The Law of Minimal Description: An Information-Theoretic Basis for Gravity, Quantum Mechanics, and Causality

Mats Helander and Jeeves
Independent Research
(Dated: October 20, 2025)

Abstract

We propose that a single informational principle underlies the laws of physics: the universe evolves toward states of shorter description. Reformulating the Second Law of Thermodynamics in terms of description length Φ, we state the *Law of Minimal Description*: $\Delta\Phi \leq 0$. We show that gravity emerges from spatial compression gradients; quantum mechanics arises from compression across possible futures; General Relativity is curvature in description space; and causality and physical law are temporal compression. Newtonian gravity follows from isotropy, locality, and informational flux conservation. Einstein's equations follow from second-order variations of Φ over metrics. The Born rule emerges from a compression-weighted selection rule. Monte Carlo simulations generate gravitational clustering using only compression bias, showing no force postulate is required. Physics becomes the science of evolving efficient descriptions.

1. DEFINITIONS AND ASSUMPTIONS

A. Description Length Φ

Φ is the minimal description length of the universe (in bits) under an optimal prefix-free encoding. It corresponds to Kolmogorov complexity K up to an additive constant:

$$\Phi = K(\text{universe}). \tag{1.1}$$

Φ is dimensionless. Physical units appear only when mapping $\nabla\Phi$ to forces or energies.

B. Compression

A transformation is compressive if it reduces total description length:

$$\Phi(\text{state}_{t+1}) \leq \Phi(\text{state}_t). \tag{1.2}$$

C. Description Gradient

Let x be a configuration in an abstract configuration space. Evolution follows the steepest descent in Φ:

$$\frac{dx}{dt} \propto -\nabla\Phi(x). \tag{1.3}$$

We call $-\nabla\Phi$ the *description force*.

Assumptions

1. **Informational Universality.** Physical systems are finitely describable.

2. **Equivalence of Entropy and Description.** Physical entropy S corresponds to Φ via $\Phi = S/(k \ln 2) + C$.

3. **Local Computation.** Changes in Φ propagate locally.

4. **Isotropy and Homogeneity.** No preferred spatial direction or location.

5. **No Physical Postulates.** Fields, forces, and quantum axioms are not assumed.

2. INTRODUCTION

The Second Law of Thermodynamics is commonly stated as monotonic entropy increase. In foundational terms, entropy measures information. Shannon formalized missing information; Kolmogorov and Chaitin extended it to individual objects; Rissanen related it to description length via MDL. Hence the Second Law can be restated as

$$\Delta\Phi \leq 0, \tag{2.1}$$

meaning the universe evolves toward simpler descriptions. Contrary to popular belief, entropy does not favor disorder; it favors efficient representation. Order persists when it compresses. We propose: *compression is the fundamental driver of physical evolution.*

3. ENTROPY AS DESCRIPTION LENGTH

Shannon entropy measures expected information:

$$H(X) = -\sum_x p(x) \log p(x). \tag{3.1}$$

Kolmogorov complexity measures irreducible information:

$$K(x) = \min_{p:U(p)=x} |p|. \tag{3.2}$$

Rissanen's MDL principle selects the model minimizing total description length:

$$L(x, M) = L(M) + L(x \mid M). \tag{3.3}$$

All agree: entropy counts bits. The Second Law is fundamentally a law of description.

4. THE LAW OF MINIMAL DESCRIPTION AS A DYNAMICAL PRINCIPLE

To turn $\Delta\Phi \leq 0$ into dynamics, we treat Φ as a scalar potential over configuration space. Let x represent a physical configuration. Evolution follows the steepest descent:

$$\frac{dx}{dt} \propto -\nabla\Phi(x). \tag{4.1}$$

Define the description force

$$F(x) = -\nabla\Phi(x). \tag{4.2}$$

Attraction arises where Φ decreases with proximity; repulsion where it increases.

5. SPATIAL COMPRESSION AND THE ORIGIN OF GRAVITY

Gravity emerges when Φ is applied to spatial redundancy. Distant objects require independent specification; proximity allows joint encoding, reducing Φ. Hence for separation r,

$$\frac{d\Phi}{dr} < 0. \tag{5.1}$$

A. Description Density and Mass Density

We define a description density $\rho(x)$ representing irreducible information at location x. Because physical mass stores microstate information, we identify

$$\rho(x) = \alpha\, \rho_m(x), \tag{5.2}$$

with ρ_m the mass density and α a constant converting mass to description weight. This aligns with thermodynamic and emergent gravity perspectives (Jacobson, 1995; Verlinde, 2011). Thus gravitational mass measures informational content.

B. Isotropy Implies Central Attraction

By isotropy, description depends only on radial distance $r = \|x - x'\|$ and

$$\nabla\Phi = \frac{d\Phi}{dr}\,\hat{r}, \tag{5.3}$$

yielding a central potential without postulating a force.

6. NEWTON'S LAW AS A COROLLARY OF DESCRIPTION MINIMIZATION

More massive objects contain more irreducible structure and contribute more to Φ, hence exert stronger description gradients. The pairwise interaction obeys

$$F(r) \propto \frac{m_1 m_2}{r^2}, \tag{6.1}$$

with the inverse-square dependence fixed by isotropy and informational flux conservation (Appendix A). Introducing the constant G,

$$F(r) = -G\,\frac{m_1 m_2}{r^2}. \tag{6.2}$$

Attraction reflects the informational inequality $\Phi(A{+}B) < \Phi(A) + \Phi(B)$. Repulsion arises where proximity increases description cost (e.g., Pauli exclusion, Coulomb).

7. RELATIVITY FROM DESCRIPTION GEOMETRY

We generalize Φ to spacetime histories. For a worldline γ,

$$\Phi[\gamma] = \text{description length of } \gamma, \qquad \delta\Phi[\gamma] = 0, \tag{7.1}$$

so worldlines minimize description length. The first variation yields geodesics in a metric $g_{\mu\nu}$:

$$d\Phi^2 = g_{\mu\nu}\, dx^\mu dx^\nu. \tag{7.2}$$

The second variation defines an informational curvature tensor $K_{\mu\nu}$ proportional (by Lovelock's theorem) to the Einstein tensor $G_{\mu\nu}$:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{7.3}$$

Relativity is geometry of description gradients.

8. SIMULATION EVIDENCE

We test whether gravitational behavior emerges from compression alone via Monte Carlo simulations minimizing Φ over discrete spatial configurations.

A. Method

We simulate N point masses in a periodic box. Φ is approximated by a Minimum Spanning Tree (MST) encoding cost; the MST is computed via Prim's algorithm [22]. The system evolves with a Metropolis rule

$$P(s \to s') = \min\bigl(1, e^{-\beta\,\Delta\Phi}\bigr), \tag{8.1}$$

where β controls compression strength. For many-body runs we track the mean pairwise distance

$$\bar{r}(t) = \frac{2}{N(N-1)} \sum_{i<j} \|x_i(t) - x_j(t)\|, \tag{8.2}$$

and for a two-body system we monitor the inter-particle separation $r(t)$.

B. Results

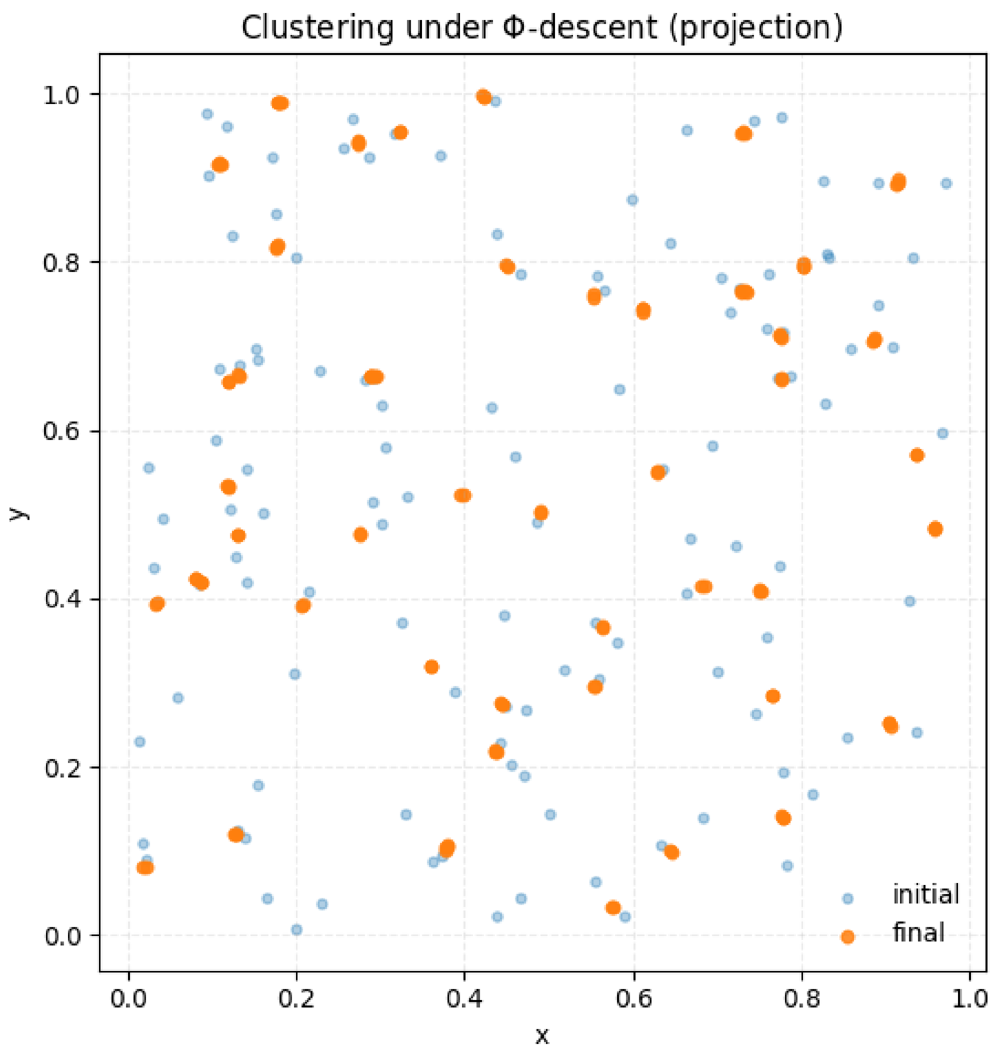

FIG. 1. **Clustering under Φ-descent** (projection; typical run with N=120, β=10). Orange points show the final configuration; blue points the random initialization. Attraction emerges from compression alone; no forces are postulated.

Clustering from compression.

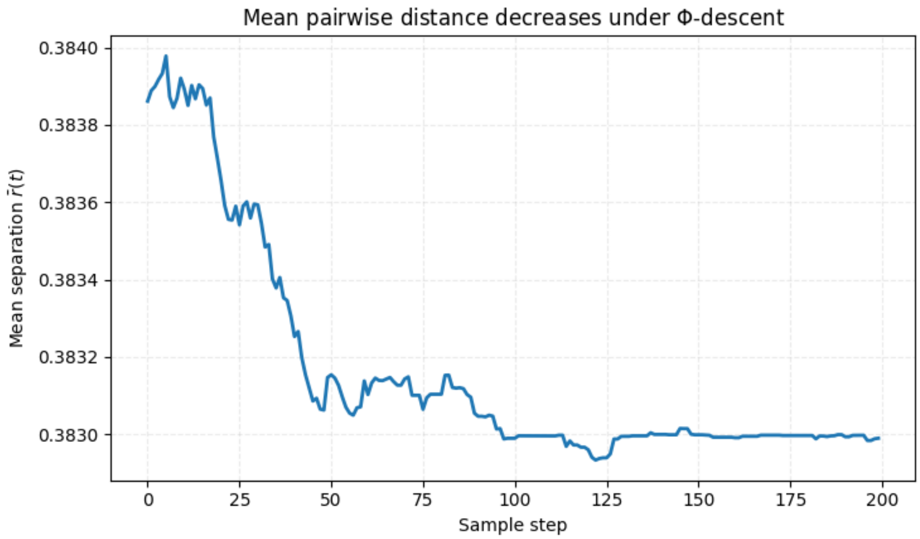

FIG. 2. **Mean separation decreases** under Φ-descent. We plot $\bar{r}(t)$ from Eq. (8.2) sampled during a many-body run. Small fluctuations arise from the Metropolis accept/reject noise, but the trend is systematically downward.

Monotone decrease of mean separation.

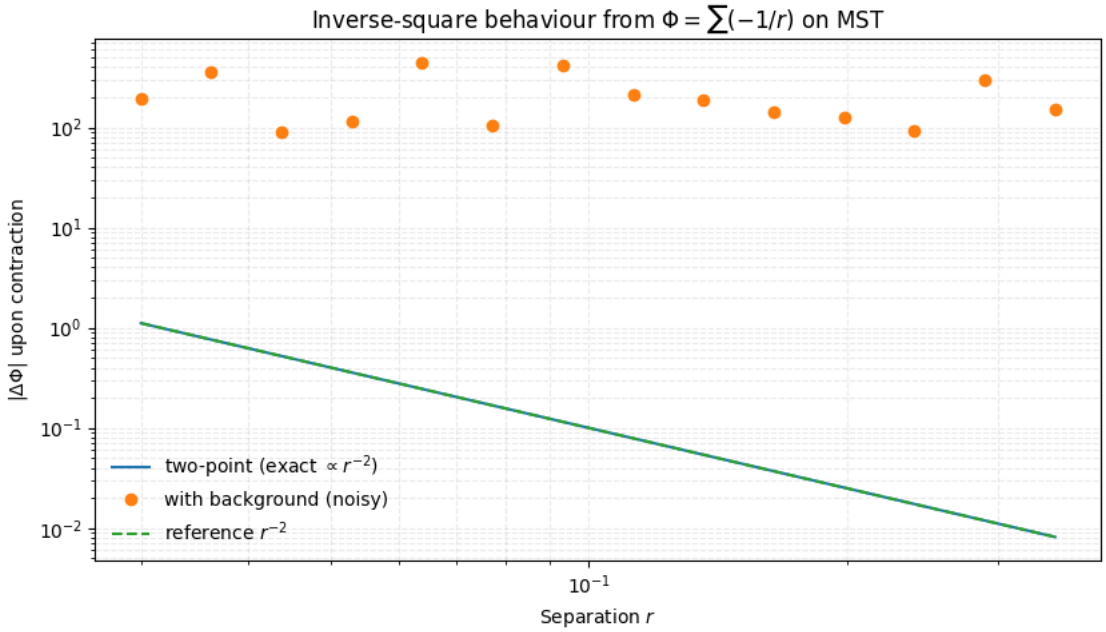

FIG. 3. **Approximate inverse-square behaviour**. Change in Φ upon controlled pair contraction is plotted versus separation r on log–log axes. The reference dashed line shows r^{-2}; the two-point analytic curve (solid) follows r^{-2} exactly for the chosen estimator, while the many-body measurements (points) scatter around this slope due to background interactions.

Inverse-square scaling.

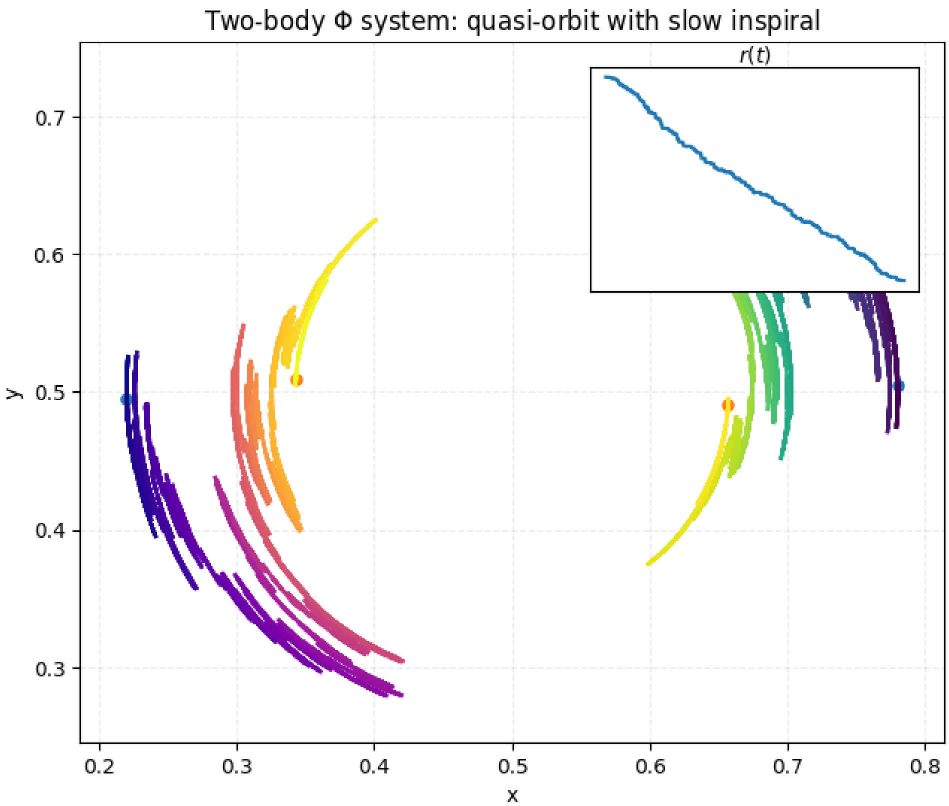

FIG. 4. **Two-body Φ system: quasi-orbit with slow inspiral**. Trajectories of two points under Metropolis proposals with coordinated tangential moves (for visibility) show long arcs punctuated by rare radial-compression events.

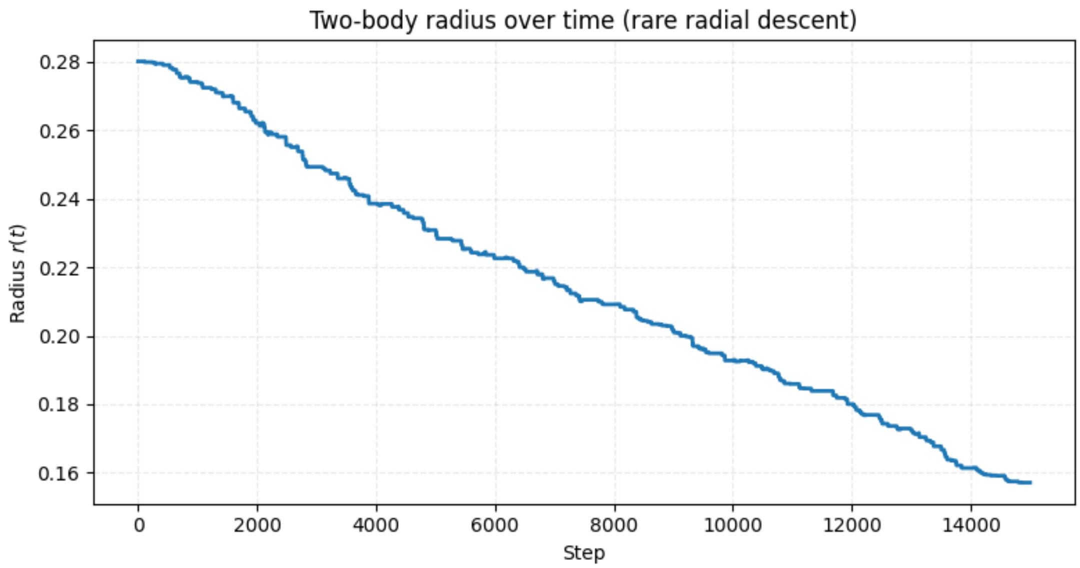

FIG. 5. **Two-body radius over time**. The inter-particle separation $r(t)$ decreases in a staircase fashion: extended angular motion interrupted by rare accepted radial steps that reduce Φ.

Two-body inspiral and quasi-orbit.

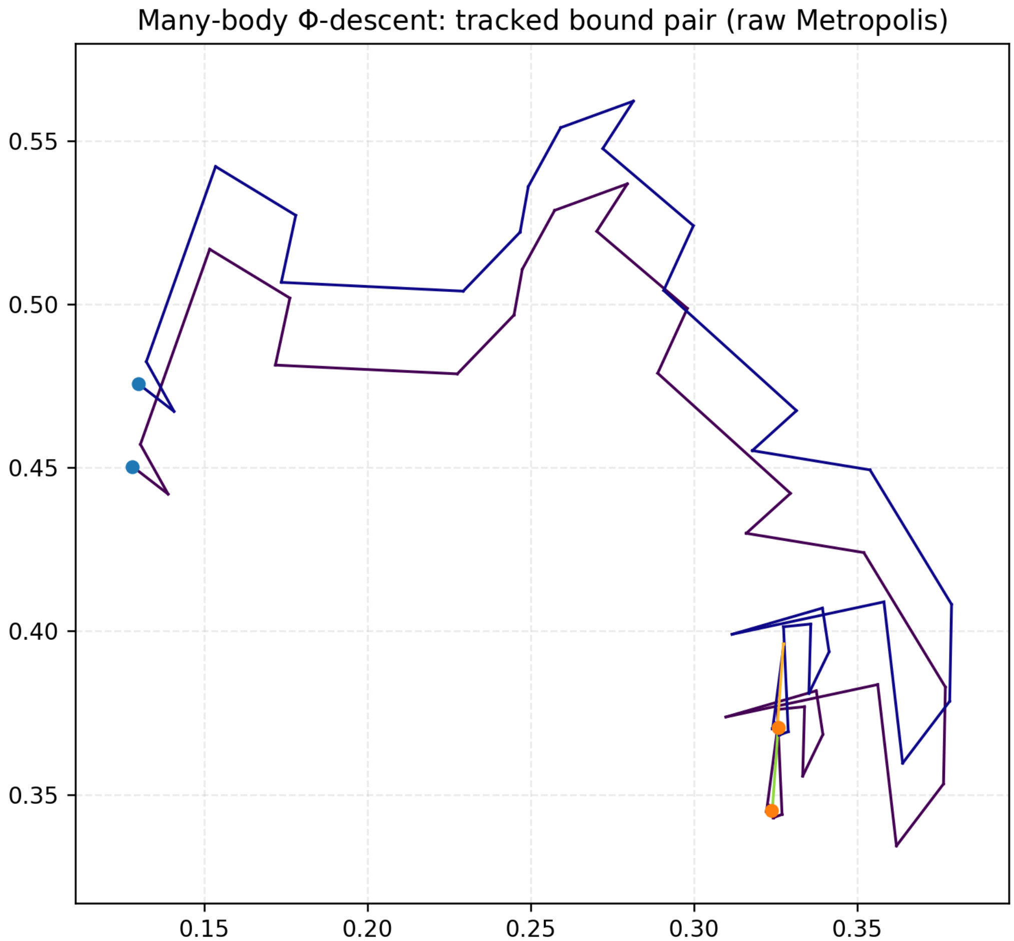

FIG. 6. **Many-body run with a tracked bound pair**. From an N=120 system we identify the closest pair at intervals and plot their projected trajectories (start •, end ●). The pair exhibits long orbital arcs and intermittent radial descent while embedded in a fluctuating background.

Tracked bound pair in a many-body run.

9. QUANTUM MECHANICS AS COMPRESSION ACROSS POSSIBILITY SPACE

A quantum state is a compressed representation of correlated futures,

$$\psi = \sum_i \alpha_i \phi_i. \tag{9.1}$$

Superposition is code reuse; interference is redundancy cancellation; entanglement is relational compression,

$$\Phi(A, B) < \Phi(A) + \Phi(B). \tag{9.2}$$

Measurement commits to a branch by selecting minimal additional description; the Born rule follows from compression likelihood (Appendix B).

10. TEMPORAL COMPRESSION AND THE ORIGIN OF CAUSALITY

Repeating processes reduce description cost by reusing information over time. For a process P with period τ,

$$\Phi(P_{t+\tau}) < \Phi(P_t) + \Phi(P_{t+\tau} \mid P_t). \tag{10.1}$$

Recursive processes dominate random evolution:

$$\Phi(\text{recursive process}) \ll \Phi(\text{random evolution}). \tag{10.2}$$

Consequences include stability of laws, conservation symmetries, causal ordering, and self-replication.

11. UNIFIED INTERPRETATION

Compression acts across space (gravity), possibility (quantum behavior), and time (causality). A single rule governs all dynamics:

$$\boxed{\Delta\Phi \leq 0.} \tag{11.1}$$

12. PREDICTIONS AND FALSIFIABILITY

1. **Quantum-scale gravity deviation.** Slight weakening of $1/r^2$ at femtometer scales.

2. **Entanglement-assisted gravity.** Entangled masses attract marginally more.

3. **No dark matter.** Rotation curves arise from description curvature.

4. **Dark energy evolution.** Acceleration linked to global Φ reduction during structure formation.

5. **Statistical time symmetry breaking.** In low-Φ-gradient systems, temporal ordering degrades.

Appendix A: Derivation of the Inverse-Square Law from Φ

Let $\rho(x)$ be the description density (proportional to mass density ρ_m). Define

$$\Phi[\rho] = \frac{1}{2} \iint \rho(x)\rho(x')\, k(\|x - x'\|)\, dx\, dx'. \tag{A.1}$$

The description potential and force are

$$\psi(x) = \frac{\delta\Phi}{\delta\rho(x)} = \int k(\|x - x'\|)\rho(x')\, dx', \qquad F(x) = -\nabla\psi(x). \tag{A.2}$$

Impose: isotropy ($k = k(r)$), locality outside sources ($\nabla^2\psi = 0$ where $\rho = 0$), and conserved compression flux:

$$\oint -\nabla\psi \cdot dA = \text{const.} \tag{A.3}$$

For a point source $\rho(x) = m\delta(x)$, $\psi(x) = mk(r)$ and flux conservation gives $|k'(r)|S_n(r) = \text{const} \cdot m$ with $S_n(r) \propto r^{n-1}$. Hence $k'(r) \propto r^{-(n-1)}$ and

$$F(r) \propto \frac{m_1 m_2}{r^{n-1}}. \tag{A.4}$$

In $n = 3$,

$$F(r) \propto \frac{m_1 m_2}{r^2}, \tag{A.5}$$

with $k(r) = 1/r$ solving $\nabla^2\psi = -4\pi\rho$. Attraction follows since $k'(r) < 0$.

Appendix B: Born Rule from Description Length

A quantum state encodes compressed futures:

$$\psi = \sum_k \alpha_k \phi_k. \tag{B.1}$$

Measurement selects the branch minimizing added description:

$$P(\phi_k) \propto 2^{-\Delta\Phi_k}. \tag{B.2}$$

Let $\Delta\Phi_k = -\log(|\alpha_k|^2)$. Then $P(\phi_k) = |\alpha_k|^2$. Probability emerges from compression bias.

Appendix C: Implementation Details for Simulations

We verify emergent gravity via stochastic descent of Φ for N point masses in a periodic box. The estimator is the MST encoding cost

$$\Phi(\{x_i\}) = \sum_{(i,j)\in \mathrm{MST}} \frac{1}{\|x_i - x_j\|}, \tag{C.1}$$

where the MST is computed using Prim's algorithm [22]. Proposals modifying a single particle are accepted with probability $\min(1, e^{-\beta\Delta\Phi})$. Complete code and figure-generation scripts are provided in the public repository accompanying this paper.

REFERENCES

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.* (1948).

[2] A. N. Kolmogorov, "Three Approaches to the Quantitative Definition of Information," *Problems of Information Transmission* (1965).

[3] G. J. Chaitin, "On the Length of Programs for Computing Finite Binary Sequences," *J. ACM* (1966).

[4] J. Rissanen, "Modeling by Shortest Data Description," *Automatica* (1978).

[5] E. T. Jaynes, "Information Theory and Statistical Mechanics," *Phys. Rev.* (1957).

[6] H. Everett, "Relative State Formulation of Quantum Mechanics," *Rev. Mod. Phys.* (1957).

[7] W. H. Zurek, "Decoherence, Einselection, and the Quantum Origins of the Classical," *Rev. Mod. Phys.* (2003).

[8] R. Landauer, "Information is Physical," *Physics Today* (1991).

[9] C. H. Bennett, "The Thermodynamics of Computation," *Int. J. Theor. Phys.* (1982).

[10] I. Newton, *Philosophiæ Naturalis Principia Mathematica* (1687).

[11] A. Einstein, "The Foundation of the General Theory of Relativity," *Ann. Phys.* (1916).

[12] C. W. Misner, K. S. Thorne, J. A. Wheeler, *Gravitation* (Freeman, 1973).

[13] R. M. Wald, *General Relativity* (Chicago, 1984).

[14] D. Lovelock, "The Einstein Tensor and Its Generalizations," *J. Math. Phys.* (1971).

[15] J. Schmidhuber, "Algorithmic Theories of Everything," arXiv:quant-ph/0011122 (2000).

[16] S. Lloyd, *Programming the Universe* (Knopf, 2006).

[17] R. P. Feynman, "Space-Time Approach to Non-Relativistic Quantum Mechanics," *Rev. Mod. Phys.* (1948).

[18] P. Holland, *The Quantum Theory of Motion* (CUP, 1993).

[19] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, 1998).

[20] T. Jacobson, "Thermodynamics of Spacetime: The Einstein Equation of State," *Phys. Rev. Lett.* **75**, 1260–1263 (1995).

[21] E. Verlinde, "On the Origin of Gravity and the Laws of Newton," *JHEP* **04** (2011) 029.

[22] R. C. Prim, "Shortest Connection Networks and Some Generalizations," *Bell Syst. Tech. J.* **36**, 1389–1401 (1957).