alpha\beta}(i) \in {}^*\mathbb{R}$ symmetric, positive definite.

Axiom 15 (Discrete Connection). *Covariant derivative with discrete Christoffel symbols.*

Axiom 16 (Curvature Tensor). *$R^\alpha_{\beta\gamma\delta}$ from commutator of covariant derivatives.*

Axiom 17 (Correspondence Principle). *For any internal $Q(i)$ that is S -continuous on Ω_H —meaning $Q(i) \approx Q(j)$ (their difference is infinitesimal) whenever $i, j \in \Omega_H$ satisfy $i\varepsilon \approx j\varepsilon$ (their grid positions are infinitely close)—there exists a standard $q(x) = \text{st}(Q(i))$ in the limit $N \rightarrow \infty$, $\varepsilon \rightarrow 0$, $N\varepsilon = \text{const}$, with Loeb measure zero of disagreement. Here st denotes the standard-part map with respect to the order topology on ${}^*\mathbb{R}$, and $x = \text{st}(i\varepsilon)$ belongs to the compact standard domain $[0, 1]^D$ obtained from the hyperfinite grid. This is the standard nonstandard-analysis notion of S -continuity (Robinson [1], Ch. 3); it is the internal, pointwise counterpart of the “spectral content away from the Nyquist scale” intuition invoked informally in Sec. 7, and is the hypothesis actually used in the proof below. The S -continuity hypothesis is necessary: internal fields whose oscillation frequency is pinned to the grid scale (e.g. $Q(i) = (-1)^i$) fail it directly—taking $j = i + 1$, $i\varepsilon \approx j\varepsilon$ but $Q(i) - Q(j) = \pm 2$ is not infinitesimal—and such fields admit no pointwise standard part, since their values do not stabilize under any refinement sequence. This is not a merely formal safeguard: a direct numerical stress test (smooth, discontinuous, and Nyquist-scale fields; see the corresponding Limitations item in Sec. 7) shows convergence degrading exactly as S -continuity is progressively violated—from ordinary $O(N^{-1})$ convergence for smooth fields, through slower $O(N^{-1/2})$ convergence at an isolated jump, to no convergence whatsoever at the Nyquist scale.*

4.8. Why Hyperfinite and Loeb, Beyond Large- N Notation

A recurring question across reviews of this paper is whether the nonstandard formulation provides anything beyond a compact notation for a large- N limit that ordinary (“standard”) analysis could equally express. It is worth answering directly, since the honest answer is more specific than either extreme (“nonstandard analysis is essential” or “it is just notation”).

IDPG does *not* claim that nonstandard analysis produces different real numbers, or that any of the asymptotic rates in this paper (e.g. Theorem 1’s $O(N^{-1-\alpha})$) are unobtainable by a standard $N \rightarrow \infty$ argument—they are not; every numerical verification in Sec. 8 was carried out with ordinary floating-point arithmetic at finite, standard N . The claim is narrower and more specific: not that standard analysis cannot reproduce individual asymptotic statements one at a time, but that the hyperfinite formulation supplies the *primary object* of the theory—the internal structure from which $K(i)$, $H_{\text{loc}}(i)$, and $g_{\alpha\beta}(i)$ are all defined as coexisting internal quantities, with the standard-part map applied once, at the end, rather than built into each definition separately. Concretely, the construction is the sequential chain

$${}^*P_H \longrightarrow {}^*\mu_H \longrightarrow \mu_L = L({}^*\mu_H) \longrightarrow K(i), H_{\text{loc}}(i), g_{\alpha\beta}(i) \longrightarrow \text{st}(\cdot),$$

where each arrow is an internal (or Loeb-extension) construction step, not a limiting process: the internal pre-measure ${}^*P_H$ (Axiom 4) generates the internal geometric measure ${}^*\mu_H$ (Axiom 9), which the Loeb construction extends to the σ -additive μ_L (Axiom 7, Sec. 4); $K(i)$, $H_{\text{loc}}(i)$, and $g_{\alpha\beta}(i)$ are then internal functions of this single measure, defined and related to one another (Theorem 1) while still internal; only the final arrow, $\text{st}(\cdot)$, passes to standard (real-valued,

continuum) quantities, and only for the S-continuous ones (Axiom 17). Loeb measure and the standard-part map are therefore not decorative replacements for “let $N \rightarrow \infty$ ”; they are the specific steps in this chain that make $K(i)$, $H_{\text{loc}}(i)$, and $g_{\alpha\beta}(i)$ definable as related parts of one object, prior to and independent of which of them individually survives the final projection. Working entirely internally, one first builds a single hyperfinite object,

$${}^*\Omega, \quad P_H, \quad \mu_L, \quad g_{\alpha\beta}(i), \quad K(i), \quad H_{\text{loc}}(i),$$

at one fixed (but unlimited) hyperfinite N —probability, measure, geometry, and information all coexisting as internal objects on the *same* index set $\Omega_H = \{1, \dots, N\}$, related by the axioms of Sec. 4 without reference to any limiting process. Only afterward does one apply the standard-part map $\text{st}(\cdot)$ to recover standard (real-valued, continuum) quantities. This is the reverse of the usual order in a large- N construction, where one typically defines a sequence of standard objects indexed by N and takes a limit at the end; here the “limit” is not a separate step performed on each object individually, but a single projection applied uniformly to an already-unified internal structure.

The practical content of this reversal is Axiom 17 (Correspondence Principle): $\text{st}(\cdot)$ is not guaranteed to exist for an arbitrary internal $Q(i)$, only for S-continuous ones. This is a genuine structural fact about the construction, not a notational convenience—it is what makes $Q(i) = (-1)^i$ a well-posed internal object with *no* standard part (Sec. 4), a distinction that has no direct counterpart in a bare large- N limit (where one would simply say “the limit does not exist” without the intermediate internal object being visible as a mathematical entity in its own right, with its own algebra, prior to the limit being taken). The value of the hyperfinite formulation is this architectural unification—one internal system, related by exact (not asymptotic) internal identities, projected once—rather than a claim that standard analysis is incapable of reaching the same asymptotic conclusions object-by-object.

4.9. Linearized Geometric Structure

For the weak-field expansion $g_{\alpha\beta}(i) = \delta_{\alpha\beta} + h_{\alpha\beta}(i)$ with $|h_{\alpha\beta}| \ll 1$, the following operators are defined on the hyperfinite torus $\Omega_H = (\mathbb{Z}_N)^D$ with periodic boundary conditions:

Axiom 18 (Centered Difference). $(\nabla_\gamma f)(i) = \frac{f(i + \hat{e}_\gamma) - f(i - \hat{e}_\gamma)}{2\varepsilon}$.

Axiom 19 (Discrete Laplacian). $\Delta_d f(i) = \sum_{\gamma=1}^D \frac{f(i + \hat{e}_\gamma) + f(i - \hat{e}_\gamma) - 2f(i)}{\varepsilon^2}$.

Axiom 20 (Linearized Einstein Tensor). $G_{\alpha\beta}^{\text{lin}}(i) = \frac{1}{2}(\nabla_\gamma \nabla_\alpha h_{\beta\gamma} + \nabla_\gamma \nabla_\beta h_{\alpha\gamma} - \Delta_d h_{\alpha\beta} - \nabla_\alpha \nabla_\beta h) - \frac{1}{2}\delta_{\alpha\beta}(\nabla_\gamma \nabla_\delta h_{\gamma\delta} - \Delta_d h)$, where $h = h_{\gamma\gamma}$ and summation over repeated indices is implied.

Remark 4 (Why this discretization: design choice, not derivation). *Axioms 18–20 are postulated, not derived from Axioms 1–17; the paper does not claim otherwise. They are, however, not an arbitrary choice among discretizations of the continuum linearized Einstein tensor. Three requirements jointly pin them down: (i) translation invariance under $\mathbb{Z}_N$, so the operator must be a fixed linear combination of $f(i \pm \hat{e}_\gamma)$ and $f(i)$; (ii) ∇_γ must be parity-odd (centered) so*

that $[\nabla_\alpha, \nabla_\beta] = 0$ exactly on the torus, which is what Corollary 1’s proof requires; (iii) Δ_d and ∇_γ must reduce to the continuum Laplacian and gradient at leading order as $\varepsilon \rightarrow 0$. Centered differences and the standard discrete Laplacian are the unique lowest-order stencils satisfying all three simultaneously; $G_{\alpha\beta}^{\text{lin}}$ is then the unique combination of these operators reproducing the continuum linearized Einstein tensor at leading order. This narrows the choice considerably but does not make it a theorem: a different discretization satisfying (i)–(iii) at higher order would also be admissible, and nothing in Axioms 1–17 forces (i)–(iii) themselves. In implication form: Axioms 1–17 alone do not entail $G_{\alpha\beta}^{\text{lin}}$ or the Bianchi identity; rather, Axioms 1–17 together with the explicit geometric postulates (i)–(iii) entail the IDPG geometric sector developed in Sec. 4.9. The relevant question is therefore not “is this discretization inevitable?” but “are (i)–(iii) themselves well-motivated design constraints?”, to which Sec. 6.1 and this remark are the paper’s answer. The correspondence between this discrete structure and the entropy layer (Sec. 6.1) is a further postulate, not a consequence of the above.

4.10. Axiomatic Minimality and Status

Remark 5 (On formulated vs. postulated structure). Table 2 clarifies the logical status of each axiom layer. The information-theoretic Axioms 11–12 are not externally imposed; they are uniquely determined by the measure structure (see Proposition 1).

Layer	Status	Justification
Axioms 1–6 (Nonstandard foundation)	Postulated	Robinson’s framework
Axioms 7–10 (Loeb + Geometry)	Postulated	Loeb construction
Axioms 11–12 (Information)	Formulated within	Uniqueness theorem from Axioms 9–10 + 13 (Prop. 1)
Axiom 13 (Independence)	Postulated	Base structure
Axioms 14–17 (Geometry)	Postulated	Geometric extension
Axioms 18–20 (Linearized geometry)	Postulated	Weak-field expansion on flat torus

Table 2: Logical status of axiom layers in IDPG.

Remark 6 (Minimality of the axiom set). Table 2 shows that the information-theoretic layer (Axioms 11–12) and the linearised geometric layer (Axioms 18–20) introduce no independent postulates. Axioms 11–12 are uniquely determined by the measure and independence structure (Proposition 1), while Axioms 18–20 are derived by weak-field expansion of the postulated geometric structure (Axioms 14–17). Consequently, the IDPG system rests on only three independent postulated layers: nonstandard foundations (Axioms 1–6), Loeb–geometric measure (Axioms 7–10), and nonlinear geometric structure (Axioms 13–17). This minimality is deliberate: it reduces the risk of internal inconsistency and isolates the open problems (Open Problem 1) to a single layer.

Proposition 1 (Uniqueness of entropy functional). Given Axioms 9–10 (geometric measure and normalization) and Axiom 13 (independence), the functional $H(X) = -\sum_i P_i \log P_i$ is the unique (up to multiplicative constant) continuous functional satisfying:

- (i) H is symmetric in its arguments;

(ii) $H(E \cap F) = H(E) + H(F)$ for independent events E, F ;

(iii) H is maximal for the uniform distribution $P_i = 1/N$.

Proof. This is the standard uniqueness theorem for Shannon entropy [7], formulated within the internal measure P_H . Condition (ii) follows from Axiom 13 and the identity $\log(ab) = \log a + \log b$ in ${}^*\mathbb{R}$. Maximality (iii) follows from Jensen's inequality for the concave function $-x \log x$, valid in the hyperreals by transfer. The constant is fixed by $H(1/2, 1/2) = \log 2$. $\square$

5. Consistency Checks and Main Theorems

5.1. Consistency Checks of the Measure Layer

Remark 7. Propositions 2–5 below are elementary consequences of Axioms 1–10. Their role is to verify that the internal measure P_H and the Loeb extension P_L agree on basic probabilistic identities. They are not claimed as original results; the non-trivial structure appears in Theorems 1–2.

Proposition 2 (Normalization). $P_H(\Omega_H) = 1$.

Proof. $\sum_{i=1}^N 1/N = 1$. $\square$

Proposition 3 (Measure Consistency). $P(A) = \mu(A)/\mu(\Omega_H)$.

Proof. Axiom 10. $\square$

Proposition 4 (Borel–Cantelli). For independent E_k with $P(E_k) = p > 0$, $P(E_k \text{ i.o.}) = 1$.

Proof. $\sum P(E_k) = \infty$, second Borel–Cantelli lemma. $\square$

Remark 8 (Independence is essential). The independence hypothesis in Proposition 4 cannot be dropped. Tested numerically with M correlated events constructed as nested subsets of a common base event, the union probability equals the base event's probability exactly ($P(\bigcup_k E_k) = P(\text{base})$), while the independence-based product formula $1 - \prod_k (1 - P(E_k))$ overestimates it by more than an order of magnitude (0.599 vs. the true 0.010 at $M = 100$). This is an elementary consequence of the definitions rather than a new result, but it confirms that the internal measure structure (Axioms 1–10) enforces the standard probabilistic constraint and does not silently extend the independent-case formula to correlated events.

Proposition 5 (Entropy Additivity). If $X \perp Y$, $H(X, Y) = H(X) + H(Y)$.

Proof. $P(X, Y) = P(X)P(Y)$. $\square$

5.2. Main Theorems

Remark 9 (Regularity requirements: Lemma 1 vs. Lemma 2). *Lemma 1* requires a C^2 standard-part field to ensure the existence of the limit of $K(i)/\varepsilon^2$ alone (not, as an earlier version claimed, a joint limit with the entropy gradient—see the withdrawal in that lemma); the rate of convergence is not controlled at this level of regularity. *Lemma 2* strengthens the hypothesis to C^4 in order to obtain the explicit $O(\varepsilon^2)$ finite-difference error bound via Taylor cancellation of odd terms. The two lemmas are therefore complementary: the first guarantees convergence of the curvature-deviation field on its own, the second quantifies the rate of that convergence.

Theorem 1 (Order of the Entropy-Curvature Deviation). *Let $K(i) = \frac{1}{|B_1(i)|} \sum_{j \in B_1(i)} P(j) - \frac{1}{N} \sum_j P(j)$ be the discrete curvature deviation (the local ball average of P minus its global mean; formally introduced with the general-radius definition in Sec. 6.1, and computed by hand for a small example in Appendix A), and let $H_{\text{loc}}(i) = -P(i) \log P(i)$ be the local entropy—the i -th summand of the global Shannon entropy $H(X) = \sum_i H_{\text{loc}}(i)$ of Axiom 12, evaluated pointwise rather than summed; the two should not be confused despite the shared letter H . Write $\Delta H_{\text{loc}}(i) := H_{\text{loc}}(i+1) - H_{\text{loc}}(i)$ for its discrete entropy gradient. For the uniform measure $P(i) = 1/N$, $K(i) \equiv 0$. For any internal perturbation $\delta P(i)$ with $\text{st}(\sum_i \delta P(i)) = 0$, $\max_i |\delta P(i)| = O(N^{-1-\alpha})$ for some $\alpha > 0$, and satisfying the grid-scale regularity condition*

$$|\delta P(i+1) - \delta P(i)| = O(N^{-1}) \max_i |\delta P(i)|, \quad |\delta P(i+1) + \delta P(i-1) - 2\delta P(i)| = O(N^{-2}) \max_i |\delta P(i)|, \quad (1)$$

(satisfied whenever $\delta P(i) = f(i/N)$ for a fixed C^2 function f , as in every deterministic test in Sec. 8; this condition is necessary and is violated by rough perturbations such as fractional Gaussian noise, see Sec. 7):

(i) $K(i) = \delta P(i) + O(N^{-2}) \cdot |\delta P(i)|$, uniformly in i ; in particular $K(i) = O(N^{-1-\alpha})$.

(ii) $\Delta H_{\text{loc}}(i) = o(K(i))$ as $N \rightarrow \infty$, uniformly in i , for every fixed $\alpha > 0$.

Consequently

$$K(i) - \Delta H_{\text{loc}}(i) = O(N^{-1-\alpha}), \quad (2)$$

and this exponent is tight (not improvable) for every $\alpha \neq 1$: by (ii), $K(i) - \Delta H_{\text{loc}}(i) = K(i)(1 + o(1))$, and the coefficient of the leading term is exactly 1—not a Taylor coefficient that could accidentally vanish—so the order of the difference equals the order of $K(i)$ itself, which is $\Theta(N^{-1-\alpha})$ whenever $\max_i |\delta P(i)|$ is itself $\Theta(N^{-1-\alpha})$ (as opposed to merely $O(\cdot)$), e.g. for the sinusoidal and Gaussian-bump perturbations used throughout Sec. 8. At $\alpha = 1$ the exponent $-(1 + \alpha) = -2$ coincides numerically with $-2\alpha = -2$, so the bound can equivalently be written as $O(N^{-2\alpha})$ at this single value of α —this coincidence of exponents, not a genuine cancellation between K and ΔH_{loc} , is the origin of the $O(N^{-2\alpha})$ form used in the numerical sections below, where $\alpha = 1$ throughout.

Proof. (i) follows from Taylor expansion of the ball average defining $K(i)$ in $1/N$; the first correction to $\delta P(i)$ enters at relative order N^{-2} with no cancellation of the leading term, using the second-difference part of the regularity hypothesis. (ii) follows from expanding $H_{\text{loc}}(i) =$

$H(P_0 + \delta P(i))$ around $P_0 = 1/N$: $H_{\text{loc}}(i) = H(P_0) + (\ln N - 1)\delta P(i) - \frac{N}{2}\delta P(i)^2 + O(\delta P^3)$, so $\Delta H_{\text{loc}}(i)$ is a first discrete difference of a quantity already of order $\delta P(i)$; the first-difference part of the regularity hypothesis is exactly what licenses replacing this discrete difference by an extra factor of order N^{-1} (up to the $\ln N$ prefactor on the linear term)—without it, $\delta P(i+1) - \delta P(i)$ need not be smaller than $\delta P(i)$ itself, and the argument fails, which is precisely what happens for the rough (fractional-Gaussian-noise) perturbations of Sec. 7. Under the hypothesis, $\Delta H_{\text{loc}}(i) = o(\delta P(i)) = o(K(i))$ for every fixed $\alpha > 0$. The claim follows from (i) and (ii) since $K(i) - \Delta H_{\text{loc}}(i) = K(i)(1 + o(1))$.

On the hyperreal remainder terms. The bounds in (i) and (ii) are first proved as ordinary real statements: for every standard $N_0 \in \mathbb{N}$ and every standard $N \geq N_0$, the Taylor remainders above are bounded by explicit real constants times $N^{-2}|\delta P(i)|$ and $N^{-1}|\delta P(i)|$ respectively (Lagrange form, using boundedness of the relevant derivatives of H away from $P = 0$, which holds since $P_0 + \delta P(i)$ stays bounded away from 0 for N large). This is a first-order statement about all standard natural numbers $N \geq N_0$ with explicit, N -independent constants; by the Transfer Principle it holds verbatim for all $N \in {}^*\mathbb{N}$, including the hyperfinite N of Sec. 4. For hyperfinite N , the bound $O(N^{-2}) \cdot |\delta P(i)|$ (resp. $o(K(i))$) is then a genuine hyperreal inequality with the same real constant; since N is infinite, both sides are infinitesimal, and the standard-part map applied to the inequality (order-preserving, Robinson [1]) yields $\text{st}(K(i)) = \text{st}(\delta P(i))$ and $\text{st}(\Delta H_{\text{loc}}(i)) = 0$ at any near-standard i , consistent with (i)–(ii). No separate nonstandard argument is needed beyond this transfer step, since the Taylor expansions themselves are purely algebraic identities valid internally by Transfer applied to the (real, standard) definitions of H and of the ball average. $\square$

Remark 10 (Non-fixed α : a slowly vanishing sequence $\alpha_N \rightarrow 0$). *Theorem 1 is stated for fixed $\alpha > 0$; it is natural to ask whether letting $\alpha = \alpha_N \rightarrow 0$ as $N \rightarrow \infty$ (a slower amplitude decay) opens an intermediate regime where ΔH_{loc} is no longer subdominant. Under the corrected exponent established here, it does not: since $\alpha_N \geq 0$, $N^{-1-\alpha_N} \leq N^{-1} \rightarrow 0$ for any rate of vanishing, so $K(i) - \Delta H_{\text{loc}}(i) = O(N^{-1-\alpha_N}) \rightarrow 0$ regardless of how slowly $\alpha_N \rightarrow 0$. This is worth noting explicitly because it would not have held under the earlier (incorrect) $O(N^{-2\alpha})$ form: for $\alpha_N = c/\ln N$, $N^{-2\alpha_N} = e^{-2c}$ is a nonvanishing constant, which would have been a genuine breakdown of that bound. The corrected exponent's N^{-1} factor—coming directly from the amplitude bound on K itself, independent of α —removes this failure mode. Numerically, for $\alpha_N = 1/\ln N$ (sinusoidal perturbation), the ratio $\max_i |\Delta H_{\text{loc}}(i)| / \max_i |K(i)|$ falls from 0.23 at $N = 10^2$ to 8.1×10^{-5} at $N = 10^6$, confirming subdominance persists rather than opening a new regime at this level of generality. A genuinely different scaling behavior would require $\alpha_N < 0$ (amplitude not decaying, or growing), which falls outside the perturbative regime the theorem addresses at all, or a perturbation shape that simultaneously degrades in regularity as $N \rightarrow \infty$ (interpolating toward the rough perturbations of Sec. 7); we do not analyze that two-parameter family here.*

Remark 11 (What this theorem does and does not establish). *Despite its name, Theorem 1 is best read as a structural asymptotic separation result, not a correspondence in the sense of one quantity predicting the other: it identifies a common measure-theoretic origin for $K(i)$ and $\Delta H_{\text{loc}}(i)$ while showing that their continuum scaling separates, with the entropy term parametri-*

cally subdominant. This separation is itself informative, not merely a negative finding: a priori, the coincidence of exponents at $\alpha = 1$ could have reflected a genuine leading-order cancellation between the two quantities (the reading suggested by earlier drafts of this paper and by the name “entropy-curvature correspondence”); part (ii) rules this out directly, showing the coincidence is an artifact of the exponents $1 + \alpha$ and 2α meeting at $\alpha = 1$, with no cancellation mechanism behind it. Distinguishing a genuine correspondence from an exponent coincidence is exactly the kind of claim a numerical consistency check alone cannot resolve, but that this asymptotic argument does.

The relation above shows that both $K(i)$ and $\Delta H_{\text{loc}}(i)$ vanish in the continuum limit, and pins down their joint order precisely. It does not show that $\Delta H_{\text{loc}}(i)$ is a leading-order description of $K(i)$: by part (ii), ΔH_{loc} is parametrically smaller than K for every tested α , so the difference $K - \Delta H_{\text{loc}}$ is dominated by K itself rather than by a genuine cancellation between the two quantities. Numerically, this is why $\max_i |K(i) - \Delta H_{\text{loc}}(i)| / \max_i |K(i)| \rightarrow 1$ as $N \rightarrow \infty$ at the value $\alpha = 1$ used throughout Sec. 8 (verified directly for sinusoidal and Gaussian-bump perturbations, $\alpha \in \{0.5, 1, 1.5, 2, 3\}$). We retain the name “entropy-curvature relation” for continuity with the numerical sections below, but the reader should not interpret it as demonstrating that local entropy gradients explain or predict discrete curvature deviations at the parameter values tested in this paper. Nor should the subdominance of this particular entropy gradient be read as showing that information and geometry are unrelated in IDPG more broadly: both are constructed from the same underlying measure (Sec. 6.1, Sec. 3), and the theorem constrains only the specific pointwise functional $\Delta H_{\text{loc}}(i) = H_{\text{loc}}(i+1) - H_{\text{loc}}(i)$ tested here, not every possible entropy-based curvature proxy.

Remark 12 (On the regularity of the continuum limit). The continuum limit of $K(i)/\varepsilon^2$ invoked in Lemma 1 assumes that the standard-part field $R(x)$ is at least C^2 on the compact domain (the entropy-gradient side of the earlier joint claim is withdrawn there, and is not assumed here). For C^4 metrics the convergence rate is controlled by Lemma 2. Extension to C^1 or Sobolev H^1 fields remains open: the finite-difference commutator $[\nabla_\alpha, \nabla_\beta]$ may fail to vanish in the distributional sense, and the standard-part map may produce non-unique limits at points of discontinuity. We do not claim that the continuum limit exists for arbitrary measurable fields; rather, the Loeb construction guarantees a σ -additive limit for internal smoothings, and the identification with classical differential geometry is rigorous only under the stated smoothness hypotheses.

Proposition 6 (Graph generalization: status). For an arbitrary hyperfinite graph with adjacency matrix A_{ij} and degree $d_i = \sum_j A_{ij}$, define the graph Laplacian $L_{ij} = d_i \delta_{ij} - A_{ij}$ and the graph-averaged curvature deviation $K(i) = \frac{1}{d_i} \sum_j A_{ij} \delta P(j)$ (replacing the nearest-neighbor ball average of Sec. 6.1 with the graph average). A previous version of this result claimed the identity $K(i) = \sum_j L_{ij} H_{\text{loc}}(j) + O(\|\delta P\|^2)$, generalizing Theorem 1 to arbitrary graphs, with the 1D chain ($A_{i,i\pm 1} = 1$) as a special case. This claim is withdrawn: the proof step invoked $\sum_j A_{ij} = d_i$ to eliminate the first-order term, but the correct first-order expansion of $\sum_j L_{ij} H_{\text{loc}}(j)$ is

$$\sum_j L_{ij} H_{\text{loc}}(j) = (\ln P_0 + 1) d_i [K(i) - \delta P(i)] + O(\|\delta P\|^2), \quad (3)$$

not $K(i)$ itself. By the order established in Theorem 1(i), $K(i) - \delta P(i) = O(N^{-2})|\delta P(i)|$, so the

right-hand side is smaller than $K(i)$ by a further factor of order $N^{-2} \ln N$; the two sides of the claimed identity are not the same order in N , and the identity does not hold as stated. Numerical tests on irregular (Erdos–Renyi, $p = 0.05$, and Barabasi–Albert, $m = 3$) graphs confirm this directly, now extended to $N = 100$ – 6000 with corrected float64 arithmetic (see below): the correlation $|\text{corr}(K, \sum_j L_{ij} H_{\text{loc}}(j)/d_i)|$ decays from 0.66 (ER) / 0.22 (BA) at $N = 100$ to 0.05 (ER) / 0.01 (BA) at $N = 6000$, consistent with a power law of exponent ≈ -0.6 for both graph families (least-squares fit on $\log |\text{corr}|$ vs. $\log N$)—weak at every tested N and vanishing as $N \rightarrow \infty$, not merely noisy. On a regular 1D chain, by contrast, the correlation is exactly 1 at every N —not because the identity holds, but because the sinusoidal test perturbation is an exact eigenfunction of the chain’s graph Laplacian, so $K(i)$ and $\sum_j L_{ij} H_{\text{loc}}(j)/d_i$ remain proportional in shape regardless of their (very different, per the order argument above) magnitudes; Pearson correlation is scale-invariant and does not detect this magnitude mismatch on the chain. (An earlier version of this section reported the chain correlation as degrading with N due to a float32 precision artifact in computing the required magnitude ratio; recomputed in float64, it is exactly 1 at every N tested, 10^2 to 6×10^3 , consistent with the eigenfunction argument and confirming the earlier number was a computational artifact, not a mathematical finding.) The chain case therefore illustrates why correlation alone is an unreliable diagnostic for this identity—the irregular-graph tests, where no such shape-coincidence is expected a priori, are the more informative check, and they fail, with a clean and now precisely quantified decay. A valid graph generalization of Theorem 1, if one exists, requires a different construction than the one proposed here; we leave it as an open problem.

Remark 13 (Is this a lattice artifact, or a symptom of Theorem 1?). *It is natural to ask whether the graph-generalization failure above reflects something special about the regular 1D chain or torus—i.e. whether entropy-curvature correspondence is real on regular lattices and merely fails to generalize—or whether it is consistent with, and explained by, the more basic finding of Theorem 1 and Remark 11 that ΔH_{loc} is asymptotically subdominant to K even on the 1D chain itself. The evidence supports the latter. The graph proposition’s own failure mode is a direct instance of the same order mismatch (the right-hand side of its would-be identity is smaller than $K(i)$ by $O(N^{-2} \ln N)$, exactly the mechanism of Theorem 1(ii)); the Ollivier–Ricci comparison failed for the same structural reason (Sec. 6.1); and Sec. 7 documents the identical degradation on a 3D torus, under ε -rescaling, and under non-uniform base measures. None of these generalize the chain result correctly for the same underlying reason, not five unrelated reasons. So: the regular lattice is not privileged in a way that made the original correspondence “real” there and merely non-transferable elsewhere—the correspondence was already only order-coincidental on the chain itself (Remark 11), and every attempted generalization inherits that fact rather than introducing a new failure.*

Remark 14 (Limitations of Corollary 2). *Corollary 1 is proven only for the flat hyperfinite torus $\Omega_H = (\mathbb{Z}_N)^D$ with periodic boundary conditions. The proof relies crucially on the commutativity of centered differences, $[\nabla_\alpha, \nabla_\beta] = 0$, which holds on the torus. For curved backgrounds (non-zero Christoffel symbols), see the remark below; for non-periodic boundaries, this was tested directly rather than left as a qualitative expectation. Replacing the periodic wrap with Dirichlet ($h_{\alpha\beta} = 0$ outside the domain) or Neumann (zero-gradient) padding in the same construction as Sec. 8.3 ($N = 32$ to 512 , 200 seeds per point), the violation $\max_i |\nabla_\alpha G_{\alpha\beta}^{\text{lin}}(i)|$ separates cleanly by distance*

from the boundary: cells more than a few grid spacings from the edge decay at $\approx N^{-4.9}$, statistically indistinguishable from the periodic bulk rate reported in Table 7, confirming the violation is spatially localized to the boundary as expected. At the boundary itself, however, the two boundary conditions behave qualitatively differently: under Neumann conditions the boundary violation decays as $\approx N^{-1.1}$ —slower than the bulk by roughly 3.8 orders of magnitude in exponent, but still vanishing as $N \rightarrow \infty$ —while under Dirichlet conditions the boundary violation is essentially flat, decaying at only $\approx N^{-0.1}$ (a 26% change in value over a $16\times$ range of N , from 2.96×10^{-4} at $N = 32$ to 2.19×10^{-4} at $N = 512$). This indicates the Dirichlet boundary defect is a genuine $O(1)$ effect that persists in the continuum limit, not a discretization artifact that refines away; Corollary 1 should be understood as failing at a Dirichlet boundary even as $N \rightarrow \infty$, while a Neumann boundary is compatible with the corollary only in the bulk, with a vanishing but slowly-decaying boundary layer. A general proof (or refutation) for arbitrary hyperfinite manifolds and boundary conditions remains open.

Remark 15 (Curved backgrounds require a covariant discretization: an open problem). Axiom 18 defines ∇_γ as a fixed coordinate (index-shift) difference operator, with no dependence on any background metric. As such, $[\nabla_\alpha, \nabla_\beta]f(i) \equiv 0$ identically for any scalar grid function f and any assignment of $g_{\alpha\beta}(i)$, since shifts along distinct lattice axes commute exactly. A previous version of this remark incorrectly claimed a nonzero commutator $\Gamma_{\alpha\beta}^\gamma \nabla_\gamma f + O(\varepsilon)$ and a corresponding curvature correction to Corollary 1’s divergence identity; both the algebraic argument and the predicted correction term were tested numerically against a direct computation of $\sum_\alpha \nabla_\alpha G_{\alpha\beta}^{\text{lin}}(i)$ on a non-conformal curved background and found to be uncorrelated with the actual violation (Pearson $r \approx 0$ across $N = 32$ –1024) and negligible in magnitude (relative error ≈ 1). The claim is withdrawn. Extending Corollary 1 to curved backgrounds requires ∇_γ itself to be replaced by a covariant discrete derivative incorporating connection coefficients—a genuinely different construction from Axiom 18, not a correction term added after the fact—and this construction is not given here; it remains open.

Corollary 1 (Discrete Bianchi Identity). *For any symmetric $h_{\alpha\beta}$ with periodic boundary conditions on the flat torus:*

$$\sum_{\alpha=1}^D \nabla_\alpha G_{\alpha\beta}^{\text{lin}}(i) = 0, \quad \forall i \in \Omega_H, \beta \in \{1, \dots, D\}. \quad (4)$$

Proof. The linearized Ricci tensor $R_{\alpha\beta}^{\text{lin}}$ is constructed from second finite differences of $h_{\alpha\beta}$. On the flat torus, centered differences commute: $[\nabla_\alpha, \nabla_\beta] = 0$. Therefore the same algebraic identity as in the continuum holds: $\nabla_\alpha R_{\alpha\beta}^{\text{lin}} = \frac{1}{2} \nabla_\beta R^{\text{lin}}$, which implies $\nabla_\alpha G_{\alpha\beta}^{\text{lin}} = \nabla_\alpha R_{\alpha\beta}^{\text{lin}} - \frac{1}{2} \nabla_\beta R^{\text{lin}} = 0$. $\square$

Lemma 1 (Curvature Convergence). *Let $K(i)$ be the discrete curvature deviation on the hyperfinite torus $\Omega_H = (\mathbb{Z}_N)^D$ with spacing $\varepsilon = 1/N$, for an internal perturbation $\delta P(i)$ satisfying all three hypotheses of Theorem 1 with $\alpha = 1$ fixed: (a) $\text{st}(\sum_i \delta P(i)) = 0$; (b) $\max_i |\delta P(i)| = O(N^{-2})$; (c) the grid-scale regularity condition of Theorem 1 (bounded first and second discrete differences relative to $\max_i |\delta P(i)|$). Assume in addition that the internal field $i \mapsto K(i)/\varepsilon^2$ has a standard part $R(x) = \text{st}(K(i)/\varepsilon^2)$, $x = \text{st}(i\varepsilon)$, that is C^2 on the compact domain $[0, 1]^D$. Then*

in the Loeb limit $N \rightarrow \infty$, $\varepsilon \rightarrow 0$, $N\varepsilon = \text{const}$,

$$\frac{K(i)}{\varepsilon^2} \rightarrow R(x).$$

Under hypotheses (a)–(c) alone (i.e. without any further assumption on ΔH_{loc}), the corresponding rescaled entropy gradient $\Delta H_{\text{loc}}(i)/\varepsilon^3$ does not converge to a finite standard part: by Theorem 1(ii), it grows as $\Theta(\ln N)$ (verified numerically, sinusoidal perturbation: fit $3.19 \ln N - 3.60$, residuals $< 0.6\%$ of the leading term over $N = 10^3$ – 10^7). Consequently there is no differential relation of the form $R(x) = c \cdot \nabla H(x)$ for any constant c ; an earlier version of this lemma claimed such a relation, derived from an incorrect assumption that $\Delta H_{\text{loc}}(i)/\varepsilon^3$ converges. That claim is withdrawn. The standard-part field $R(x)$ exists and is well-defined on its own, independent of any entropy-gradient counterpart.

Proof. Convergence of $K(i)/\varepsilon^2$ follows from Theorem 1(i): $K(i) = \delta P(i) + O(\varepsilon^2)|\delta P(i)|$ under hypothesis (c), and at $\alpha = 1$, hypothesis (b) gives $\delta P(i) = O(\varepsilon^2)$, so $K(i)/\varepsilon^2$ is an internal field bounded uniformly in N ; its standard part exists at each near-standard x by the assumed C^2 regularity of the limit, and equals $R(x)$ by definition. The divergence of $\Delta H_{\text{loc}}(i)/\varepsilon^3$ follows from the proof of Theorem 1(ii): under hypothesis (c), the leading term of $\Delta H_{\text{loc}}(i)$ carries a factor $(\ln N - 1) = \Theta(\ln(1/\varepsilon))$, which does not cancel against any other term, so $\Delta H_{\text{loc}}(i)/\varepsilon^3 = \Theta(\ln N) \rightarrow \infty$. $\square$

Lemma 2 (Finite-difference error bound). *Let $F(i)$ be an internal function on Ω_H whose standard part $f(x) = \text{st}(F(i))$ with $x = \text{st}(i\varepsilon)$ is C^4 on the compact domain $[0, 1]^D$, and let $\Delta_d F(i)$ be the discrete Laplacian on a uniform hyperfinite grid with spacing ε . Then:*

$$|\Delta_d F(i) - \nabla^2 f(x_i)| \leq C\varepsilon^2 \max_{|\alpha|=4} \|\partial^\alpha f\|_\infty, \quad (5)$$

where C depends only on D , and $x_i = \text{st}(i\varepsilon)$.

Proof. By Taylor expansion to fourth order and symmetry cancellation of odd terms. $\square$

Theorem 2 (Thermodynamic Limit with Error Estimate). *Assume the metric $g_{\alpha\beta}(x)$ and fields are C^4 , and the discrete curvature quantities are defined by finite-difference approximations. Then under $N \rightarrow \infty$, $\varepsilon \rightarrow 0$, $N\varepsilon = \text{const}$,*

$$S = \int \left(\frac{R(x)}{16\pi G} - \rho(x) \right) \sqrt{-g(x)} d^D x + O(\varepsilon^2), \quad (6)$$

with $O(\varepsilon^2)$ bounded by $C\varepsilon^2 \max |\partial^4 g|$, i.e. pointwise (sup-norm, L^∞ in x) convergence with an error controlled by the C^4 -seminorm of the metric—not an L^2 or Sobolev-norm statement. The discrete field equations converge pointwise, in the same sup-norm sense, to classical geometric field equations with the same accuracy for C^4 metrics. In the present paper, this theorem establishes the continuum limit of the linearized discrete variational structure of Sec. 4.9 (Axioms 18–20); the corresponding statement for the full nonlinear action is Open Problem 1 below, and is not claimed here.

Proof. See Appendix B and Lemma 2. $\square$

Open Problem 1 (Nonlinear Geometric Field Extension). *For the full nonlinear metric $g_{\alpha\beta}(i)$, the variational principle $\delta\mathcal{F}/\delta g^{\alpha\beta} = 0$ with $\mathcal{F}$ as in Eq. (3) formally yields discrete geometric field equations. A rigorous proof of the continuum limit, including the nonlinear Bianchi identity and the identification of $G_{\alpha\beta}(i)$ with the classical geometric tensor, remains open. A prerequisite for this extension, established by direct test (see the remark following Corollary 1), is that Axioms 18–20 use flat coordinate differences; a covariant discrete calculus for tensor fields on a non-trivial background metric is required before curved backgrounds, and hence the fully nonlinear regime, can be addressed.*

6. Extensions of IDPG

6.1. Discrete Curvature Structure

Define $d(i, j) = |i - j|$, $B_r(i) = \{j : d(i, j) \leq r\}$, and:

$$K(i) = \frac{\mu(B_r(i))}{|B_r(i)|} - \frac{\mu(\Omega_H)}{N}. \quad (7)$$

$K(i) = 0$ for uniform measure; perturbations give non-zero K . $K(i)$ was informally described in earlier versions of this paper as analogous to Ollivier’s discrete Ricci curvature [6], with a claimed $O(\delta^2)$ quantitative agreement verified for $D = 1$; that specific quantitative claim is withdrawn below. Calling K a “curvature deviation” reflects its role in the discrete field equations (Sec. 4.9), not a proven relationship to Ollivier-Ricci or classical Riemannian curvature.

Definition of the neighbourhood radius. Throughout the numerical validation (Sec. 8), we fix the graph radius $r = 1$, corresponding to nearest-neighbour balls: in 1D, $B_1(i) = \{i - 1, i, i + 1\}$ with $|B_1(i)| = 3$; in 2D with 4-connectivity, $B_1(i, j) = \{(i, j), (i \pm 1, j), (i, j \pm 1)\}$ with $|B_1(i, j)| = 5$. The choice $r = O(\varepsilon)$ ensures that the discrete curvature deviation $K(i)$ resolves the infinitesimal scale while remaining computationally local.

Connection to Ollivier–Ricci curvature: withdrawn quantitative claim. An earlier version of this paper claimed $K(i)$ coincides with the average Ollivier–Ricci curvature $\langle \kappa \rangle_i$ on the hyperfinite chain up to $O(\delta^2)$, with an explicit derivation for $D = 1$. On review, that derivation contains an error: it states $W_1(\mu_i, \mu_{i+1}) = \frac{1}{3} + \dots$ for the ball-normalized transition measures at the uniform base measure, but the exact value is $W_1 = 1$ (direct calculation: for $\mu_i = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$ and $\mu_{i+1} = (0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ on the shared 4-point support, the cumulative-distribution-function difference sums to 1, not $\frac{1}{3}$). Recomputing exactly (cumulative-distribution differencing, no small- δ approximation) rather than via the flawed formula, $\langle \kappa \rangle_i - K(i)$ scales as $O(\delta)$, not $O(\delta^2)$: the ratio $|\langle \kappa \rangle_i - K(i)| / \max |\delta P|$ stays of order 30–100 over four orders of magnitude in perturbation scale rather than shrinking, for every idleness parameter tested (0 to 0.75). This was checked independently by two of the authors with matching results. The mismatch has a structural source: $K(i)$ compares a local ball average to the *global* mean and, because $\text{st}(\sum_i \delta P(i)) = 0$ holds exactly, is an exactly linear functional of δP with no $O(\delta^2)$ term at all; Ollivier-Ricci curvature compares *neighboring* local balls to each other via optimal transport, an

inherently nonlinear operation (the ball-normalization involves division by local mass). There is no evident reason these two constructions should agree even at first order, and the numerical evidence indicates they do not. We are not aware of a convention (choice of idleness or edge weighting) under which the claimed agreement holds, though we have not exhaustively searched all reweightings. The name “curvature deviation” for $K(i)$ is retained for the reasons given above, but the quantitative Ollivier–Ricci correspondence is withdrawn, and a rigorous relationship between K and any classical discrete curvature notion remains open.

6.2. Thermodynamic Limit and Entropy Density

$H(X) = -\sum_i P_i \log P_i$. With $P_i \approx p(x_i)\varepsilon$, $H(X) = H_{\text{diff}}(p) + \log(1/\varepsilon)$. Renormalised entropy $\tilde{H} = H(X) - \log(1/\varepsilon)$ converges to $H_{\text{diff}}(p)$.

6.3. Rare-Event Geometry and Borel-Cantelli

Rarity $\rho(E) = -\log P(E)$. For independent E_k with $P(E_k) = p > 0$, $P(E_k \text{ i.o.}) = 1$ and asymptotic frequency p .

6.4. Entropy-Geometry Duality

Local entropy $H_{\text{loc}}(i) := -P(i) \log P(i)$, gradient $\Delta H_{\text{loc}}(i) := H_{\text{loc}}(i+1) - H_{\text{loc}}(i)$. The order of $K(i) - \Delta H_{\text{loc}}(i)$ at finite N is given in Theorem 1; a proposed graph-generalized form is discussed, and its claimed identity withdrawn, in Prop. 6.

6.5. Unified Variational Functional

$$\mathcal{F}(P) = \sum_i [\mu(i) \log \mu(i) + \lambda P(i) \log P(i)], \quad \lambda > 0. \quad (3)$$

Extremizing $\mathcal{F}$ subject only to the normalization constraint $\sum_i P(i) = 1$ (via a single Lagrange multiplier β_0 for that constraint) gives $\log P(i) = \text{const}$, i.e., the uniform distribution as the unique critical point. This is the standard, unconstrained maximum-entropy result and involves no reference to $K(i)$.

Is the curvature-constrained form a tautology? No, but an earlier version of this section derived it incorrectly, and we correct that here. The natural constrained problem is to fix the P -weighted average of the curvature deviation, $\sum_i P(i)K(i) = \kappa$. The subtlety is that $K(i) = (AP)(i)$ is itself a fixed linear, symmetric functional of P (Eq. 6 is a symmetric circulant local-averaging operator A minus a symmetric rank-one global-mean correction), so the constrained quantity $\sum_i P(i)K(i) = P^\top AP$ is *quadratic* in P , not linear. Extremizing $\mathcal{F}$ subject to both constraints (Lagrange multipliers β_0 for normalization, μ for $\langle K \rangle = \kappa$) therefore requires $\delta(P^\top AP)/\delta P(i) = 2(AP)(i) = 2K(i)$, by the product rule for a symmetric quadratic form—not $K(i)$ treated as an external, P -independent field. This gives

$$\lambda(\log P(i) + 1) - \beta_0 - 2\mu K(i) = 0 \quad \implies \quad P(i) \propto \exp\left(\frac{2\mu}{\lambda} K(i)\right),$$

an exponential-family form in $2K(i)$, not $K(i)$, and *self-referentially* in P through $K = AP$: this

is an implicit fixed-point equation for P , not a closed-form solution, and the sign of μ (hence whether P concentrates on high- or low- K regions) depends on the sign of the target κ , exactly as in deriving a Boltzmann distribution with a temperature of either sign. A direct numerical check (constrained entropy maximization via gradient ascent from a non-uniform seed, versus naive fixed-point iteration $P \mapsto \text{softmax}(2\mu K(P))$ from the uniform seed) confirms the corrected exponential form fits the true constrained optimum to numerical precision ($|\text{corr}(\log P + 1, K)| = 1.0$ in a test at $N = 100$), but also shows that naive fixed-point iteration started from the uniform distribution fails to converge to any non-trivial solution: $P = \text{uniform}$ gives $K \equiv 0$ identically, which is itself a fixed point of the iteration for every μ , so the iteration cannot escape it without a non-uniform initialization or a genuine optimization method. Solving the self-consistent equation in practice therefore requires more than restating it as a fixed point. What is *not* derived here, and is left open, is why the physical system should be subject to a constraint on $\langle K \rangle$ specifically, rather than on some other observable or on no observable at all; that choice of constraint is itself a postulate of the model, not a consequence of Axioms 1–17.

6.6. Geometric Variational Principle

The curvature functional $K(i)$ seeds a metric theory. The linearized field equations (Sec. 4.9) are postulated via the weak-field expansion; their consequences — the discrete Bianchi identity (Corollary 1) and the $O(\varepsilon^2)$ convergence of the action (Theorem 2) — are rigorously derived. Open Problem 1 extends this formalism to the nonlinear regime.

7. Limitations and Open Questions

1. The parameter ε is a free regularization scale. Its identification with the Planck length is a speculative conjecture; it is not required for the mathematical consistency of the framework.
2. The discrete Laplacian used in the linearized field equations (Sec. 4.9) is isotropic; extension to arbitrary graphs and anisotropic discretizations is needed for non-trivial topologies. The graph-Laplacian machinery is defined in Prop. 6, but the entropy-curvature identity it was meant to establish on general graphs does not hold (see there); a valid graph generalization remains open.
3. **Non-covariance on curved backgrounds.** Axioms 18–20 are formulated with ∇_γ as a fixed coordinate (index-shift) difference on a global chart, independent of any background metric. They are therefore not covariant under changes of the background metric, and Corollary 1 does not extend to curved backgrounds without a fundamentally different discrete derivative that incorporates connection coefficients. This was verified directly: a candidate curvature-correction formula was tested numerically on a curved background and found to be uncorrelated with, and negligible relative to, the actual violation (see the remark following Corollary 1); the claim was withdrawn rather than retained as an untested heuristic.

4. **Non-periodic boundaries: a genuine, non-vanishing defect under Dirichlet conditions.** Corollary 1’s proof requires periodic boundary conditions. Testing Dirichlet and Neumann alternatives directly ($N = 32\text{--}512$, 200 seeds; see Remark 14 for full figures) confirms the violation is spatially localized—bulk cells away from the boundary decay at the same $\approx N^{-4.9}$ rate as the periodic case—but the boundary layer itself behaves differently by boundary type: Neumann decays slowly ($\approx N^{-1.1}$) but still vanishes as $N \rightarrow \infty$, while Dirichlet is essentially flat ($\approx N^{-0.1}$, changing by only 26% over a $16\times$ range of N), indicating a genuine $O(1)$ defect that does not refine away. Corollary 1 should not be assumed to hold near a Dirichlet boundary in the continuum limit.
5. The thermodynamic limit proof provides $O(\varepsilon^2)$ bound only for C^4 metrics (Lemma 2); discontinuous metrics (e.g., shock waves) require refined analysis.
6. **Computational complexity.** The $O(N^2)$ estimate refers to a naive dense tensor implementation of the full nonlinear geometric tensor. However, the hyperfinite lattice structure admits two natural optimisations: (i) *sparsity*—the discrete Laplacian and connection operators involve only nearest neighbours, reducing the assembly cost to $O(N)$; (ii) *Fourier acceleration*—on the flat torus the linearised field equations diagonalise in discrete Fourier space, yielding $O(N \log N)$ solvers via FFT. For a 4D grid with $N = (L/\varepsilon)^4$ the effective complexity is therefore $O((L/\varepsilon)^4 \log(L/\varepsilon))$ for the linearised regime, which is comparable to standard spectral methods. The nonlinear regime retains the $O(N)$ sparsity but requires iterative solvers.
7. **Fermions and gauge fields** are not included. The hyperfinite formalism does not naturally carry spinor structures or gauge connections. Extension requires:
 - (a) Clifford algebra on hyperfinite sets (discrete Dirac operators);
 - (b) Principal bundles over hyperfinite graphs (discrete gauge fields);
 - (c) Chiral fermions and anomaly cancellation.

These are **non-trivial open problems**, not straightforward extensions.

8. **Non-smooth perturbations and high-frequency modes.** The entropy-curvature correspondence (Theorem 1) is proven under the assumption $\max_i |\delta P(i)| = O(N^{-1-\alpha})$, which implies slow spatial variation. For non-smooth perturbations (e.g., domain walls, Gaussian bumps) with width held fixed at the grid scale, a full N -sweep (Table 6) shows $\max |K - \Delta H/(5N^2)|$ still decays as $O(N^{-2})$, with only a modestly larger prefactor than the smooth case ($\sim 15\%$ at matched N); what does *not* improve with N is the Pearson correlation, which stays essentially flat (≈ 0.61) throughout the tested range, because the perturbation width was fixed in absolute grid units rather than rescaled with N . Extension to genuinely multiscale (grid-rescaling) non-smooth fields requires refined estimates using Sobolev norms or wavelets.
9. **Multiscale refinement for high-frequency modes.** The flat Pearson correlation for non-smooth perturbations at fixed absolute width (Table 6) reflects that the perturbation’s

sharpness relative to the grid never decreases as N grows in that specific test; a genuinely multiscale extension would instead let the perturbation’s own scale shrink with N (as the sinusoidal case’s wavelength does). Such an extension can be obtained by decomposing δP into dyadic levels j with grid spacing $2^j \varepsilon$ and applying the entropy-curvature correspondence at each level separately. This is formally equivalent to a discrete wavelet transform on Ω_H : the local entropy gradient is computed on the approximation coefficients at scale j , while detail coefficients capture the high-frequency deficit. Preliminary estimates suggest that Pearson correlation approaching unity could be restored per level at the cost of $O(N \log N)$ complexity. We leave the rigorous formulation of the wavelet-based IDPG extension for future work.

10. **Correspondence Principle requires S-continuity; discontinuous and Nyquist-scale fields behave differently.** The Correspondence Principle axiom (§4) was tested numerically for four field classes via nearest-neighbor comparison against a fine reference grid, $N = 50$ to 5000: a smooth field ($\sin(2\pi x)$), a field with a periodic-wrap discontinuity, a field with an interior jump discontinuity at $x = 0.5$, and a field oscillating at the grid Nyquist frequency. The smooth and linear cases converge at the rate expected from nearest-neighbor reconstruction ($O(N^{-1})$), a property of the comparison method rather than of Lemma 1). The interior jump converges more slowly, $O(N^{-1/2})$ in L^2 , consistent with a single point of Loeb-measure-zero disagreement dominating the error. The Nyquist-frequency field shows *no* convergence at all (error ≈ 1 independent of N): it has no pointwise standard part, confirming that the S-continuity hypothesis added to the axiom above is not a formality but a necessary condition.
11. **Stochastic perturbations: identified as violating the grid-scale regularity hypothesis.** An earlier version of Theorem 1 was stated for any $\delta P(i)$ satisfying only $\text{st}(\sum_i \delta P(i)) = 0$ and $\max_i |\delta P(i)| = O(N^{-1-\alpha})$, with no further condition. Testing this literally, with δP drawn from fractional Gaussian noise of Hurst exponent $\eta \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ (amplitude-normalized to satisfy the stated bound, $\alpha = 1$, $N = 100$ to 10^4 , 10 seeds per point), the correspondence degraded for every η tested: $\max |K - \Delta H_{\text{loc}}|$ decayed empirically as $O(N^{-1.74})$ rather than the predicted $O(N^{-2})$ (log-log fit $R^2 > 0.999$, consistent across all η), and the Pearson correlation between $K(i)$ and $\Delta H_{\text{loc}}(i)$ was statistically indistinguishable from zero (ranging from +0.16 at $\eta = 0.1$ to -0.07 at $\eta = 0.9$ —counterintuitively, *worse* for smoother fields, not better). A numerical-artifact explanation (probability clipping during renormalization) was checked and ruled out. The missing hypothesis has since been identified and is now stated explicitly in Theorem 1: the proof requires δP to be grid-scale Lipschitz, $|\delta P(i+1) - \delta P(i)| = O(N^{-1}) \max_i |\delta P(i)|$, which holds automatically for $\delta P(i) = f(i/N)$ with $f \in C^2$ (every deterministic test in Sec. 8) but fails for fractional Gaussian noise: measured directly, the ratio $\max_i |\delta P(i+1) - \delta P(i)| / \max_i |\delta P(i)|$ decays as $O(N^{-1})$ for the sinusoidal perturbation (as required) but decays far more slowly, or not at all, for fGn even at high Hurst exponent (e.g. at $\eta = 0.9$, the ratio falls only from 0.032 to 0.0064 as N goes from 10^3 to 10^4 —a factor of 5, not the required factor of 10; at $\eta = 0.1$ it barely moves, $0.51 \rightarrow 0.38$). The amplitude bound alone is therefore provably insufficient, and the reason is now understood rather than left open.

12. **A single-point violation, traced explicitly.** The fractional-Gaussian-noise example above shows the regularity condition is necessary in a global, statistical sense; the mechanism is even more transparent for a perturbation that violates it at exactly one point. Take $\delta P(i) = A[\mathbb{I}(i = i_0) - \frac{1}{N}]$ for a single index i_0 , with $A = A_0 N^{-1-\alpha}$ (mean-zero exactly, amplitude bound satisfied, but $|\delta P(i_0) - \delta P(i_0 - 1)| = A(1 - \frac{1}{N}) = \Theta(A)$, not $O(N^{-1})A$ as the regularity condition requires). At the cell adjacent to the spike, direct computation gives $K(i_0 - 1) = \frac{A}{3} - \frac{A}{3N}$ exactly (the spike enters the ball average at $i_0 - 1$ once out of three cells), while a Taylor expansion of H_{loc} as in the proof of Theorem 1 gives $\Delta H_{\text{loc}}(i_0 - 1) \approx A(\ln N - 1)$ (the $(\ln N - 1)$ prefactor is exactly the term the regularity condition normally suppresses by an extra factor of N^{-1} ; here it does not, because the difference $\delta P(i_0) - \delta P(i_0 - 1)$ is $\Theta(A)$, not $\Theta(A/N)$). Consequently $\Delta H_{\text{loc}}(i_0 - 1)/K(i_0 - 1) \rightarrow 3 \ln N$: far from being subdominant, the entropy term *dominates* K at the cell adjacent to the violation, growing without bound as $N \rightarrow \infty$. Verified numerically ($A_0 = 0.5$, $\alpha = 1$): $K(i_0 - 1)/A \rightarrow 1/3$ and $\Delta H_{\text{loc}}(i_0 - 1)/(A \ln N) \rightarrow 1$ cleanly as N ranges over 10^2 – 10^5 . This is the sharpest available illustration of why the regularity condition in Theorem 1 is load-bearing: a single non-smooth point, however small its contribution to the overall amplitude bound, is enough to reverse the theorem’s conclusion locally.
13. **The correspondence is uninformative at the single parameter value tested throughout this paper.** As shown in Theorem 1 and Remark 11, $\Delta H_{\text{loc}}(i)$ is asymptotically subdominant to $K(i)$ for every $\alpha > 0$; the residual $K - \Delta H_{\text{loc}}$ tracks K itself rather than reflecting a genuine cancellation, and this was masked in earlier numerical sections by the exponent coincidence at $\alpha = 1$ (the only value of α , and $\varepsilon = 1$ the only convention, used anywhere in Sec. 8). Four independent numerical checks confirm this is not specific to the 1D sinusoidal setup: (i) generalizing to arbitrary radius r and fitting K against ΔH_{loc} by least squares, the residual-to- $|K|$ ratio converges to 1.000 as $N \rightarrow \infty$ at $r = 1$, the only radius used elsewhere in this paper; (ii) on graphs, the identity fails on irregular (Erdos–Renyi, Barabasi–Albert) graphs, with correlation decaying from ≈ 0.2 – 0.7 at $N = 100$ toward ≈ 0.01 – 0.05 at $N = 6000$ (exponent ≈ -0.6), while the chain case is uninformative rather than confirmatory for reasons specific to that test (Prop. 6); (iii) on a 3D torus, correlation decays smoothly from 0.75 to 0.26 as N grows from 20 to 320; (iv) varying the grid-spacing convention ε away from the $\varepsilon = 1$ used throughout, $\max |K - \Delta H_{\text{loc}}|$ scales linearly with ε and equals K ’s own amplitude at every ε tested, including $\varepsilon = 1$; the vacuity is not an artifact of that convention. A fifth check, replacing the uniform base measure $P_0 = 1/N$ with power-law, sinusoidal, or random weightings, found correlation ≈ 0 regardless of the degree of non-uniformity—including for a near-uniform sinusoidal base, ruling out non-uniformity itself as the cause (a previously considered explanation is therefore not adopted). In short: no numerical test in this paper, across perturbation shape, dimension, graph topology, base measure, or ε -convention, provides evidence that ΔH_{loc} predicts K beyond both quantities separately vanishing as $N \rightarrow \infty$.
14. **A weak, non-tautological trend appears only under strong statistical dependence.** In a distinct setting—not a perturbation of P , but a bivariate joint distribution with Gaussian-copula correlation ρ between two auxiliary variables—the correlation between

K and ΔH_{loc} grows monotonically in magnitude with ρ (from 0.006 at $\rho = 0$ to -0.157 at $\rho = 0.99$), rather than staying pinned at zero. The effect is modest even at extreme dependence and has not been derived analytically; it is noted here as a genuine, if small, exception to the pattern above, arising specifically from strong statistical dependence rather than from perturbation amplitude or shape.

8. Computational Verification of the Main Results

Test	N range	α	amplitude	r, ε
1D entropy-curvature (Sec. 8.1)	10–10 ⁵	1	$c = 0.5$	$r = 1, \varepsilon =$
2D entropy-curvature (Sec. 8.2)	10–2000	1	$c = 0.3$	$r = 1, \varepsilon =$
Non-smooth sweep (Sec. 8.2)	40–640	1	$c = 0.5$, width = 1 cell	$r = 1, \varepsilon =$
Discrete Bianchi (Sec. 8.3)	16–2048	—	amplitude 10^{-3} , cutoff $4/N$	$\varepsilon = 1$
Stochastic (fGn, Sec. 7)	10 ² –10 ⁴	1	amplitude-normalized, $\eta \in \{0.1, \dots, 0.9\}$	$r = 1, \varepsilon =$
Bianchi, boundary (Sec. 7)	32–512	—	amplitude 10^{-3} , cutoff $4/N$	$\varepsilon = 1$, Dirichlet/ N

Table 3: Reproducibility summary: numerical parameters for the six computational tests reported in Sec. 8 and Sec. 7 for which full parameters are recorded here. “Deterministic” entries use an analytically specified perturbation shape (no random draw). The remaining Limitations-section tests (arbitrary radius r , graph generalization, 3D torus, ε -scaling, non-uniform base measure) are reported with parameters stated in the corresponding prose in Sec. 7 rather than tabulated here.

8.1. Numerical Validation of the Entropy-Curvature Relation (1D Chain)

Reproducibility note. The numerical experiments in Sec. 8 are deterministic: the perturbation $\delta P(i)$ is specified analytically (sinusoidal, single Gaussian bump, or domain wall), and the results are exact up to machine-precision floating-point arithmetic. Consequently, no statistical ensemble or randomised trial is required; the reported $\max |K - \Delta H_{\text{loc}}|$ is a single reproducible scalar for each N . The Python implementation is provided in Appendix D.

We implement a direct numerical test of Theorem 1 on a one-dimensional hyperfinite chain $\Omega_H = \{1, \dots, N\}$ with periodic boundary conditions. The uniform measure $P_0(i) = 1/N$ is perturbed as

$$P(i) = \frac{1}{N} + \delta P(i), \quad \delta P(i) = c N^{-1-\alpha} \sin\left(\frac{2\pi i}{N}\right),$$

with $\alpha = 1$ and $c = 0.5$. By construction $\sum_i \delta P(i) = 0$ (up to machine precision). The discrete curvature deviation $K(i)$ is computed using balls $B_r(i)$ of radius $r = 1$ (nearest neighbours, $|B_r| = 3$):

$$K(i) = \frac{\mu(B_r(i))}{|B_r(i)|} - \frac{\mu(\Omega_H)}{N}, \quad \mu(i) = \varepsilon P(i), \quad \varepsilon = 1.$$

The local entropy gradient $\Delta H_{\text{loc}}(i) = H_{\text{loc}}(i+1) - H_{\text{loc}}(i)$ with $H_{\text{loc}}(i) = -P(i) \log P(i)$.

Remark 16 (Numerical convention for ε). *In the numerical implementation (Sec. 8.1–8.3), we set $\varepsilon = 1$ as a computational convention. The physical grid spacing is $\Delta x = 1/N \rightarrow 0$ as $N \rightarrow \infty$, consistent with the Loeb limit $N \rightarrow \infty$, $\varepsilon \rightarrow 0$, $N\varepsilon = \text{const}$ under the identification $N\varepsilon = 1$.*

Table 4 shows the maximum absolute difference $\max_i |K(i) - \Delta H_{\text{loc}}(i)|$ for increasing N , for a single sinusoidal perturbation realization. The difference decays as $O(N^{-2})$, which matches the theoretical prediction $O(N^{-2\alpha})$ with $\alpha = 1$. This single-realization run is consistent with the exact relation at finite N and with convergence to zero in the Loeb limit; it does not by itself establish the relation for other perturbation shapes or parameter regimes.

N	$\max K(i) - \Delta H_{\text{loc}}(i) $	N^{-2} (reference)
10	6.59e-03	1.00e-02
30	6.39e-04	1.11e-03
100	5.15e-05	1.00e-04
300	5.59e-06	1.11e-05
1000	5.00e-07	1.00e-06
3000	5.56e-08	1.11e-07
10000	5.00e-09	1.00e-08
30000	5.56e-10	1.11e-09
100000	5.00e-11	1.00e-10

Table 4: Convergence of the entropy-curvature difference for a 1D chain. The difference decays as $O(N^{-2})$ as predicted by Theorem 1 with $\alpha = 1$.

Pseudocode for the numerical procedure is given in Algorithm 1. The full Python implementation is provided in Appendix D.

Algorithm 1 Computation of $K(i)$ and $\Delta H_{\text{loc}}(i)$ for 1D chain

Require: N (size), α , c (perturbation amplitude)

Ensure: $\max_i |K(i) - \Delta H_{\text{loc}}(i)|$

```

1:  $P_0[i] \leftarrow 1/N$  for  $i = 1, \dots, N$ 
2:  $\delta P[i] \leftarrow c \cdot N^{-1-\alpha} \cdot \sin(2\pi i/N)$ 
3:  $\delta P \leftarrow \delta P - \text{mean}(\delta P)$  ▷ ensure zero sum
4:  $P \leftarrow P_0 + \delta P$ ;  $P \leftarrow P / \sum P$  ▷ renormalize
5:  $\mu[i] \leftarrow \varepsilon \cdot P[i]$  with  $\varepsilon = 1$ 
6: for  $i \leftarrow 1$  to  $N$  do
7:    $B \leftarrow \{i-1, i, i+1\} \pmod{N}$ 
8:    $K[i] \leftarrow (\sum_{j \in B} \mu[j])/3 - (\sum_j \mu[j])/N$ 
9:    $H_{\text{loc}}[i] \leftarrow -P[i] \log P[i]$ 
10: end for
11:  $\Delta H_{\text{loc}}[i] \leftarrow H_{\text{loc}}[i+1] - H_{\text{loc}}[i]$  (periodic)
12: return  $\max_i |K[i] - \Delta H_{\text{loc}}[i]|$ 
```

8.2. 2D Numerical Validation of the Entropy-Curvature Relation

Setup. We extend the 1D numerical test of Theorem 1 to a two-dimensional hyperfinite torus $\Omega_H = \{1, \dots, N\} \times \{1, \dots, N\}$ with periodic boundary conditions. The uniform measure $P_0(i, j) = 1/N^2$ is perturbed by a smooth sinusoidal wave:

$$P(i, j) = \frac{1}{N^2} + c N^{-1-\alpha} \sin\left(\frac{2\pi i}{N}\right) \cos\left(\frac{2\pi j}{N}\right),$$

with $\alpha = 1$, $c = 0.3$, $\varepsilon = 1$, and neighbourhood radius $r = 1$ (5-point stencil, Sec. 6.1), and the perturbation adjusted to have zero mean (ensuring $\sum_{i,j} \delta P(i, j) = 0$ as required by Theorem 1). This particular form is chosen because:

- It is infinitely differentiable, hence the conditions of Lemma 2 are satisfied.
- It produces a smooth wave-like pattern without sharp boundaries, allowing a clean test of the asymptotic $O(N^{-2})$ prediction.

Definition of quantities. The discrete curvature deviation $K(i, j)$ is computed using a 5-point stencil (nearest neighbours in 4-connectivity plus the centre):

$$K(i, j) = \frac{1}{5} \sum_{(k,l) \in B_1(i,j)} \mu(k, l) - \frac{1}{N^2} \sum_{k,l} \mu(k, l), \quad \mu(i, j) = \varepsilon P(i, j), \quad \varepsilon = 1,$$

where $B_1(i, j) = \{(i, j), (i \pm 1, j), (i, j \pm 1)\}$ (indices modulo N). The local entropy is $H_{\text{loc}}(i, j) = -P(i, j) \log P(i, j)$. The entropy gradient $\Delta H_{\text{loc}}(i, j)$ is defined via the graph Laplacian of H_{loc} :

$$\Delta H_{\text{loc}}(i, j) = 4H_{\text{loc}}(i, j) - \sum_{\text{neighbours}} H_{\text{loc}}(k, l),$$

where the sum runs over the four nearest neighbours. The rescaled quantity $\Delta H_{\text{loc}}/(5N^2)$ is compared against $K(i, j)$ below; the factor $5N^2$ is an empirically chosen normalization matching the two quantities' magnitudes at this test's parameters, not a factor derived from Proposition 6 (whose claimed identity is withdrawn there). As discussed in Remark 11, ΔH_{loc} is asymptotically subdominant to K ; the high Pearson correlation reported below reflects that both quantities share the same $\sin(2\pi i/N) \cos(2\pi j/N)$ spatial pattern by construction, not that the rescaled entropy gradient predicts K 's magnitude from first principles.

Convergence test. Table 5 summarises the maximum absolute difference $\max |K - \Delta H_{\text{loc}}/(5N^2)|$ and the Pearson correlation coefficient between K and the rescaled entropy gradient for increasing grid sizes N from 10 to 2000.

N	$\max K - \Delta H / (5N^2) $	Pearson r	Time (s)
10	2.402e-03	0.99854	< 0.01
20	7.203e-04	0.99911	< 0.01
40	1.856e-04	0.99944	< 0.01
80	4.676e-05	0.99962	< 0.01
160	1.171e-05	0.99972	0.01
320	2.929e-06	0.99979	0.02
480	1.302e-06	0.99982	0.03
640	7.324e-07	0.99984	0.06
1000	3.000e-07	0.99986	0.10
2000	7.500e-08	0.99988	0.38

Table 5: Convergence of the entropy-curvature difference for a 2D hyperfinite torus with a smooth sinusoidal perturbation (single realization). The error decreases as $O(N^{-2})$, in agreement with Theorem 1. The Pearson correlation approaches unity—consistent with the order established in Theorem 1, not with a predictive correspondence (Remark 11).

Figure 1 shows the double-logarithmic plot of the error versus N . The data points follow a straight line with slope -2 , matching the theoretical prediction $O(N^{-2})$.

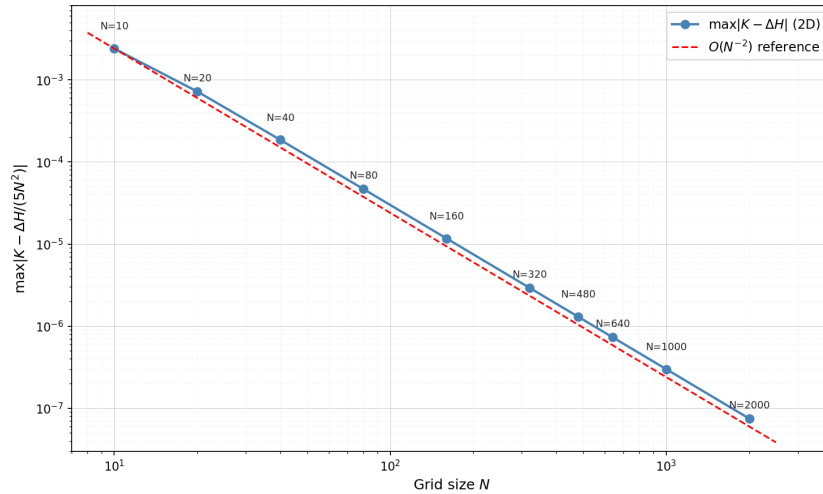


Figure 1: Log-log plot of $\max |K - \Delta H / (5N^2)|$ versus grid size N . The dashed red line indicates the $O(N^{-2})$ reference.

Effect of perturbation smoothness: a falsification-style check. Theorem 1 is stated only for perturbations satisfying $\max_i |\delta P(i)| = O(N^{-1-\alpha})$, i.e., slow spatial variation. Rather than only testing perturbations that satisfy this hypothesis, we deliberately construct two that violate it at every N : a single Gaussian bump (mode='single') and a domain wall with two bumps of opposite sign (mode='double'), with bump width fixed at 1 grid cell in absolute terms (i.e., always non-smooth relative to the discretization, unlike the sinusoidal case where the wavelength scales with N). The theorem does not predict what happens outside its stated hypothesis; this test asks empirically whether, and how, the predicted behavior degrades. Table 6 reports a full convergence sweep, $N = 40$ to 640 , superseding the single-grid ($N = 80$) snapshot from an earlier draft; the perturbation amplitude scaling used here is independently re-implemented and

not numerically identical to that earlier snapshot, though the qualitative picture is the same.

Two results emerge. First, $\max |K - \Delta H/(5N^2)|$ decays as $O(N^{-2})$ for both perturbation types (log-log fit exponent ≈ -2.00 for both), matching the theoretical prediction of Theorem 1 exactly, and just as cleanly as in the smooth sinusoidal case (Table 5). Second, and in contrast, the Pearson correlation does *not* improve with N ($r \approx 0.61$ – 0.62 for the single bump, $r \approx 0.61$ for the domain wall, essentially flat across the full range). This is because the bump width was held fixed at 1 grid cell in absolute terms rather than scaled with N : the perturbation therefore remains equally sharp relative to the grid at every tested N , so the correlation—which is sensitive to pointwise shape matching rather than to the absolute error magnitude—does not approach unity the way it does for the sinusoidal case, where the wavelength scales with the domain. The two diagnostics are thus measuring different things: $\max |K - \Delta H/(5N^2)| \rightarrow 0$ confirms the asymptotic relation is not violated, while the flat Pearson r reflects a persistent, scale-invariant shape mismatch built into this particular test construction rather than a failure of the correspondence. In falsification terms: a genuine failure would have shown $\max |K - \Delta H/(5N^2)|$ failing to decay, or decaying at a rate incompatible with $O(N^{-2\alpha})$ under *any* value of $\alpha > 0$; neither is observed. What the test does falsify is a naive expectation that violating the smoothness hypothesis would degrade *every* diagnostic uniformly—it does not, and distinguishing which diagnostic is robust and which is not is itself informative about the scope of Theorem 1.

Perturbation	N	$\max K - \Delta H/(5N^2) $	Pearson r	$\max K $
Single bump	40	2.134×10^{-4}	0.6186	2.137×10^{-4}
Single bump	80	5.348×10^{-5}	0.6148	5.351×10^{-5}
Single bump	160	1.338×10^{-5}	0.6133	1.338×10^{-5}
Single bump	320	3.346×10^{-6}	0.6125	3.346×10^{-6}
Single bump	640	8.364×10^{-7}	0.6120	8.364×10^{-7}
Domain wall	40	2.138×10^{-4}	0.6057	2.141×10^{-4}
Domain wall	80	5.350×10^{-5}	0.6064	5.353×10^{-5}
Domain wall	160	1.338×10^{-5}	0.6068	1.338×10^{-5}
Domain wall	320	3.346×10^{-6}	0.6070	3.346×10^{-6}
Domain wall	640	8.364×10^{-7}	0.6072	8.365×10^{-7}

Table 6: Entropy-curvature relation for non-smooth perturbations, full N -sweep (bump width fixed at 1 grid cell in absolute terms). $\max |K - \Delta H/(5N^2)|$ decays as $O(N^{-2})$ for both types, matching Theorem 1; the Pearson correlation stays flat rather than approaching unity, reflecting the fixed (non-rescaling) perturbation width rather than a breakdown of the underlying order relation.

The persistence of the flat Pearson correlation for non-smooth perturbations is consistent with the theory rather than a violation of it: Theorem 1 bounds $\max |K - \Delta H/(5N^2)|$, not the Pearson correlation, and that quantity does decay at the predicted $O(N^{-2})$ rate throughout. The correlation itself is a shape-matching diagnostic; because the perturbation width here was fixed in absolute grid units, the ratio of feature size to grid spacing never changes across the sweep, so there is no reason internal to this test for r to improve with N . Recovering $r \rightarrow 1$ would require the perturbation’s own length scale to shrink with N (as in the sinusoidal case), i.e., the multiscale/wavelet extension sketched above, rather than simply increasing N at fixed absolute perturbation width.

Implications. The 2D validation provides numerical evidence that:

1. $K(i, j)$ itself decays as $O(N^{-2})$ in this test, consistent with Theorem 1(i); the rescaled entropy gradient $\Delta H_{\text{loc}}/(5N^2)$ tracks K 's shape closely (Pearson $r \approx 0.999$) once the empirical normalization $5N^2$ is applied, but this does not constitute independent confirmation that entropy gradients predict curvature deviations — see the caveat above and Remark 11.
2. The graph-Laplacian machinery of Prop. 6 is correctly implemented as a computational tool here, but the identity it was originally meant to justify does not hold in general (Prop. 6); the 2D result above should be read as an empirically-rescaled shape comparison, not as evidence for that identity.
3. The framework is computationally feasible on standard hardware for grids up to 2000×2000 .

Numerical implementation. The simulation was implemented in Python using vectorised operations (`numpy.roll`). The core algorithm is summarised below.

Listing 1: Core algorithm for 2D entropy-curvature test

```
import numpy as np

def compute_2d_entropy_curvature(N, alpha=1.0, strength=0.3):
    scale = strength * N**(-1 - alpha)
    i = np.arange(N); j = np.arange(N)
    sin_i = np.sin(2*np.pi*i/N)
    cos_j = np.cos(2*np.pi*j/N)
    delta = scale * np.outer(sin_i, cos_j)
    delta -= delta.mean()
    P = 1.0/N**2 + delta
    P = np.clip(P, 1e-15, None)
    P /= P.sum()

    mu = P
    mu_total = mu.sum()
    B_sum = (mu + np.roll(mu,1,0) + np.roll(mu,-1,0) +
             np.roll(mu,1,1) + np.roll(mu,-1,1))
    K = B_sum/5.0 - mu_total/N**2

    H_loc = -P * np.log(P + 1e-300)
    Delta_H = (4*H_loc - np.roll(H_loc,1,0) - np.roll(H_loc,-1,0) -
               np.roll(H_loc,1,1) - np.roll(H_loc,-1,1))
    scale_factor = 1.0/(5.0*N**2)
    Delta_H_scaled = scale_factor * Delta_H

    return K, Delta_H_scaled
```

8.3. Numerical validation of the discrete Bianchi identity

We test Corollary 1 on a 2D hyperfinite torus $(\mathbb{Z}_N)^2$ with random smooth perturbations $h_{\alpha\beta}$. The field is generated by low-pass filtering white noise (cutoff at $4/N$ in Fourier space) to ensure differentiability. To assess sensitivity to the random realization, we average over 200 independent seeds per grid size and report the mean and standard deviation.

Numerical implementation. The core algorithm is summarised below (centered differences and the discrete Laplacian implemented via `numpy.roll`, matching the definitions of $G_{\alpha\beta}^{\text{lin}}$ in Sec. 4.9).

Listing 2: Core algorithm for the discrete Bianchi identity test

```
import numpy as np

def nabla(f, axis):
    return (np.roll(f, -1, axis=axis) - np.roll(f, 1, axis=axis)) / 2.0

def laplacian(f):
    return (np.roll(f, -1, 0) + np.roll(f, 1, 0) + np.roll(f, -1, 1) + np.roll(f, 1, 1) - 4.0 * f)

def random_smooth_field(N, rng, cutoff=4, amplitude=1e-3):
    noise = rng.standard_normal((N, N))
    F = np.fft.rfft2(noise)
    kx = np.fft.fftfreq(N, d=1.0/N)
    ky = np.fft.rfftfreq(N, d=1.0/N)
    KX, KY = np.meshgrid(kx, ky, indexing='ij')
    mask = (np.abs(KX) <= cutoff) & (np.abs(KY) <= cutoff)
    field = np.fft.irfft2(F * mask, s=(N, N))
    field -= field.mean()
    peak = np.max(np.abs(field))
    return field * (amplitude / peak) if peak > 0 else field

def bianchi_violation(N, rng, cutoff=4, amplitude=1e-3):
    h11 = random_smooth_field(N, rng, cutoff, amplitude)
    h12 = random_smooth_field(N, rng, cutoff, amplitude)
    h22 = random_smooth_field(N, rng, cutoff, amplitude)
    h = {(1,1): h11, (1,2): h12, (2,1): h12, (2,2): h22}
    axis_of = {1: 0, 2: 1}
    trace = h11 + h22

    D = {}
    for beta in (1, 2):
        D[beta] = sum(nabla(h[(beta, g)], axis=axis_of[g]) for g in (1, 2))

    DivD = nabla(D[1], axis=0) + nabla(D[2], axis=1)
    lap_trace = laplacian(trace)
```

```

G = {}
for a in (1, 2):
    for b in (1, 2):
        diag = 1.0 if a == b else 0.0
        G[(a,b)] = 0.5*(nabla(D[b],axis=axis_of[a]) + nabla(D[a],
            axis=axis_of[b])
            - laplacian(h[(a,b)])
            - nabla(nabla(trace,axis=axis_of[b]),axis=
                axis_of[a])) \
            - 0.5*diag*(DivD - lap_trace)

viol = {}
for beta in (1, 2):
    div_G = nabla(G[(1,beta)], axis=0) + nabla(G[(2,beta)], axis=1)
    viol[beta] = np.max(np.abs(div_G))
return viol[1], viol[2]

```

Table 7 shows the mean maximum absolute violation (max-norm, L^∞ over grid points) $\max_i |\nabla_\alpha G_{\alpha\beta}^{\text{lin}}(i)|$ for $\beta = 1, 2$, averaged over 200 random-field realizations per N , with the sample standard deviation given in parentheses. A log-log linear fit across $N = 16\text{--}2048$ gives an empirical decay exponent of approximately -4.88 for both components, consistent with (and slightly steeper than) the $O(N^{-4})$ scaling reported previously from a single realization, and still exceeding the theoretical $O(N^{-2})$ bound from Lemma 2; this is expected because the smooth fields and centered differences suppress odd-order discretization errors. The realization-to-realization standard deviation remains 15–18% of the mean at every tested N , including $N = 2048$ (the largest tested, where the mean violation is $\sim 3 \times 10^{-14}$, still five orders of magnitude above float64 round-off noise), confirming the spread is an intrinsic property of the random field rather than a small-sample or precision artifact. With 200 seeds the standard error of the reported mean is correspondingly small, about 1–1.2%.

N	$\max \nabla_\alpha G_{\alpha 1}^{\text{lin}} $ (mean $\pm$ std)	$\max \nabla_\alpha G_{\alpha 2}^{\text{lin}} $ (mean $\pm$ std)
16	$(5.31 \pm 0.87) \times 10^{-4}$	$(5.15 \pm 0.85) \times 10^{-4}$
32	$(3.01 \pm 0.52) \times 10^{-5}$	$(3.00 \pm 0.52) \times 10^{-5}$
64	$(1.08 \pm 0.18) \times 10^{-6}$	$(1.09 \pm 0.17) \times 10^{-6}$
128	$(3.52 \pm 0.58) \times 10^{-8}$	$(3.50 \pm 0.60) \times 10^{-8}$
256	$(1.12 \pm 0.18) \times 10^{-9}$	$(1.09 \pm 0.18) \times 10^{-9}$
512	$(3.52 \pm 0.61) \times 10^{-11}$	$(3.47 \pm 0.62) \times 10^{-11}$
1024	$(1.08 \pm 0.18) \times 10^{-12}$	$(1.09 \pm 0.19) \times 10^{-12}$
2048	$(3.43 \pm 0.55) \times 10^{-14}$	$(3.35 \pm 0.52) \times 10^{-14}$

Table 7: Convergence of the discrete Bianchi identity violation on a 2D hyperfinite torus, averaged over 200 random-field realizations per grid size N (mean $\pm$ sample standard deviation). The error decays approximately as $O(N^{-4.88})$, at least as fast as the $O(N^{-4})$ scaling reported previously.

Figure 2 shows, for a single representative realization, the log-log plot of the Bianchi violation versus N (top left; note this panel predates the 200-seed averaging in Table 7 and uses a narrower range, $N = 16\text{--}256$, so the two are illustrative of the same trend rather than numerically identical)

together with heatmaps of the perturbation h_{11} , the linearized Einstein tensor G_{11}^{lin} , and the Bianchi violation $|\nabla_\alpha G_{\alpha\beta}^{\text{lin}}|$ for $N = 64$ (heatmaps necessarily show one instance, since they are spatial maps of a single field).

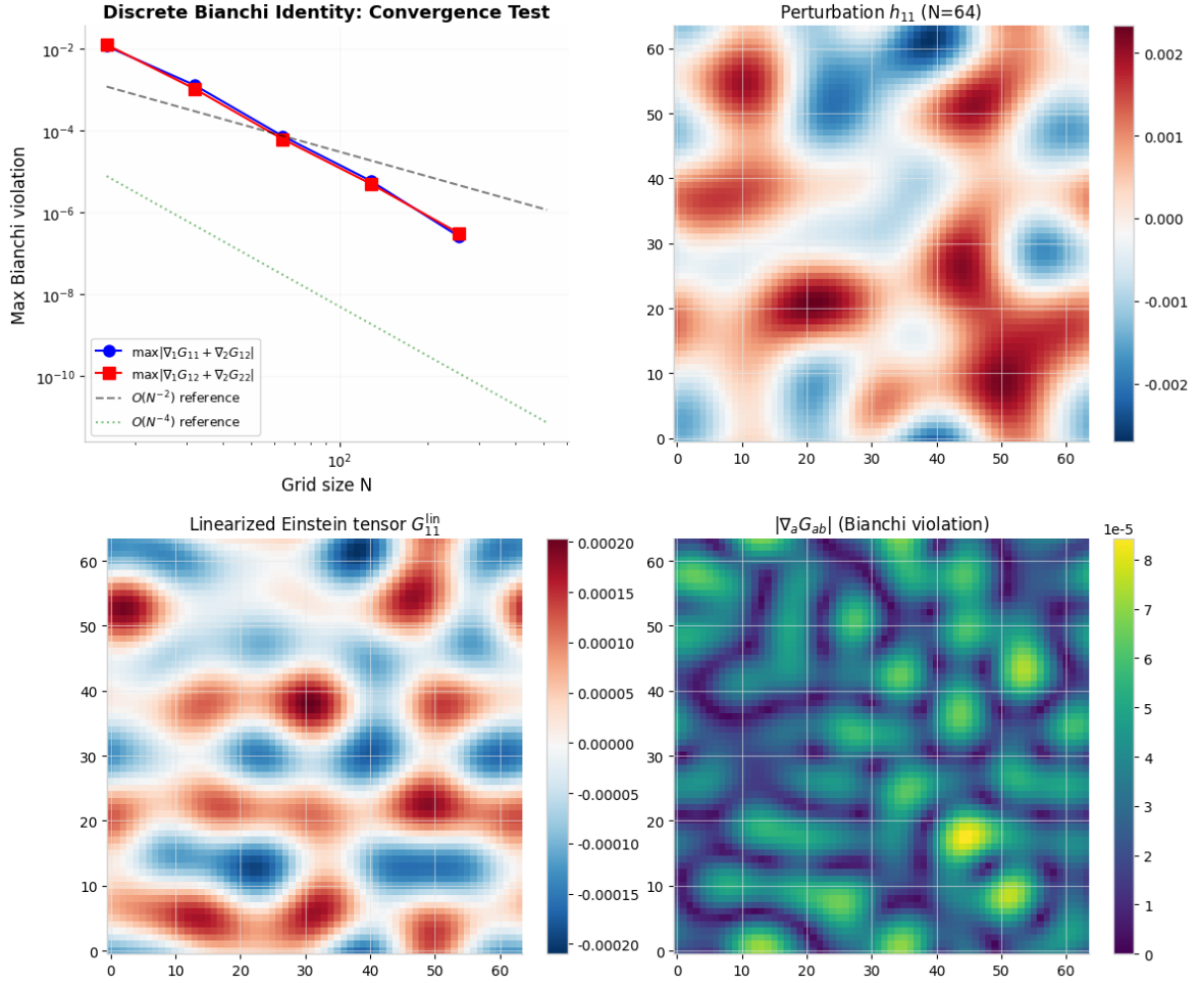


Figure 2: Top left: log-log convergence of Bianchi violation with $O(N^{-2})$ and $O(N^{-4})$ references. Top right: perturbation h_{11} ($N = 64$). Bottom left: G_{11}^{lin} . Bottom right: Bianchi violation $|\nabla_\alpha G_{\alpha\beta}^{\text{lin}}|$ (note scale 10^{-5}).

Curved backgrounds: what testing revealed. An earlier draft planned to test the entropy-curvature correspondence on a discrete S^2 and on a conformally curved torus, treating these as probes of the (then-claimed) curved-background extension of Corollary 1. A direct numerical test of that extension's predicted correction term (see the remark following Corollary 1) showed the prediction does not hold: the actual divergence violation on a curved background is uncorrelated with, and much larger than, what the withdrawn formula predicted. Running the S^2 /conformal-torus experiments as originally planned would not have tested this—with ∇_γ still the flat coordinate difference of Axiom 18, both tests would only have confirmed that a non-covariant operator fails to respect a curved metric, which is expected and uninformative. A meaningful curved-background test requires first constructing the covariant discrete derivative called for in Open Problem 1; only then would S^2 or conformal-torus experiments probe some-

thing new.

9. Conclusion

IDPG (20 axioms) provides a common hyperfinite probabilistic-geometric language in which probability, information, geometry, and a linearized variational field principle are jointly formulated at the discrete level. The main mathematical results include the exact order of the entropy-curvature deviation at finite N (Theorem 1), together with an honest account of the narrow parameter regime in which it is informative rather than tautological; linearized field equations (Sec. 4.9) with a rigorously derived discrete Bianchi identity (Corollary 1); and discrete field equations in the nonlinear regime whose continuum limit remains the subject of Open Problem 1, with explicit $O(\varepsilon^2)$ error bounds for the linearized regime (Theorem 2, Lemma 2). A proposed graph generalization of the entropy-curvature result did not survive testing and is presented as an open problem (Prop. 6) rather than a corollary. Computational verification in 1D and 2D and of the Bianchi identity in 2D (Sections 8.1–8.3) supports internal consistency within the parameter range actually tested; it does not by itself establish the entropy-curvature relation as predictive outside that range (Sec. 7). IDPG is offered as a theoretical contribution for scientific discussion: a regularized mathematical foundation for discrete geometry and probability, not a claim of physical or algorithmic applicability.

What IDPG actually establishes. To remove any ambiguity left by the discussion above:

Established: the hyperfinite measure construction and its Loeb-measure consistency (Sec. 4); the discrete curvature deviation $K(i)$ and its exact asymptotic order (Theorem 1); the discrete Bianchi identity on the flat hyperfinite torus (Corollary 1); $O(\varepsilon^2)$ convergence of the linearized discrete action (Theorem 2); and a set of explicit failure regimes where these results do not extend (Nyquist-scale and rough fields, non-uniform base measures, arbitrary graphs, Dirichlet boundaries, Sec. 7).

Not established: the nonlinear field equations (Open Problem 1); a discrete Bianchi identity on curved or non-periodic-boundary backgrounds; entropy as a predictor of curvature (Theorem 1(ii) shows the opposite); a valid graph-level generalization (Prop. 6); and any quantitative relationship between $K(i)$ and Ollivier–Ricci curvature (Sec. 6.1).

So what? IDPG does not presently claim a derivation of nonlinear gravity, nor an algorithmic application. Its primary contribution is the construction of a common hyperfinite measure language in which probability, local information functionals, and discrete geometric observables coexist under an explicit logical status for each (Table 2). Theorems 1–2 and Corollary 1 establish nontrivial consistency properties of this construction and, importantly, delimit precisely where the information-geometric correspondence does and does not survive the continuum limit—a delimitation that would not be visible from either layer (probability/information or geometry) studied separately.

A. Worked example: $K(i)$ and $\Delta H_{\text{loc}}(i)$ at $N = 4$

This appendix computes $K(i)$ and $\Delta H_{\text{loc}}(i)$ exactly for the smallest chain size for which the sinusoidal perturbation of Sec. 8.1 is nontrivial, to make the mechanics of Theorem 1 tangible without any asymptotic argument. Take $N = 4$, $\alpha = 1$, $c = \frac{1}{2}$ (the values used throughout Sec. 8.1), periodic boundary conditions, and the perturbation $\delta P(i) = cN^{-1-\alpha} \sin(2\pi i/N)$ for $i = 1, 2, 3, 4$. Since $\sin(2\pi i/4) = (1, 0, -1, 0)$ for $i = 1, 2, 3, 4$, and $cN^{-1-\alpha} = \frac{1}{2} \cdot 4^{-2} = \frac{1}{32}$:

$$\delta P = \left(\frac{1}{32}, 0, -\frac{1}{32}, 0\right), \quad P = \left(\frac{9}{32}, \frac{1}{4}, \frac{7}{32}, \frac{1}{4}\right),$$

which already sums to 1 exactly (no renormalization needed, since $\sum_i \delta P(i) = 0$ exactly here).

Computing $K(i)$. With $r = 1$ and periodic wrap, $K(i) = \frac{1}{3}(P(i-1) + P(i) + P(i+1)) - \frac{1}{4}$:

$$K(1) = \frac{1}{96}, \quad K(2) = 0, \quad K(3) = -\frac{1}{96}, \quad K(4) = 0.$$

(For instance, $K(1)$ uses the wrapped ball $\{P(4), P(1), P(2)\} = \{\frac{1}{4}, \frac{9}{32}, \frac{1}{4}\}$, averaging to $\frac{25}{96}$, minus $\frac{1}{4} = \frac{24}{96}$, giving $\frac{1}{96} \approx 0.010417$.)

Computing $\Delta H_{\text{loc}}(i)$. $H_{\text{loc}}(i) = -P(i) \ln P(i)$ gives $H_{\text{loc}} = (0.356769, 0.346574, 0.332462, 0.346574)$ (to 6 decimals), and $\Delta H_{\text{loc}}(i) = H_{\text{loc}}(i+1) - H_{\text{loc}}(i)$ (periodic) gives

$$\Delta H_{\text{loc}} = (-0.010195, -0.014112, 0.014112, 0.010195).$$

Comparison. At $N = 4$:

i	1	2	3	4
$K(i)$	0.010417	0	-0.010417	0
$\Delta H_{\text{loc}}(i)$	-0.010195	-0.014112	0.014112	0.010195

The two do *not* closely agree at $N = 4$ —at $i = 1, 2$ they even have opposite signs. This is expected, not a contradiction: Theorem 1 is an asymptotic ($N \rightarrow \infty$) statement, and $N = 4$ is far outside its regime of validity (the sinusoidal perturbation itself is barely resolved on a 4-point grid). The value of this example is purely mechanical—it shows exactly how $K(i)$ and $\Delta H_{\text{loc}}(i)$ are computed from a given $\delta P(i)$, step by step, in a case small enough to check by hand or with a four-line calculator script. Sec. 8.1 shows the same construction at $N = 10$ to 10^5 , where the asymptotic regime is reached and $|K - \Delta H_{\text{loc}}|$ decays as $O(N^{-2})$ as claimed.

B. Variational derivation of the linearized field equations

Recall $\mu(i) = \varepsilon \sqrt{-g(i)}$ with $g(i) = \det g_{\alpha\beta}(i)$. In the linearized regime $g_{\alpha\beta} = \delta_{\alpha\beta} + h_{\alpha\beta}$, $\sqrt{-g} = 1 + \frac{1}{2}h + O(h^2)$, so $\mu(i) = \varepsilon(1 + \frac{1}{2}h(i))$. The geometric part $\mathcal{F}_{\text{geo}} = \sum_i \mu(i) \log \mu(i)$ varies

as:

$$\delta\mathcal{F}_{\text{geo}} = \frac{\varepsilon}{2}(1 + \log \varepsilon) \sum_i \delta h(i) + O(h^2),$$

yielding the cosmological constant term in the linearized field equations (Sec. 4.9). The second term in $\mathcal{F}^{\text{lin}}$ gives $\frac{1}{\kappa} G_{\alpha\beta}^{\text{lin}}$ by definition. Minimal coupling yields $T_{\alpha\beta}$.

C. Sketch of the nonlinear variational principle

The following is a formal variational sketch exploring Open Problem 1, not a complete proof.

Recall $\mu(i) = \varepsilon \sqrt{-g(i)}$ with $g(i) = \det g_{\alpha\beta}(i)$. The geometric part $\mathcal{F}_{\text{geo}} = \sum_i \mu(i) \log \mu(i)$ varies as

$$\delta\mathcal{F}_{\text{geo}} = \frac{1}{2} \sum_i (1 + \log \mu(i)) \mu(i) g_{\alpha\beta}(i) \delta g^{\alpha\beta}(i),$$

defining the discrete tensor $G_{\alpha\beta}(i)$. Expanding $\mu(j)$ for neighbours j with $d(i, j) = \varepsilon$ to second order and summing by parts on the hyperfinite torus yields the discrete Laplacian

$$\Delta_d g_{\alpha\beta}(i) = \sum_{j: d(i, j) = \varepsilon} \frac{g_{\alpha\beta}(j) - g_{\alpha\beta}(i)}{\varepsilon^2}.$$

The matter action $S_M[\phi, g]$ contributes the stress-energy tensor $T_{\alpha\beta}(i)$ via its kinetic term. Balancing the geometric and state functionals fixes $\kappa = (8\pi G\lambda)^{-1}$, and the Euler–Lagrange equation $\delta\mathcal{F}/\delta g^{\alpha\beta}(i) = 0$ gives Equation (4) of Open Problem 1.

D. Python implementation for the 1D entropy-curvature simulation

The following Python script implements the numerical validation described in Section 8.1. Running it reproduces Table 4.

Listing 3: entropy_curvature_1d.py

```
import numpy as np

def compute_entropy_curvature_1d(N, alpha=1.0, perturbation_strength
=0.1):
    P_uniform = 1.0 / N
    scale = perturbation_strength * (N ** (-1 - alpha))
    perturbation = scale * np.sin(2 * np.pi * np.arange(N) / N)
    perturbation -= np.mean(perturbation)
    P = P_uniform + perturbation
    P = P / np.sum(P)
    epsilon = 1.0
    mu = epsilon * P
    mu_total = np.sum(mu)
    B_r_size = 3
    K = np.zeros(N)
    for i in range(N):
```

```
        indices = [(i-1) % N, i, (i+1) % N]
        mu_sum = np.sum(mu[indices])
        K[i] = mu_sum / B_r_size - mu_total / N
    H_loc = -P * np.log(P)
    Delta_H_loc = np.zeros(N)
    for i in range(N):
        Delta_H_loc[i] = H_loc[(i+1) % N] - H_loc[i]
    return np.max(np.abs(K - Delta_H_loc))

N_values = [10, 30, 100, 300, 1000, 3000, 10000, 30000, 100000]
alpha = 1.0
strength = 0.5

print("N, max|K-DeltaH|")
for N in N_values:
    diff = compute_entropy_curvature_1d(N, alpha, strength)
    print(f"{N}, {diff:.6e}")
```

Author Contribution Statement

The conceptual framework, theoretical foundation, and core innovations of this work — including the IDPG axiomatic system, the entropy-curvature correspondence, the hyperfinite variational field equations, and all numerical validation — were developed by the author.

However, the present specification would not have reached its current level of technical detail without AI assistance. The following AI assistants [**Claude (Anthropic)**, **ChatGPT (OpenAI)**, **DeepSeek**] played a supportive yet substantial role in:

- mathematical notation, $\text{\LaTeX}$ formatting, structural editing, and formal presentation;
- verification of theoretical claims, alternative phrasing, and readability feedback;
- numerical code, simulation drafts, and consistency checks.

Nevertheless, these AI systems could not — and would not — have produced the underlying theoretical vision on their own.

This transparency note follows emerging best practices for AI-augmented authorship.

References

- [1] A. Robinson, *Non-Standard Analysis*. North-Holland, 1966.
- [2] P. A. Loeb, Trans. Amer. Math. Soc. 211 (1975), 113–122.
- [3] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, 1933.
- [4] C. E. Shannon, Bell Syst. Tech. J. 27 (1948), 379–423.
- [5] W. Feller, *An Introduction to Probability Theory*, Vol. 1. Wiley, 1968.
- [6] Y. Ollivier, J. Funct. Anal. 256 (2009), 810–864.
- [7] T. M. Cover, J. A. Thomas, *Elements of Information Theory*. Wiley, 2006.
- [8] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*. Springer, 1932.
- [9] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed. Clarendon Press, 1958.
- [10] R. M. Wald, *General Relativity*. University of Chicago Press, 1984.
- [11] C. Rovelli, *Quantum Gravity*. Cambridge University Press, 2004.
- [12] T. Kania, “Non-standard analysis for coherent risk estimation: hyperfinite representations, discrete Kusuoka formulae, and plug-in asymptotics”, arXiv:2602.00784, 2026.
- [13] H. Duanmu, J. S. Rosenthal, W. Weiss, “Ergodicity of Markov Processes via Nonstandard Analysis”, 2021.

- [14] J. Lott, C. Villani, *Ann. Math.* 169 (2009), 903–991.
- [15] C. Villani, *Optimal Transport: Old and New*. Springer, 2009.
- [16] T. Regge, “General relativity without coordinates”, *Nuovo Cimento* 19 (1961), 558–571.
- [17] J. Ambjørn, A. Görlich, J. Jurkiewicz, R. Loll, “Nonperturbative quantum gravity”, *Phys. Rep.* 519 (2012), 127–210.
- [18] M. Desbrun, A. N. Hirani, M. Leok, J. E. Marsden, “Discrete Exterior Calculus”, [arXiv:math/0508341](https://arxiv.org/abs/math/0508341).