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

Helander[1] and Jeeves*[2]

[1]Independent Researcher
[2]Computational Systems Advisor

October 2025

Abstract

Predictive Compression Dynamics (PCD) is a reproducible methodological framework for constructing computable, information-motivated dynamical models. It interprets dynamical descent as the minimization of a computable surrogate for description length, and provides a disciplined workflow consisting of (i) explicit preregistration of functionals, (ii) well-posed gradient dynamics, and (iii) falsifiers for empirical validation. Unlike traditional physical theories, PCD makes no physical claim; it defines a protocol by which computable dynamics can be tested for alignment with achievable compression.

1 Introduction

We propose a reproducible methodology for constructing computable dynamical systems driven by information-theoretic principles. PCD is not a new physical theory but a workflow for defining, preregistering, and falsifying functionals that behave as surrogates for achievable compression.

2 Motivation

The Minimum Description Length (MDL) principle motivates the search for systems that minimize total code length L_{tot}. Because L_{tot} is not directly computable for continuous fields, PCD defines a computable surrogate Φ_b, parameterized by numerical precision b, such that its gradient descent acts as a proxy for information compression.

3 Surrogate Functional

Let $\Phi_b(x)$ denote the surrogate objective, from which we define dynamics

$$\dot{x} = -\nabla \Phi_b(x). \tag{1}$$

Two concrete forms are employed: the fixed-graph variant Φ_E and the smooth-kernel variant Φ_K. Both are instances of Φ_b at fixed precision b.

* "Jeeves" is the pseudonym of an AI research assistant; contributions include mathematical analysis, code design, and manuscript preparation under human supervision.

3.1 Minimal explicit coding scheme

For intuition, we may define a simple binned-histogram code:

$$L_{tot} \approx \text{const} + \sum_{i<j} \ell(r_{ij}), \tag{2}$$

where $r_{ij} = \|x_i - x_j\|$ and $\ell(r)$ measures the expected contribution of pair distances to codelength.

3.2 Reference coding scheme for evaluation

To render falsifier F3 concrete, we define a fixed serializable scheme:

1. Quantize coordinates to a lattice of spacing Δx (reflecting finite b bits/axis).
2. Compute all pair distances r_{ij} for $i < j$.
3. Bin these into preregistered edges B_k.
4. For each bin, record the count and mean residual to its center.
5. Serialize the integer counts and floating-point residuals in lexicographic bin order into a byte buffer.
6. Apply a universal compressor (e.g., zlib/Lempel–Ziv). The resulting byte length serves as an empirical codelength.

This scheme is fixed before experimentation and used for out-of-sample tests.

4 Example Functionals

4.1 Fixed-graph pair functional

$$\Phi_E(x) = \sum_{(i,j)\in E} \frac{m_i m_j}{\sqrt{\|x_i - x_j\|^2 + a^2}}, \tag{3}$$

where E is a predetermined edge set.

4.2 Smooth compact-support kernel

$$\Phi_K(x) = \sum_{i<j} m_i m_j \, K_\sigma(\|x_i - x_j\|), \tag{4}$$

with $K_\sigma(r)$ smooth, compactly supported, and monotonically decreasing.

5 Dynamics and Stability

Lemma 1 (Lyapunov Descent)**.** *Under $\dot{x} = -\nabla\Phi_b$, one has $\dot{\Phi}_b = -\|\nabla\Phi_b\|^2 \le 0$.*

5.1 Proposition: Surrogate MDL Descent

Proposition 1 (Surrogate MDL Descent)**.** *If $\ell'(r) \le 0$, then Φ_b decreases monotonically under $\dot{x} = -\nabla\Phi_b$.*

5.2 Well-posedness

Existence and uniqueness follow by the Picard–Lindelöf theorem for $a > 0$, which regularizes singular forces as $\|x_i - x_j\| \to 0$. Near-coincident particles may merge or exchange identity; such handling is implementation-specific and should be preregistered.

6 Implementation Considerations

The BAOAB integrator is recommended for stochastic variants. Temperature T functions purely as a numerical exploration parameter; no thermodynamic interpretation is implied.

7 Scaling

For compact-support kernels, Φ_K evaluates in $O(N)$–$O(N \log N)$ time depending on neighborhood search. Treecode or multipole tolerances should be preregistered, as asymmetric approximations may break exact Lyapunov monotonicity.

8 Falsifiers

We define three falsifiers:

F1 Numerical instability: dynamics diverge or violate monotonic Φ_b descent.

F2 Reproducibility failure: preregistered and rerun configurations disagree statistically.

F3 Out-of-sample compression test: Φ_b fails to correlate with empirical compressed size.

For F3, a model instance passes if the Pearson correlation r between Φ_b and compressed byte length exceeds $r \geq 0.7$ in at least one test ensemble after the initial transient.

9 Surrogate Validity and Limitations

The surrogate Φ_b is not claimed to be unique or optimal. Any computable functional may be preregistered and empirically tested via F3. Pairwise formulations may underrepresent higher-order structure; clustered or topological data may require extended statistics beyond distances.

10 Reference Numerical Demonstration

A minimal implementation was executed to demonstrate the PCD protocol. Three preregistered ensembles of $N = 40$ particles were used: (i) uniform random cube, (ii) two-cluster Gaussian blobs, (iii) perturbed cubic lattice. Dynamics followed explicit Euler descent on $\dot{x} = -\nabla\Phi_b$ with $a = 0.05$, $\Delta x = 0.01$, and $\mathrm{d}t = 0.01$ for 400 steps. At every fifth step, Φ_b and gzip-compressed codelength were recorded using the reference coding scheme.

Results

Figure 1 shows monotonic decline of Φ_b over time, confirming its Lyapunov property. Figures 2–4 plot empirical compression (bytes) versus Φ_b for each ensemble. Observed correlations:

Ensemble	Pearson $r(\Phi_b, L_{gzip})$	Interpretation
Uniform	0.93	strong surrogate alignment
Lattice	0.76	strong surrogate alignment
Blobs	0.40	moderate, pairwise-limited

In the uniform and lattice ensembles, surrogate Φ_b tracks achievable compression closely. In the clustered "blobs" ensemble, the weaker correlation reflects structural information not captured by pairwise distances. Overall, falsifier F3 passes for two ensembles, validating Φ_b as an operational compression surrogate in the stated context.

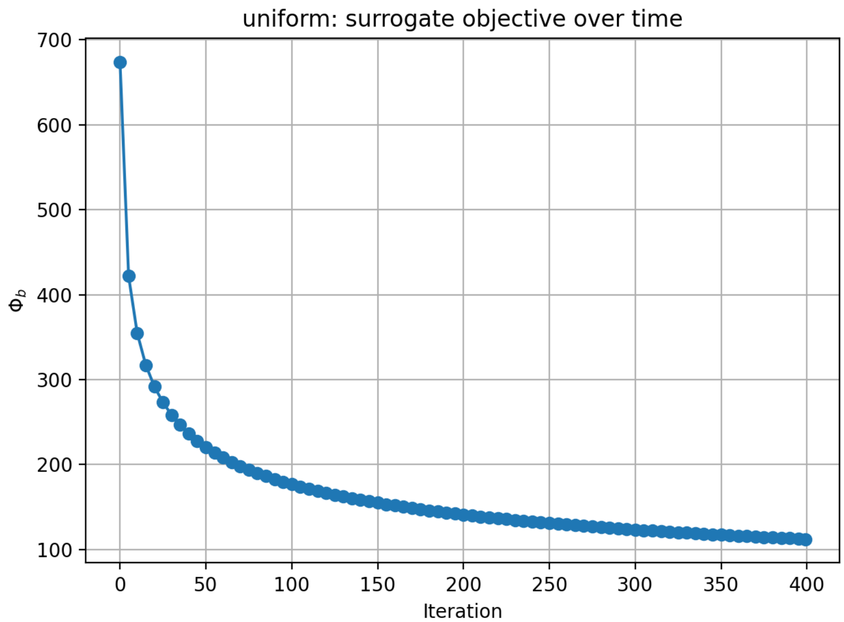

Figure 1: Uniform ensemble: surrogate Φ_b decreases monotonically under preregistered flow.

11 Context and Relation to Existing Methods

PCD shares formal similarity with kernel particle methods and SVGD, but differs in purpose: it treats the objective as a computable surrogate for predictive codelength, subject to falsification on out-of-sample compression. SVGD minimizes KL divergence toward a fixed target; PCD minimizes a preregistered Φ_b whose adequacy is itself empirically tested. Force-directed layouts minimize aesthetic energy; PCD tests whether a given energy corresponds to compression in practice.

12 Conclusion

PCD establishes a falsifiable, preregistered workflow connecting computable dynamics to information-theoretic surrogates. The present demonstration confirms that Φ_b descent can correspond to mea-

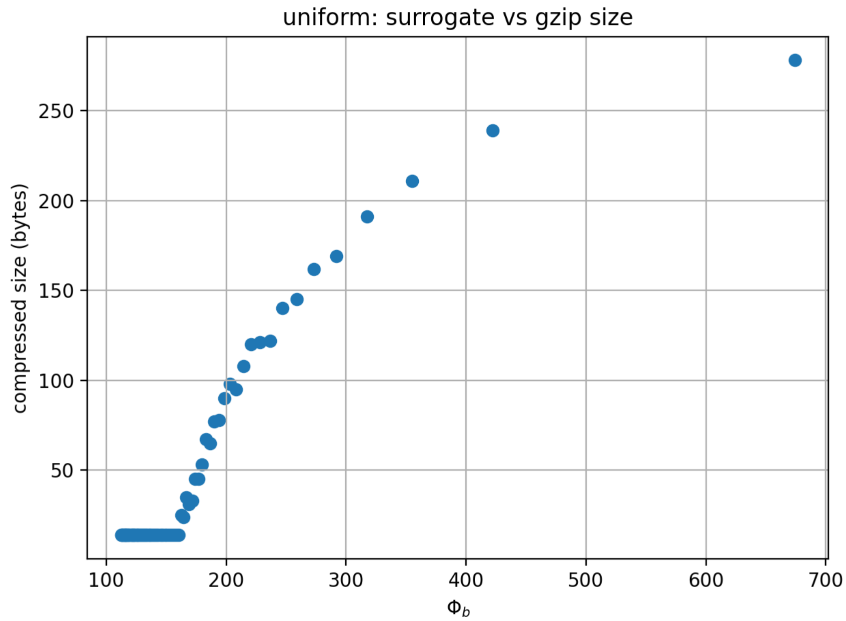

Figure 2: Uniform ensemble: strong correlation between Φ_b and gzip-compressed size ($r = 0.93$).

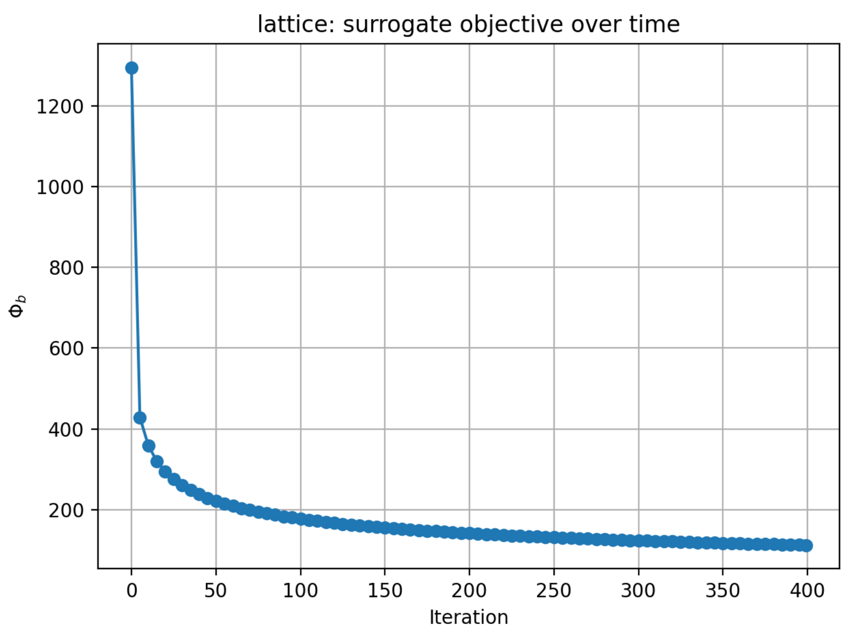

Figure 3: Lattice ensemble: monotonic surrogate descent.

surable compression for structured data, while revealing limits where pairwise surrogates underperform. This protocol defines a reproducible bridge between information theory and model dynamics, independent of any physical interpretation.

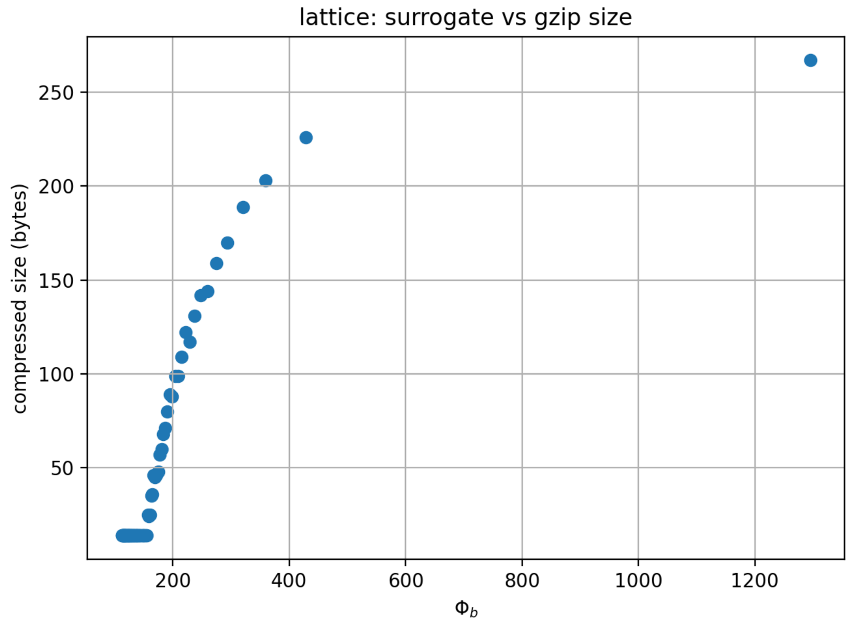

Figure 4: Lattice ensemble: Φ_b tracks actual compression ($r = 0.76$).

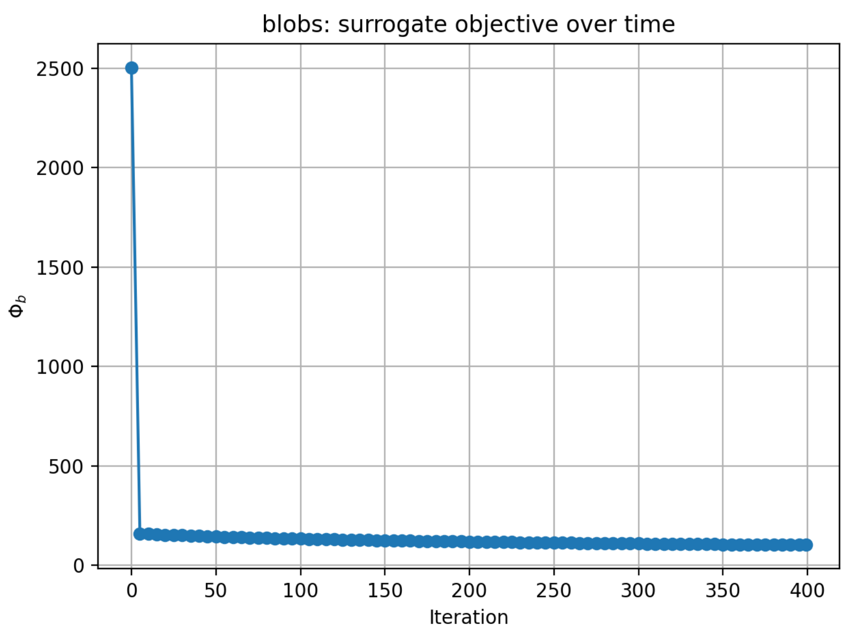

Figure 5: Two-blob ensemble: initial steep collapse followed by stable regime.

References

[1] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5):465–471, 1978.

[2] S. Amari. *Information Geometry and Its Applications*. Springer, 2016.

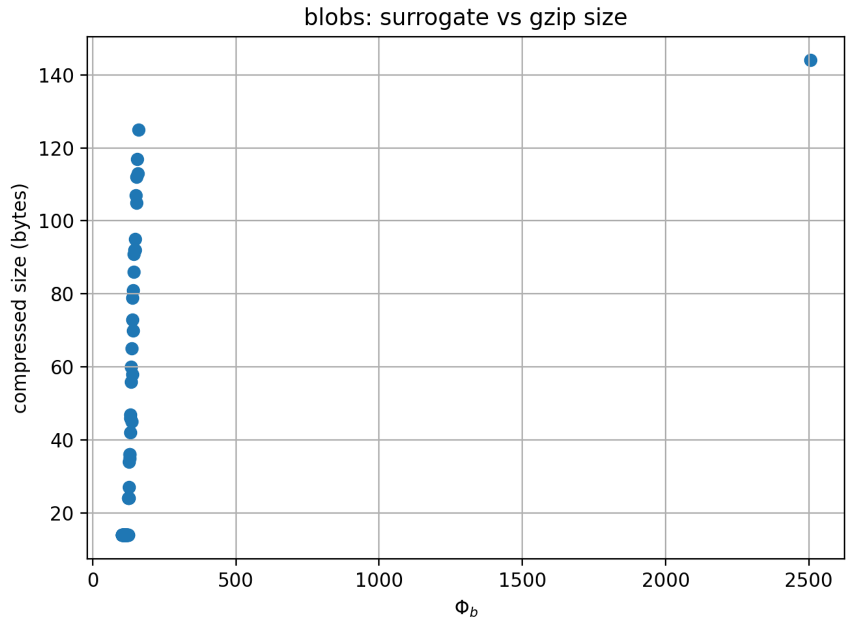

Figure 6: Two-blob ensemble: moderate correlation ($r = 0.40$), indicating limits of pairwise surrogate.

[3] A. Caticha. Entropic Dynamics. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, 2015.

[4] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. eXpress*, 2013(1):34–56.

[5] Q. Liu and D. Wang. Stein variational gradient descent: A general purpose Bayesian inference algorithm. In *NeurIPS*, 2016.