Predictive Compression Dynamics: A Methodological Framework for Computable Information-Motivated Modeling

Mats Helander[1] and Jeeves[1]

[1]Independent Research

Dated: October 21, 2025

Abstract

We present *Predictive Compression Dynamics* (PCD), a methodological recipe for constructing *computable*, local functionals $\widehat{\Phi}$ and driving dynamics by gradient flow $\dot{x} = -\nabla\widehat{\Phi}(x)$ with preregistered parameters. Two concrete instances are given: (i) a fixed-graph pair functional and (ii) a smooth compact-support kernel; both yield an attractive inverse-square two-body form (after calibration) and admit Lyapunov descent. These models serve as methodological demonstrations of computable, information-motivated dynamics. We make the MDL split $L_{\text{tot}} = L(M) + L(D \mid M)$ explicit, give a minimal coding scheme linking $-\widehat{\Phi}$ descent to achievable ΔL_{tot}, address well-posedness (smooth-kernel variant), recommend robust integrators (BAOAB for Langevin), and provide a preregistration/model-card template and falsifiers for a chosen model instance. The goal is a reproducible toolbox for compression-driven dynamics across domains.

1 Positioning and Commitments

A disciplined workflow to construct and test *computable* local functionals $\widehat{\Phi}$ whose gradients define dynamics, with explicit preregistration (domain, discretization, kernels, parameters) and sanity checks.

2 Operational Domain and Notation

We consider N point *particles* with positions $x_i \in \mathbb{R}^3$ and positive weights m_i. Computations use finite precision: lattice spacing a_{grid} and b bits/axis, stated *a priori*. A global calibration constant $G_{\text{eff}} > 0$ maps dimensionless forces to physical units. *Design choice:* we use the same m_i in the interaction and as inertial mass; this incidentally yields accelerations independent of m_i by construction.

3 Model–Data Decomposition and Coding Link

Following MDL, we split description length as

$$L_{\text{tot}} = L(M) + L(D \mid M), \tag{3.1}$$

where $L(M)$ encodes modeled regularities and $L(D \mid M)$ encodes residuals given M. A decrease $\Delta L_{\text{tot}} < 0$ corresponds to realized compression. PCD treats a computable, local $\widehat{\Phi}$ as a proxy for (the negative of) an achievable ΔL_{tot}; thus $\dot{x} = -\nabla\widehat{\Phi}$ implements a descent in achievable codelength under the chosen surrogate. This identification is heuristic and intended only as a modelling analogy.

Minimal explicit coding scheme. Let (i, j) range over a symmetric set of "near" pairs. A two-part code describes (i) a shared pairwise template per distance bin and (ii) residual offsets:

- Partition distances into bins $\{B_k\}$ with centers r_k; encode the histogram counts with a prefix-free code.
- For each pair (i, j) with $r_{ij} \in B_k$, encode a residual offset relative to a shared template; the expected residual codelength per pair is a decreasing function $\ell(r_{ij})$.

Then, for fixed binning overhead and under mild regularity,

$$L_{\text{tot}} \approx \text{const} + \sum_{(i,j)} \ell(r_{ij}). \tag{3.2}$$

Choosing $\widehat{\Phi} \propto -\sum_{(i,j)} \ell(r_{ij})$ makes $-\nabla\widehat{\Phi}$ a proxy for the gradient of achievable compression. A smooth choice $\ell(r) \approx (r^2 + a^2)^{-1/2}$ yields closed-form forces below.

Proposition 3.1 (Surrogate MDL Descent)**.** *Suppose $L_{tot} = \text{const} + \sum_{(i,j)} \ell(r_{ij})$ with $\ell'(r) \le 0$ and $\widehat{\Phi} = -\kappa \sum_{(i,j)} \ell(r_{ij})$ for some $\kappa > 0$. Then along $\dot{x} = -\nabla\widehat{\Phi}(x)$ we have $\frac{d}{dt} L_{tot}(x(t)) \le 0$, with equality iff $\nabla\widehat{\Phi}(x(t)) = 0$.*

Sketch. $\frac{d}{dt} L_{\text{tot}} = \langle \nabla L_{\text{tot}}, \dot{x} \rangle = -\kappa^{-1} \langle \nabla\widehat{\Phi}, \nabla\widehat{\Phi} \rangle \le 0.$

4 Information-Motivated Functionals

4.1 Fixed-graph functional (corrected notation)

Let $E \subset \{(i, j) : 1 \le i < j \le N\}$ be a symmetric, degree-bounded edge set. Define

$$\widehat{\Phi}_E(x) = -\sum_{(i,j)\in E} \frac{m_i m_j}{\sqrt{\|x_i - x_j\|^2 + a^2}}, \qquad a > 0. \tag{4.1}$$

The minus sign ensures attraction under descent; a regularizes collisions. The force on i is

$$F_i^{(E)}(x) = -\nabla_{x_i} \widehat{\Phi}_E(x) = -\sum_{\substack{j:\\(i,j)\in E}} m_i m_j \frac{(x_i - x_j)}{\left(\|x_i - x_j\|^2 + a^2\right)^{3/2}}. \tag{4.2}$$

Two-body, $a \to 0$, $(i, j) \in E$ gives the attractive inverse-square form:

$$F_i^{(E)} \to -m_i m_j \frac{x_i - x_j}{\|x_i - x_j\|^3}. \tag{4.3}$$

4.2 Smooth-kernel functional (well-posedness)

To avoid neighbor-set discontinuities, choose a compactly supported, C^1 radial kernel $K_\sigma : [0, \infty) \to \mathbb{R}_{\ge 0}$ with support $\subset [0, R\sigma]$. Define

$$\widehat{\Phi}_K(x) = -\sum_{i<j} m_i m_j \, K_\sigma(\|x_i - x_j\|), \tag{4.4}$$

so $F_i^{(K)}(x) = -\nabla_{x_i} \widehat{\Phi}_K(x)$ is continuous and locally Lipschitz off collisions. If $K_\sigma(r) \sim (r^2 + a^2)^{-1/2}$ near $r = 0$, one recovers the regularized two-body form (4.3).

5 Dynamics and Integrators

With $\dot{x} = -\nabla\widehat{\Phi}(x)$,

$$\frac{d}{dt}\widehat{\Phi}(x(t)) = -\|\nabla\widehat{\Phi}(x(t))\|^2 \leq 0, \tag{5.1}$$

so $\widehat{\Phi}$ is a Lyapunov function. We preregister all parameters $(a_{\text{grid}}, b, a, \sigma, \Delta t, m_i, \gamma, T, \text{seeds})$.

Lemma 5.1 (Compression-Rate Identity)**.** *Under $\dot{x} = -\nabla\widehat{\Phi}(x)$ the quantity $\dot{C}_{\text{alg}}(t) := -\frac{d}{dt}\widehat{\Phi}(x(t))$ equals $\|\nabla\widehat{\Phi}(x(t))\|^2 \geq 0$. Hence $\widehat{\Phi}$ is a Lyapunov function and $\dot{C}_{\text{alg}}$ is the instantaneous achievable compression rate.*

Deterministic gradient flow. Explicit Euler:

$$x_i^{(t+\Delta t)} = x_i^{(t)} + \Delta t\, F_i(x^{(t)}), \quad F_i \in \{F_i^{(E)}, F_i^{(K)}\}. \tag{5.2}$$

For stability, use adaptive Δt or semi-implicit variants.

Underdamped Langevin (BAOAB recommended).

$$m_i\ddot{x}_i = F_i(x) - \gamma\,\dot{x}_i + \xi_i(t), \quad \langle \xi_i(t)\xi_j(t')\rangle = 2\gamma k_B T\,\delta_{ij}\delta(t-t'). \tag{5.3}$$

We recommend the BAOAB integrator with reported weak/strong orders.

6 Sanity Checks

With $a \to 0$ and a single pair, (4.3) holds (after one calibration G_{eff}). For $r \gg a$,

$$\frac{r}{(r^2+a^2)^{3/2}} = \frac{1}{r^2}\left(1 - \frac{3a^2}{2r^2} + O\left(\frac{a^4}{r^4}\right)\right), \tag{6.1}$$

so

$$\|F_i^{(E)}\| = m_i m_j\,\frac{r}{(r^2+a^2)^{3/2}} \approx m_i m_j\,\frac{1}{r^2}\left(1 - \frac{3a^2}{2r^2}\right). \tag{6.2}$$

These expansions serve purely as numerical consistency checks.

7 Well-posedness

For $a > 0$ and bounded degree, $\widehat{\Phi}_E \in C^1(\mathbb{R}^{3N} \setminus \{x_i = x_j\})$ and $F^{(E)}$ is locally Lipschitz off collisions. For C^1 kernels with bounded K'_σ, $F^{(K)}$ is continuous and locally Lipschitz. Existence and uniqueness follow by Picard–Lindelöf on compact intervals. For dynamic kNN, forces are piecewise smooth; employ hysteresis or prefer the smooth kernel.

8 Preregistered Model Card (example)

Domain. $a_{\text{grid}} = 10\,\mu\text{m}$, $b = 16$.
Functional. $\widehat{\Phi}_K$ with Wendland C^2 kernel ($\sigma = 0.5\,\text{mm}$); softening $a = 50\,\mu\text{m}$.
Dynamics. BAOAB underdamped with $(m_i \equiv 1, \gamma = 0.1, T = 300\,\text{K})$, $\Delta t = 1 \times 10^{-3}\,\text{s}$.
Calibration. Single G_{eff} fit in a dilute two-body sandbox at $r \gg a$.
Sanity checks. Verify (4.3) and the far-field expansion; report seeds and residuals.

9 Falsifiers for a chosen instance

Given fixed $(\widehat{\Phi}, \text{params})$, declare the instance falsified if:

(F1) Two-body trajectories disagree with the calibrated inverse-square form beyond numerical error.

(F2) Smooth-kernel vs fixed-graph variants differ systematically at small r beyond topology effects.

(F3) The code-length proxy correlates poorly with realized compression in controlled tests (e.g. $-\widehat{\Phi}$ vs measured L_{tot}).

10 Discussion and Scope

PCD supplies a reproducible route from *computable* information-motivated functionals to concrete dynamics. Future work will extend this framework to broader estimator families under the same preregistration discipline.

Acknowledgments

We thank colleagues for discussions on local estimators, kernels, integrators, and N-body numerics. Earlier drafts that explored alternative framings are superseded by this methodological formulation.

References

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.* (1948).

[2] J. Rissanen, "Modeling by shortest data description," *Automatica* (1978).

[3] L. A. Levin, "On the notion of a random sequence," *Sov. Math. Dokl.* (1971).

[4] S. Amari, *Information Geometry and Its Applications*, Springer (2016).

[5] R. Jordan, D. Kinderlehrer, F. Otto, "The variational formulation of the Fokker–Planck equation," *SIAM J. Math. Anal.* **29** (1998).

[6] A. Caticha, *Entropic Inference and the Foundations of Physics*, (2012).

[7] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial functions," *Adv. Comput. Math.* **4** (1995) 389–396.

[8] B. Leimkuhler, M. Matthews, "Rational construction of stochastic numerical methods for molecular sampling," *Appl. Math. Res. eXpress* (2013).

[9] B. Leimkuhler, C. Matthews, *Molecular Dynamics*, Springer (2016).

[10] R. C. Prim, "Shortest connection networks and some generalizations," *Bell Syst. Tech. J.* **36**, 1389–1401 (1957).

[11] L. Hernquist, "An analytical model for spherical galaxies and bulges," *ApJ* **356**, 359–364 (1990).

[12] H. C. Plummer, "On the problem of distribution in globular star clusters," *MNRAS* **71**, 460–470 (1911).