

Article

The Physics, Information, and Computation of Perennial Learning: Kolmogorov Complexity, Information Distance, and Port-Hamiltonian Thermodynamics

Chandrajit Bajaj 

Department of Computer Science and Oden Institute for Computational Engineering & Sciences,
The University of Texas at Austin, Austin, TX 78712, USA; bajaj@cs.utexas.edu

Abstract

Real-world autonomous agents learn under nonstationarity, safety constraints, and finite energetic budgets. We develop a framework for perennial learning—agents that continuously refine their models while provably controlling the cost of forgetting—by unifying three classical pillars: Kolmogorov complexity, which equates scientific discovery with algorithmic compression; Landauer’s principle, which assigns a minimal thermodynamic cost of $k_B T \ln 2$ per erased bit to every irreversible model update; and port-Hamiltonian (PH) dynamics, whose $(J - R)\nabla H$ decomposition separates zero-cost reversible inference from costly irreversible forgetting by construction. The Maxwell demon analogy is formalized: each learning episode is a Szilard cycle in which information acquisition, belief transport, and memory erasure must balance thermodynamically. The information-distance framework, comprising the normalized information distance (NID) and normalized compression distance (NCD), provides a computable geometry for measuring learning progress and guiding curriculum design. We separate the ideal uncomputable regularizer based on prefix complexity from the practical compressor/MDL (minimum description length) surrogate that appears in optimization and prove a calibration lemma linking the two under a mild uniform-accuracy assumption. Under explicit regularity, compact-sublevel, and non-energy-extracting assumptions, we prove a passivity speed limit for curriculum-induced contractions of the effective feasible set. Under local asymptotic normality, we reprove that Fisher information is a local posterior codelength proxy rather than an exact theorem about algorithmic entropy. A conditional sequential information-budget proposition shows that the per-stage sample requirement scales as $\tilde{O}(\Delta k_t / \lambda_*)$, where Δk_t is the number of materially changed model coordinates (not the total model complexity k_t); the $k^3 \rightarrow \Delta k$ improvement is conditional on a warm-start assumption and a chosen cold-start baseline. A double-integrator running example with a moving obstacle illustrates the architecture.

Keywords: perennial learning; Port-Hamiltonian systems; Kolmogorov complexity; normalized compression distance; curriculum learning; information thermodynamics; Differential Policy Optimization (DPO)



Academic Editor: Ming Li

Received: 21 March 2026

Revised: 3 May 2026

Accepted: 9 May 2026

Published: 13 May 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

1. Introduction

1.1. Dedication

This article is dedicated to Professor Paul M.B. Vitányi on the occasion of his 80th birthday. It is offered in recognition of his foundational contributions to the theory of computation, Kolmogorov complexity, normalized information distance, compression-based

similarity and clustering, model selection and structure functions, reversible computation, and the broad application of algorithmic information-theoretic ideas across computer science, mathematics, cognition, and the sciences [1–5].

This paper follows and attempts to build upon the algorithmic information-theoretic program associated with Li and Vitányi by treating learning as compression, transfer as information distance, and practical adaptation as a computable description-length surrogate for ideal Kolmogorov complexity.

1.2. *The Need for Perennial Learning*

Autonomous agents deployed in the physical world confront an unavoidable reality: the data distribution shifts continuously. Autonomous vehicles must accommodate evolving traffic patterns, weather, and infrastructure modifications [6,7]. In healthcare, wearable devices and adaptive prosthetics must recalibrate to a patient’s changing biomechanics. Industrial predictive-maintenance systems trained on data from new machinery become unreliable as components age. Smart-city infrastructure evolves over decades, with traffic, energy-consumption, and population profiles drifting continuously [8]. Even natural language processing requires models that track cultural and linguistic evolution without forgetting foundational structure.

Perennial learners deployed in consumer and cyber-physical settings also accumulate long-lived sensing histories, so physical safety must be considered together with information security and privacy. Our framework does not solve secure storage, encryption, or access control; rather, the compression/erasure accounting identifies which state summaries must persist, which can be reversibly transported, and which should be deleted.

Perennial learning—synonymous with lifelong or continual learning—refers to agentic systems that learn, search, and discover from a stream of data, adapting to new patterns while mitigating catastrophic forgetting [6,7]. Classical continual-learning methods such as elastic weight consolidation, synaptic intelligence, gradient episodic memory, and Progress and Compress focus on retaining useful parameters across task shifts [9–12]. They are algorithmically successful but mostly agnostic about the thermodynamic cost of a model update, how safe sets evolve during learning, and how the update rule should respect the underlying system structure.

1.3. *Computational Hardness: Why Perennial Refinement Is Necessary*

Most motion-planning problems are PSPACE-hard [1,13,14]. No polynomial-time algorithm can solve the general planning problem exactly. Approximate solutions must therefore be iteratively refined as new observations arrive. The quality of the approximation at any time step depends on the information accumulated thus far—precisely what Kolmogorov complexity measures—while the cost of updating the approximation as the environment changes is precisely what Landauer’s principle quantifies. Navigation under uncertainty exemplifies this: evaluating risk requires simultaneously sensing, identifying, and assessing hazardous objects at multiple spatial scales, from local obstacle avoidance (1:1) to global route planning (1:80,000), each with its own identifiability challenges and constraint geometry [8].

1.4. *Learning Has Its Own Hamiltonian*

The concept of learning Hamiltonians to solve tasks efficiently has deep classical roots in variational calculus [15,16]. The brachistochrone problem—finding the curve $z(x)$ that minimizes descent time—illustrates that nature “learns” optimal trajectories by extremizing an action functional. Our framework makes this literal: a port-Hamiltonian inference engine minimizes a variational free energy (the “action” of inference) subject to dissipative constraints (the “friction” of forgetting). Information thermodynamics interprets prediction and

adaptation as irreversibility-budgeted information processing [17–19]; PH modeling provides a native decomposition into lossless interconnection, explicit dissipation, and power ports [20]; linearly solvable control, path-integral control, and maximum-entropy RL show that information and control costs share a common variational language [21–23].

1.5. Contributions

This paper develops a perennial-learning viewpoint in which ideal algorithmic descriptions, practical code proxies, and port-Hamiltonian learning dynamics are kept conceptually distinct. The manuscript now makes that separation explicit.

Status map

Formal results: Theorems 1–3 and Propositions 1–5 under the stated assumptions. Design principles: PH channel separation, NCD-gated curriculum scheduling, the practical algorithmic-entropy diagnostic, and the warm-start sequential update rule. Analogies: the Maxwell/Szilard interpretation and Landauer accounting for abstract model updates unless a physical implementation model is specified.

The main technical contributions are as follows:

- (i) A clean separation between the ideal Kolmogorov-complexity description of model revision and the practical compressor/MDL surrogate used in optimization, together with a locally scoped calibration statement on the finite serialized model family actually visited by the curriculum.
- (ii) A PH “perennial inference engine” architecture that interprets reversible transport, dissipative forgetting, safety barriers, and Casimir locks as separate channels rather than as a single monolithic update.
- (iii) A standard passivity/dissipation estimate specialized to curriculum-induced feasible-set contraction, whose novelty lies in its interpretation as a rate limiter for safe curriculum scheduling.
- (iv) A local Fisher/Laplace/MDL proxy for posterior codelength contraction and a sequential information-budget proposition that clarifies when warm starts can reduce the effective per-stage sample requirement.
- (v) A concrete moving-obstacle double-integrator toy case study with NCD-gated curriculum scheduling, actual simulated trajectories, practical entropy/Fisher/overwrite diagnostics, a passivity-violation stress test, and a compact baseline table.

1.6. Roadmap

Section 2 presents the notations, scope, and standing assumptions, restoring the full Kolmogorov-complexity background and stating explicitly where the manuscript uses ideal K notions and where it uses the practical proxy L_Z . Section 3 turns algorithmic update costs into a computable optimal-control surrogate, formalizes the Maxwell/Szilard interpretation as a modeling lens, and develops the information-distance geometry of discovery. Section 4 presents the PH architecture, the toy example, the safety hierarchy, and the passivity speed limit interpreted as a curriculum-admissibility test. Section 5 develops the local Fisher/Laplace codelength proxy, NCD/NID-guided curriculum scheduling, and the sequential information budget. Section 6 specifies a structure-preserving discretization, algorithmic template, and a concrete numerical case study. Section 7 assembles the full engine and states the quantitative headline result together with the broader empirical validation agenda. Section 8 closes with limitations, lifted/reduced PH learning, and open problems.

Related works

The present paper is closest in spirit to MDL/Bayesian coding views of learning; information-bottleneck and rate-distortion perspectives on representation change; continual-learning methods such as EWC [9], SI [10], GEM [11], and Progress and Compress [12]; and safe-control/safe-RL methods built around confidence sets, reachability, or barrier certificates. Our point is not to replace these lines of work but to organize them through a PH separation of reversible transport, irreversible overwrite, and safety channels. In control, path-integral control [22] and linearly solvable Markov decision processes [21,23] link information and control costs but do not address lifelong overwrite budgets. Safe RL via Gaussian-process confidence sets [24] and Hamilton–Jacobi reachability [25] handle safety but not the code-length cost of updating safety knowledge. Structure-preserving ML [26–28] exploits PH structure for prediction but not for budgeted perennial learning. Geometric optimization and completeness/solvability questions relevant to structured constraints are discussed in [29–31].

2. Background, Scope, and Standing Assumptions

2.1. Description Complexity, Proxies, and Landauer Cost

Definition 1 (Prefix Kolmogorov complexity ([32], Ch. 2)). *Let U be a universal prefix-free Turing machine. The **prefix Kolmogorov complexity** of a binary string x is*

$$K(x) := \min\{\ell(p) : U(p) = x\}, \quad (1)$$

where $\ell(p)$ denotes the length of program p .

Definition 2 (Conditional complexity [33]). $K(x | y) = \min\{\ell(p) : U(p, y) = x\}$.

Definition 3 (Algorithmic mutual information). $I(x : y) = K(x) - K(x | y) = K(y) - K(y | x) + O(\log)$.

Throughout, $O(\log)$ absorbs terms of order $O(\log(K(x) + K(y)))$.

For optimization, we use a computable codelength $L_Z(x)$, obtained from a fixed lossless compressor or MDL code [4,34]. Unless otherwise stated, the primary compressor is LZMA/XZ with preset 6 applied to a canonical UTF-8 byte string. Each environment is serialized as sorted-key JSON metadata, followed by ordered, quantized obstacle boundary samples and, for grid tasks, row-major signed-distance samples. The notation L_Z is reserved for this practical proxy throughout this paper, and Appendix A provides the serialization, quantization, and sensitivity checks against gzip/zlib level 6 and a parameter-vector encoding. Exact K is uncomputable [32]; therefore, any formal theorem about an implemented algorithm must use L_Z or another explicit surrogate.

The Shannon entropy of a computable distribution P over a countable set \mathcal{X} is $H(P) = \sum_x P(x) \log(1/P(x))$.

Theorem 1 (Entropy–complexity bridge ([32], Thm. 8.1.1)). *For every recursive distribution P ,*

$$0 \leq \sum_x P(x) K(x) - H(P) \leq c_P, \quad (2)$$

where c_P depends only on the length of the shortest program computing P .

Remark 1 (Implication for perennial learning). *A learning agent maintains a model parameterized by θ , inducing a distribution P_θ over observations. The agent’s “understanding” of the*

data is the extent to which P_θ compresses observations, i.e., the extent to which $\sum P_\theta(x)K(x)$ approaches $H(P_\theta)$. Discovery corresponds to finding shorter descriptions: when the agent discovers structure, $K(\text{model})$ may increase modestly but $H(\text{data} \mid \text{model})$ decreases dramatically, and the net algorithmic entropy decreases.

Theorem 2 (Landauer’s principle [35]). Any logically irreversible device that erases b bits must dissipate at least $Q_{\text{erase}} \geq b k_B T \ln 2$ of heat into a bath at temperature T . Reversible reorganization avoids this lower bound in principle [2,36,37].

Remark 2 (What the Landauer lower bound actually applies to). Theorem 2 lower-bounds the heat cost of the irreversible part of an update in a concrete physical device. When the manuscript applies Landauer accounting to abstract model updates, it does so as a formal bookkeeping convention for irreversible overwrite unless a computation substrate is explicitly modeled. Thus, $k_B T \ln 2$ per erased bit is not asserted to be the literal energy paid by a software optimizer. The PH architecture below is designed to separate reversible transport from irreversible overwrite as clearly as possible [38].

2.2. Port-Hamiltonian Systems

Definition 4 (Port-Hamiltonian systems [20]). Let $\mathcal{X} \subseteq \mathbb{R}^{d_x}$ and let $H : \mathcal{X} \rightarrow \mathbb{R}$ be C^1 . A **port-Hamiltonian system** (PH) has the form

$$\dot{x} = (J(x) - R(x))\nabla H(x) + G(x)u(t), \tag{3}$$

where $J(x) = -J(x)^\top$ is the interconnection structure, $R(x) = R(x)^\top \succeq 0$ is the dissipation matrix, and $G(x)$ is the input-port map. The power-conjugate output is $y(x) = G(x)^\top \nabla H(x)$.

Proposition 1 (Passivity identity [20]). Along any sufficiently smooth trajectory of (3),

$$\frac{d}{dt}H(x(t)) = -\nabla H(x(t))^\top R(x(t))\nabla H(x(t)) + y(t)^\top u(t). \tag{4}$$

In particular, when $u \equiv 0$, the Hamiltonian is nonincreasing.

Definition 5 (Casimir lock [20]). A smooth function $C : \mathcal{X} \rightarrow \mathbb{R}$ is a **Casimir invariant** if $J(x)\nabla C(x) = 0$ for all $x \in \mathcal{X}$. Casimir invariants are preserved along every PH trajectory regardless of the choice of H . We use the phrase **Casimir lock** for an engineered invariant that the learner is not allowed to overwrite.

No theorem in this paper proves that PH is uniquely optimal among all learning architectures. The PH decomposition is used because it is structurally interpretable: reversible interconnection and irreversible dissipation are explicit, and hard invariants can be encoded at the Poisson and dissipative bracket levels [39,40].

2.3. Standing Assumptions

Assumption 1 (PH regularity and compact sublevels). The maps H, J, R, G are locally Lipschitz on \mathcal{X} , H is C^2 , and all trajectories of interest remain in a compact sublevel set $\mathcal{X}_{E_0} := \{x \in \mathcal{X} : H(x) \leq E_0\}$. On each such set,

$$D_{\max}(E_0) := \sup_{x \in \mathcal{X}_{E_0}} \nabla H(x)^\top R(x)\nabla H(x) < \infty. \tag{5}$$

Assumption 2 (Observation model and local asymptotic normality). *Observations satisfy*

$$y_i = h_{\xi^*}(x_i) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \Gamma_y), \tag{6}$$

where $\Gamma_y \succ 0$ and $h_{\xi}(x)$ is C^2 in (x, ξ) . For the local Fisher proxy, we assume that the posterior around the current iterate is well approximated by a Gaussian with a precision matrix equal to the accumulated Fisher information.

Assumption 3 (Local/probabilistic proxy calibration). *Whenever a theorem transfers an ideal K -based statement to the implemented code, the comparison is restricted to the finite serialized model family actually visited by the curriculum, denoted by Ξ_T . With probability at least $1 - \delta$ over the sampling/serialization process,*

$$|L_Z(\xi) - K(\xi)| \leq c_Z(\delta), \quad \forall \xi \in \Xi_T. \tag{7}$$

Outside Ξ_T , the manuscript makes no global calibration claim.

Assumption 4 (Sequential drift and warm start). *Between curriculum stages $t - 1$ and t , only an active block $S_t \subseteq \{1, \dots, d_{\xi}\}$ of size $\Delta k_t := |S_t|$ changes materially, and the restricted Fisher information on that block satisfies $\mathcal{I}_t|_{S_t} \succeq \lambda_* I$ for some $\lambda_* > 0$. The estimator is initialized from the previous stage rather than from scratch.*

Remark 3 (Scope). *Assumption 3 is deliberately local: it is a statement about the finite family of serialized models actually visited by the curriculum rather than a universal theorem about arbitrary compressors or arbitrary task classes. The main optimization problem and all numerical examples can be read entirely in terms of L_Z without appealing to K at all.*

3. Thermodynamic Formulation of Perennial Learning

In this section, we first explain how the lifelong regularizer uses a computable code length. Next, we describe how the optimal-control objective is a surrogate for algorithmic rate distortion while introducing an entropy-style discovery diagnostic. Finally, we formalize the Maxwell demon analogy and develop information distance as the geometry of discovery.

3.1. Ideal and Practical Lifelong Regularizers

Let ξ_t denote the model or solver description at curriculum step t .

Definition 6 (Ideal and practical overwrite cost). *The ideal bit-overwrite count at step t is $\Delta K_t^* := [K(\xi_t) - K(\xi_{t-1})]_+$. The practical overwrite count is $\widehat{\Delta K}_t := [L_Z(\xi_t) - L_Z(\xi_{t-1})]_+$. The corresponding lifelong penalties are*

$$\Omega_{\text{life}}^* := \sum_{t=1}^T \Delta K_t^*, \quad \Omega_{\text{life}}^Z := \sum_{t=1}^T \widehat{\Delta K}_t. \tag{8}$$

Proposition 2 (Calibration transfer). *Under Assumption 3, $|\Delta K_t^* - \widehat{\Delta K}_t| \leq 2c_Z$ for each t , and therefore $|\Omega_{\text{life}}^* - \Omega_{\text{life}}^Z| \leq 2Tc_Z$.*

Proof. The map $a \mapsto [a]_+$ is 1-Lipschitz, so

$$|\Delta K_t^* - \widehat{\Delta K}_t| \leq |(K(\xi_t) - K(\xi_{t-1})) - (L_Z(\xi_t) - L_Z(\xi_{t-1}))| \leq 2c_Z.$$

Summing over t proves the second bound. \square

3.2. From Algorithmic Rate Distortion to a Computable Optimal Control Problem (OCP)

A critical constraint on any learning process is the principle of nonincreasing mutual information.

Theorem 3 (Nonincreasing mutual information ([32], Thm. 8.1.4)). *For deterministic processing $z = f(x)$,*

$$I(z; y) \leq I(x; y) + K(f) + O(1). \tag{9}$$

Data processing cannot create information; any apparent information gain from computation is bounded by the complexity of the computation itself.

The algorithmic rate-distortion function sharpens the compression–fidelity tradeoff.

Definition 7 (Algorithmic rate distortion ([32,41], Def. 8.1.5)). *For an individual object x and distortion measure d ,*

$$r_x(\delta) = \min_y \{K(y) : d(x, y) \leq \delta\}. \tag{10}$$

This provides a per-object compression–fidelity tradeoff rather than a statistical one. The PH solver’s variational free energy (Section 4) instantiates exactly this tradeoff: the agent seeks the shortest model ($\min K(\theta)$) whose predictions match the data within tolerance.

Definition 8 (Thermodynamic work of model update ([32], Thm. 8.2.4)). *The minimal thermodynamic work required to transform model state x to model state y is*

$$W(y | x) = K(x) - K(y). \tag{11}$$

If the new model is simpler ($K(y) < K(x)$), the system *extracts* work—it has discovered structure. If the new model is more complex ($K(y) > K(x)$), the system *pays* work—it is storing more information.

Remark 4 (Speed–dissipation tradeoff). *The time-bounded complexity satisfies $K^t(x) > K(x)$ in general. A demon (learning agent) with limited computation time pays $K^t(x) - K(x)$ in excess dissipation. This gap constrains the regret of Differential Policy Optimization (DPO) [42] (Section 5.4) and is consistent with the reversible/adiabatic-computation viewpoint that time and space resources can be traded against energy expenditure [2].*

In optimization, K is replaced with its computable surrogate. We therefore define the running cost using observable distortion and practical codelength increments [43].

Definition 9 (Computable perennial OCP). *At stage t , with state x , control u , observation y_t , model ξ_t , and meta-parameters η (the top-level policy/design variables; see Section 4.1), define*

$$\begin{aligned} \ell_t(x, u; \mathcal{E}_t) = & \underbrace{\lambda_{\text{err}} \frac{1}{2} \|y_t - h_{\xi_t}(x)\|_{\Gamma_y^{-1}}^2}_{\text{predictive distortion}} + \underbrace{\lambda_{\text{code}} \widehat{\Delta K}_t}_{\text{irreversible update proxy}} \\ & + \lambda_{\text{safe}} \underbrace{\psi_{\mathcal{F}_t}(x)}_{\text{safety barrier/margin}} + \lambda_u \|u\|^2. \end{aligned} \tag{12}$$

The stage objective is

$$\mathcal{J}_t(\eta) := \mathbb{E} \left[\int_0^{T_t} \ell_t(x(s), u(s); \mathcal{E}_t) \, ds + \phi_t(x(T_t)) \right]. \tag{13}$$

3.3. Algorithmic Entropy as a Discovery Diagnostic

We now define the algorithmic entropy of physical systems following Li and Vitányi [32], §§8.5–8.6 and introduce a computable surrogate for monitoring discovery [44].

Definition 10 (Algorithmic entropy ([32], Def. 8.6.2)). *The algorithmic entropy of a macrostate x is*

$$S_A(x) = (k_B \ln 2)(K(x) + H_x), \quad (14)$$

where $K(x)$ is the prefix complexity of the macroscopic description and $H_x = S_B(x)/(k_B \ln 2)$ is the log-volume of the macrostate.

The decomposition splits total physical entropy into the complexity of what we know (the regularity) and our ignorance about the microstate given the macroscopic description.

Remark 5 (Diagnosis: dS_A/dt sign criterion). *For a regular (compressible) microstate, increasing measurements cause H_x to decrease rapidly while $K(x)$ increases slowly, so the net S_A decreases—discovery is occurring. For a random (incompressible) microstate, H_x decreases but $K(x)$ increases at the same rate, so S_A remains flat—the agent is memorizing noise. The sign of dS_A/dt is the algorithmic thermodynamic diagnosis of identifiability: regions where $dS_A/dt < 0$ are identifiable; regions where $dS_A/dt \approx 0$ are not.*

Definition 11 (Coarse-grained algorithmic entropy ([32], Def. 8.6.6)).

$$H_\mu^n(\omega) = \inf_{i \leq n} \{H_\mu(\Gamma_{\omega_{1:i}})\}, \quad (15)$$

which satisfies the second law (strong entropy growth) and converges: $H(\omega) = \lim_{n \rightarrow \infty} H^n(\omega) = \inf_{x \in \{0,1\}^*} \{H(\Gamma_x) : \omega \in \Gamma_x\}$.

The optimal measurement precision n_0 is the level at which additional complexity $K(\omega_{1:n})$ no longer yields a sufficient decrease in $\log \mu(\Gamma_{\omega_{1:n}})$. This defines a natural stopping criterion for measurement refinement in the curriculum (Section 5.2).

For monitoring purposes, we use a computable surrogate.

Definition 12 (Practical algorithmic-entropy surrogate). *Let m_t denote the current compressed model description, and let H_t denote a coarse uncertainty volume (e.g., log-volume of a posterior credible ellipsoid). Define*

$$\widehat{S}_A(t) := (k_B T \ln 2)(L_Z(m_t) + H_t). \quad (16)$$

Remark 6 (Status of the entropy diagnostic). *Figure 1 should be read as a design diagnostic rather than as a theorem in which $\widehat{S}_A(t)$ is monotone in every learning problem. The operational point is as follows: monitor whether uncertainty is being converted into compression or merely into parameter count.*

3.4. Maxwell's Demon: The Learning Agent as a Thermodynamic Engine

The Maxwell/Szilard discussion below is used as the motivation and as a bookkeeping analogy. The formal claims of this paper are the explicit propositions and assumptions rather than the analogy itself.

The Szilard engine [45] provides the following canonical thought experiment: a demon with memory observes a single-molecule gas, records the molecule's position, and extracts $k_B T \ln 2$ of work per observation cycle. The apparent violation of the second law is resolved by Bennett [36]: the demon must *erase* its memory to complete the cycle, incurring a Landauer cost $\geq k_B T \ln 2$ [46].

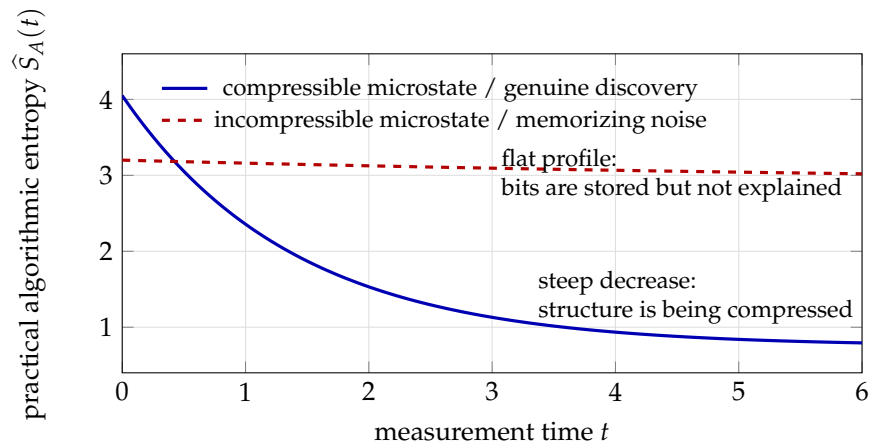


Figure 1. Schematic diagnostic; not empirical data. **Algorithmic-entropy diagnostic (Section 3.3).** Decreasing $\hat{S}_A(t)$ indicates that measurements are reducing uncertainty faster than model codelength grows (genuine discovery); a nearly flat curve indicates memorization of noise. The figure is purely schematic; it is intended to visualize a diagnostic criterion rather than report empirical data.

Claim 1 (Thermodynamic balance sheet ([32], §8.6.1, Claim 8.6.1)). *For a combined system (engine + demon) with algorithmic entropy $S_A = (k_B \ln 2)(K(x) + H_y)$, where $K(x)$ is the complexity of the demon’s memory and H_y is the Boltzmann part of the engine, the net heat gained over one cycle satisfies*

$$\Delta Q = (S_A^f - S_A^i)T = \Delta Q^+ + \Delta Q^- \leq 0, \tag{17}$$

where $\Delta Q^+ = (S_B^f - S_B^i)T$ is the heat from the engine’s entropy change and $\Delta Q^- = (K(f) - K(i)) \cdot k_B T \ln 2$ is the heat lost to the demon’s memory update.

One possible mapping to perennial learning is shown in Table 1.

Table 1. A nomenclature mapping between the Maxwell’s Demon and the Learning Agent.

Demon Component	Learning-Agent Analog
Demon’s memory	Meta-parameters η (policy/design variables)
Measuring molecule speed	Observing data
Sorting fast/slow	Updating posterior
Memory erasure	Overwriting old beliefs (Ω_{life})
Heat-bath temperature T	Environment stochasticity
Szilard cycle	One episode of online learning

From the demon analysis, the physically correct form of the lifelong regularizer is

$$\Omega_{\text{life}}(\eta) \propto \sum_{\text{tasks}} [K(\xi_{\text{new}}) - K(\xi_{\text{old}})]_+, \tag{18}$$

penalizing cumulative positive complexity increase across task transitions.

Remark 7 (Intelligent vs. unintelligent erasure). *An unintelligent demon erasing a random-looking string x pays n bits (the full length). An intelligent demon that recognizes that x encodes a compressible object compresses to $K(x) \ll n$ before erasing, paying much less. With limited time, the demon pays $K^t(x) > K(x)$. This speed–dissipation tradeoff (Remark 4) constrains the DPO regret bound and aligns with the reversible-computation perspective that extra time/space can be exchanged for reduced energetic cost [2].*

3.5. Information Distance and the Geometry of Discovery

We develop the information distance as the metric for measuring learning progress and guiding the curriculum, following Li and Vitányi [32], §8.3–8.4. On the computational side, the normalized-compression viewpoint of Li et al. [3] and the clustering-by-compression program of Cilibrasi and Vitányi [5] are central rather than peripheral: they supply the operational surrogate used later for curriculum gating.

Definition 13 (Information distance ([32], §8.3)). *The max distance between strings x and y is*

$$E_1(x, y) = \max\{K(x | y), K(y | x)\}. \quad (19)$$

The sum distance is

$$E_3(x, y) = K(x | y) + K(y | x) \pm O(\log). \quad (20)$$

The sum distance E_3 measures the total irreversible bit flow during a reversible computation from x to y and is the correct cost metric for the PH solver, which consumes observations (bits in) and discards old beliefs (bits out).

Theorem 4 (Universality ([32], Thm. 8.3.2)). *E_1 is minimal among all admissible distances: every computable distance between x and y is at least $E_1(x, y)$ up to an additive constant.*

Definition 14 (Normalized information distance (NID) ([32], Def. 8.4.1)).

$$e(x, y) = \frac{\max\{K(x | y), K(y | x)\}}{\max\{K(x), K(y)\}}. \quad (21)$$

This takes values in $[0, 1]$ and satisfies metric properties ([32], Thm. 8.4.1).

The practical approximation replaces K with the output length of a compressor Z , yielding the normalized compression distance (NCD) [3,5]:

$$\text{NCD}_Z(x, y) = \frac{Z(xy) - \min\{Z(x), Z(y)\}}{\max\{Z(x), Z(y)\}}. \quad (22)$$

We address how information and compression distances connect to perennial learning:

- (a) Task relatedness for curriculum design: $e(\mathcal{E}_t, \mathcal{E}_{t+1})$ measures how much the task structure changes between curriculum steps.
- (b) Model-change tracking: $e(\xi_t, \xi_{t+1})$ measures how much the PH solver parameters have changed; a large NID implies a higher Landauer cost.
- (c) Discovery progress: $e(\text{model}_t, \text{ground truth})$ monotonically decreases, signaling convergence toward the ground truth.
- (d) Safe feasibility-set change: $e(\mathcal{F}(\kappa_t), \mathcal{F}(\kappa_{t+1})) \leq \delta_{\text{safe}}(\tau)$ bounds the safe curriculum step size, where κ parameterizes the constraint geometry (defined in Example 1) and δ_{safe} is the NCD-derived threshold (defined in Section 5.2).

4. The Perennial Inference Engine

This section turns the abstract ingredients into a dynamical architecture. The key claim is architectural rather than variational: if observation, reversible transport, dissipation, and hard invariants are kept as separate channels, then update cost and safety become inspectable instead of being hidden inside a generic optimizer.

4.1. Architecture

The meta-policy π_η maps an environment \mathcal{E} to PH solver parameters $\xi = \pi_\eta(\mathcal{E})$. The (q, p) -form PH solver is

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} \nabla_p H_\xi \\ -\nabla_q H_\xi \end{pmatrix} + \begin{pmatrix} 0 \\ -R_\xi \nabla_p H_\xi \end{pmatrix} + \begin{pmatrix} 0 \\ G_\xi u_\xi \end{pmatrix}. \tag{23}$$

The stochastic extension is

$$dx = [(J(x) - R(x))\nabla H(x) + G(x)u(t)] dt + \Sigma(x) \circ dW_t, \tag{24}$$

where W_t is a Wiener process and \circ denotes the Stratonovich integral [47]. Figure 2 shows the architecture of a perennial inference engine.

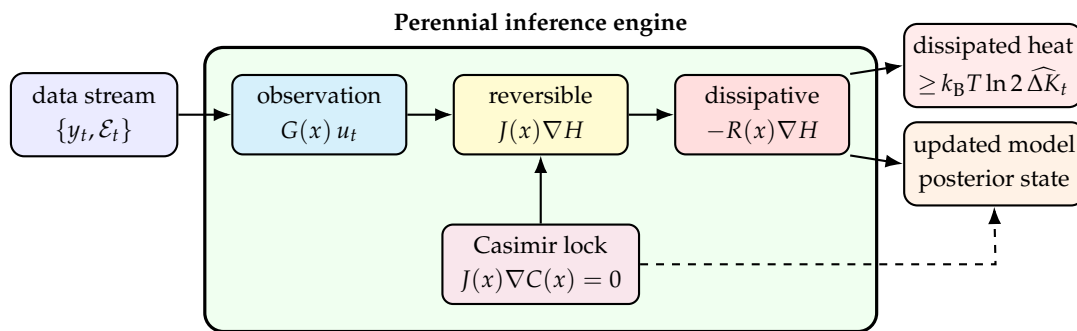


Figure 2. Perennial inference engine architecture display. Three explicit components in differential form handle data input, reversible transport, and irreversible (dissipative) overwrite. Only the dissipative branch carries a Landauer lower bound; the Casimir lock (dashed arrow) enforces the “never overwrite” structural invariant without dissipating energy.

Table 2 summarizes the thermodynamic interpretation. The passivity inequality (Proposition 1) ensures that, in the absence of external input, the energy (and hence the “distance to danger”) can only decrease—this is the energy certificate for safety. In the stochastic extension, the passivity inequality would be replaced with weak passivity [48].

Table 2. Thermodynamic interpretation of PH terms.

PH Term	Thermodynamic Role	Kolmogorov Cost
$J\nabla H$ (symplectic)	Reversible inference; reorganizes beliefs	Zero Landauer cost
$-R\nabla_p H$ (dissipation)	Irreversible forgetting; erases outdated beliefs	$k_B T \ln 2$ per erased bit
$G_\xi u_\xi$ (control port)	Observation intake; new data enters the system	Information gain $\leq K(u)$
ΣdW_t (noise port)	Environmental stochasticity	Irreducible uncertainty

The three knowledge regimes introduced by PHAST [49] connect to Kolmogorov cost:

- KNOWN: $V(q), M(q), D(q)$ are given; low $K(\xi)$; Casimir-protectable.
- PARTIAL: some structure is given, some is learned; moderate $K(\xi)$.
- UNKNOWN: everything is learned from data; high $K(\xi)$; full Landauer cost.

The meta-optimization over environments reads

$$\min_\eta \mathbb{E}_{\mathcal{E} \sim P} [\mathcal{J}(\eta; \mathcal{E})] + \beta \Omega_{\text{struct}}(\eta) + \gamma \Omega_{\text{life}}(\eta), \tag{25}$$

where the OCP functional is

$$\mathcal{J}(\eta; \mathcal{E}) = \mathbb{E}_{x_0 \sim \mu_{\mathcal{E}}} \left[\int_0^T \ell(x_{\mathcal{E}}^{\pi_\eta(\mathcal{E})}(t), \mathcal{E}) dt + \phi(x_{\mathcal{E}}^{\pi_\eta(\mathcal{E})}(T), \mathcal{E}) \right]. \tag{26}$$

4.2. Running Toy Example: A Planar Double Integrator with a Moving Obstacle

The running example is a planar double integrator with state $x = (q, p)$, position $q \in \mathbb{R}^2$, momentum $p \in \mathbb{R}^2$, and goal $q_g \in \mathbb{R}^2$. At curriculum stage τ , a circular obstacle has center $c(\tau)$ and radius $r(\tau)$.

Example 1 (Moving-obstacle PH controller). *Let*

$$H(q, p; \kappa) = \frac{1}{2} p^\top M^{-1} p + \frac{\omega_g}{2} \|q - q_g\|^2 - \alpha \log g_\varepsilon(q; \kappa), \tag{27}$$

with

$$g_\varepsilon(q; \kappa) := \|q - c(\tau)\|^2 - r(\tau)^2 + \varepsilon, \quad \kappa(\tau) := (c(\tau), r(\tau)). \tag{28}$$

The barrier is smoothed by $\varepsilon > 0$ to avoid singular stiffness. The plant uses the canonical PH structure

$$J_{\text{plant}} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}, \quad R_{\text{plant}}(q) = \text{diag}(0, 0, \beta_{\parallel}, \beta_{\perp}(q)),$$

where $\beta_{\perp}(q)$ increases near the obstacle in the normal direction. Because the canonical double integrator has no nontrivial Casimirs, the Casimir lock is implemented on the controller side by augmenting the interconnection with a memory coordinate z whose bracket is chosen so that a momentum-budget quantity remains invariant. Hard invariants need not belong to the physical plant alone; they may live in the learning/control interconnection.

The example is reused in Sections 5 and 6. The obstacle moves slowly, the barrier reshapes accordingly, and the learner warm starts from the previous stage rather than relearning from scratch.

4.3. Safety Hierarchy: Casimir, Barrier, Dissipation

Safety in the perennial-learning context requires three nested guarantees:

- (1) Instantaneous safety: $x(t) \in \mathcal{F}(\kappa(t))$ at every t .
- (2) Transitional safety: when \mathcal{F} contracts, the system reaches the new set without leaving the old one during transition.
- (3) Informational safety: the agent’s model of the boundary $\partial\mathcal{F}$ has low enough K to reliably distinguish safe from unsafe.

As shown in Figure 3, there are three safety mechanisms that operate at different time scales and with different update costs:

1. Casimir lock: a permanent structural invariant that should never be overwritten during normal operation.
2. Barrier shaping: a tunable potential that changes when the geometry of the safe set changes.
3. Dissipation shaping: a graded slowdown that governs how aggressively the state may approach the boundary.

We list out tradeoffs among three safety mechanisms in Table 3. For the toy example, the stage-wise effective feasible set is

$$\mathcal{F}_{\text{eff}}(\tau) := \{x \in \mathcal{X} : H(x; \kappa(\tau)) \leq E_{\text{max}}(\tau), C_j(x) = c_j \text{ for all locked invariants}\}. \tag{29}$$

Mechanism 1: Dynamic potential shaping

Decompose the potential as

$$V_{\tilde{\zeta}}(q; \mathcal{E}) = V_{\text{task}}(q; \mathcal{E}) + V_{\text{barrier}}(q; \kappa), \tag{30}$$

where the log-barrier enforces constraint feasibility:

$$V_{\text{barrier}}(q; \kappa) = -\alpha \sum_i \log(g_i(q; \kappa)). \tag{31}$$

As $q \rightarrow \partial\mathcal{F}$, $V_{\text{barrier}} \rightarrow \infty$. Passivity ($dH/dt \leq 0$) prevents the trajectory from gaining energy to climb the barrier.

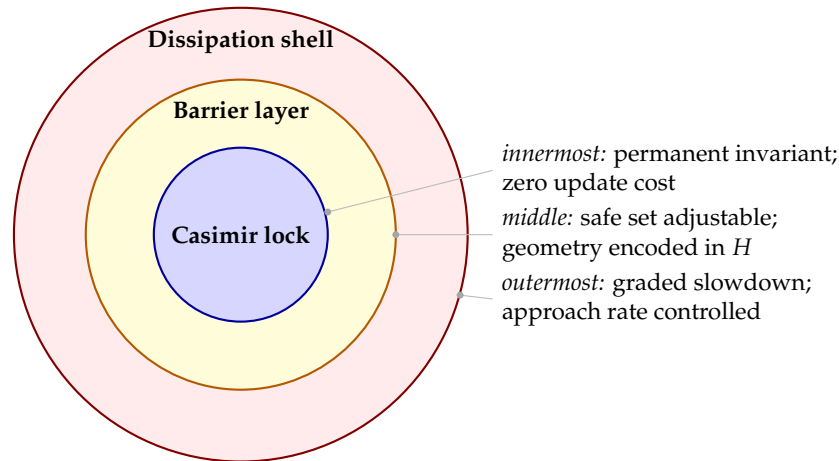


Figure 3. Schematic diagnostic, not empirical data. Safety hierarchy (Section 4.3). Permanent structure belongs in Casimirs, geometric changes belong in barriers, and approach-rate control belongs in dissipation. Combining all three into a single penalty obscures both safety and update cost.

Mechanism 2: Anisotropic dissipation shaping

$$R_{\zeta}(q, p; \mathcal{E}) = R_0(\mathcal{E}) + R_{\text{margin}}(q; \kappa), \tag{32}$$

where the margin dissipation increases near the boundary:

$$R_{\text{margin}}(q; \kappa) = \beta \sum_i \frac{1}{g_i(q; \kappa)^2} \nabla g_i \nabla g_i^{\top}. \tag{33}$$

This is *anisotropic*: strong damping normal to the constraint surface and weak damping tangentially. The trajectory can explore along the boundary (learning its geometry for identifiability) while being strongly damped against crossing it.

Mechanism 3: Casimir invariants as hard safety constraints

Casimir functions $C(x)$ satisfying $\{C, H_{\zeta}\} = 0$ (Definition 5) are preserved for *all* Hamiltonians, surviving potential reshaping, dissipation adjustment, and curriculum changes. From a Kolmogorov perspective, Casimirs incur zero marginal Landauer cost: once encoded in the Poisson-bracket structure, they require no ongoing memory, measurement, or erasure.

Table 3. Safety-mechanism tradeoffs.

Type	Enforcement	Landauer Cost	Use for
Casimir ($\{C, H\} = 0$)	Permanent, exact	Zero	Conservation laws, hard actuator limits
Barrier (V_{barrier})	Soft, adjustable	$K(\Delta V)$ per update	Moving obstacles, tightening clearances
Dissipation (R_{margin})	Graduated slowdown	Moderate (operating cost)	Safety margins, approach speed limits

Remark 8 (Informational cost of safety). *Updating the agent’s model of what is safe carries an overwrite cost. In a physical implementation, the irreversible component is lower-bounded by $\Delta Q_{\text{safety}} \geq k_B T \ln 2 \cdot K(\mathcal{F}(\kappa_{t+1}) \mid \mathcal{F}(\kappa_t))$ for the erased portion of the description. For abstract software updates, this expression should be read as a formal overwrite budget rather than as a literal device-level energy audit.*

4.4. Passivity Speed Limit for Constraint Tightening

The next proposition is a standard passivity/dissipation estimate tailored to curriculum-induced feasible-set contraction. Its novelty here does not lie in the inequality itself but in the interpretation of D_{max} as a rate limiter for safe curriculum scheduling.

Proposition 3 (Passivity speed limit). *Under Assumption 1, consider the trajectory in (3) satisfying the non-energy-extracting condition*

$$y(t)^\top u(t) \geq 0 \quad \text{for almost every } t \in [0, T]. \tag{34}$$

Suppose $x(0) \in \mathcal{X}_{E_0}$ with $H(x(0)) = E_0$, and any transition into the updated feasible set requires the energy to reach some level at most $E_1 < E_0$ while the trajectory remains inside \mathcal{X}_{E_0} . Then, every such transition satisfies

$$T \geq \frac{E_0 - E_1}{D_{\text{max}}(E_0)}. \tag{35}$$

The curriculum cannot tighten constraints faster than the system can dissipate energy.

Proof. By Proposition 1,

$$\dot{H}(x(t)) = -\nabla H^\top R \nabla H + y^\top u \geq -D_{\text{max}}(E_0),$$

where the inequality uses (34) and $D_{\text{max}}(E_0)$. Integrating from 0 to T , $E_1 - E_0 \geq -D_{\text{max}}(E_0)T$, which rearranges to (35). \square

Remark 9 (Interpretation and limits). *Proposition 3 is a worst-case lower bound. If an external controller is allowed to extract energy ($y^\top u < 0$), if \mathcal{X}_{E_0} is not compact, if $D_{\text{max}}(E_0)$ is not finite, or if stochastic excitation/exploration policies inject net power, the proposition no longer applies in the stated form. Section 6.4 gives a numerical stress test: the direct $A \rightarrow C$ obstacle jump has $\text{NCD}_Z = 0.265 > \delta_{\text{safe}} = 0.245$ and, when forced with a lagged barrier, produces negative clearance, whereas the admitted NCD-gated chain maintains positive clearance.*

4.5. Coverage, Identifiability, and Entropy Reduction Rate

The occupation measure $\rho_{\eta, \mathcal{E}} = \frac{1}{T} \int_0^T \delta_z(t) dt$ determines which microstates have their $K(x)$ reduced by observation. The coverage discrepancy is

$$\text{Cov}(\eta; \mathcal{E}) = D(\rho_{\eta, \mathcal{E}}, \nu_{\mathcal{E}}), \tag{36}$$

where the target measure $\nu_{\mathcal{E}}$ weights regions of high S_A (i.e., regions with much to discover), where $-dS_A/dt$ is achievable (i.e., where the microstate is compressible rather than random).

The Fisher/observability proxy for algorithmic-entropy reduction is (using the same symbol conventions as Section 5.1)

$$\mathcal{I}_T(\xi) := \int_0^T (\partial_\xi h_\xi(x(t)))^\top \Gamma_y^{-1} (\partial_\xi h_\xi(x(t))) dt. \tag{37}$$

Claim 2 (Fisher proxy for local posterior contraction). Let $\lambda_{\min}(\mathcal{I}_T(\xi))$ denote the smallest eigenvalue of the accumulated Fisher information matrix defined in (37). Large $\lambda_{\min}(\mathcal{I}_T(\xi))$ suggests that observations are informative in the weakest identified direction and therefore that the local posterior codelength can contract efficiently along the trajectory.

Note that Claim 2 is a local design claim, not a theorem about exact algorithmic entropy S_A . Proposition 4 (Section 5.1) gives the qualified Laplace/MDL version under Assumption 2.

Remark 10 (Safety–identifiability alignment). Near constraint boundaries, dynamics are typically more structured (the boundary imposes regularity), so $-\text{d}S_A/\text{d}t$ is higher near boundaries than in the interior. The most informative regions for identifiability are thus exactly where the agent most needs to learn constraint geometry. The spatio-temporal dissipation shaping (33) ensures a slow, safe approach. This alignment is task-dependent rather than universal (see Remark 12).

5. Identifiability, Curriculum, and Sequential Complexity

This section focuses on the Fisher-information claim and the measurement-complexity claim.

5.1. Fisher Information as a Local Codelength Proxy

For a trajectory $x(\cdot)$ and using the observation model from Assumption 2, the accumulated Fisher information is

$$\mathcal{I}_T(\xi) := \int_0^T (\partial_{\xi} h_{\xi}(x(t)))^{\top} \Gamma_y^{-1} (\partial_{\xi} h_{\xi}(x(t))) \text{d}t. \tag{38}$$

To avoid singularities, we work with the ridge-stabilized matrix $\mathcal{I}_T^{\rho} := \mathcal{I}_T + \rho I$ for any $\rho > 0$.

Proposition 4 (Local Fisher proxy for posterior codelength contraction). Under Assumption 2, let the local negative log posterior around the current iterate be approximated by

$$-\log p(\xi \mid \text{data}_{1:T}) = \text{const} + \frac{1}{2} (\xi - \hat{\xi}_T)^{\top} \mathcal{I}_T^{\rho} (\xi - \hat{\xi}_T) + o(\|\xi - \hat{\xi}_T\|^2).$$

Define the local uncertainty-codelength score $U_T^{\rho} := \frac{1}{2} \log \det(\mathcal{I}_T^{\rho})$. Then,

$$\frac{d_{\xi}}{2} \log(\lambda_{\min}(\mathcal{I}_T) + \rho) \leq U_T^{\rho} \leq \frac{d_{\xi}}{2} \log(\lambda_{\max}(\mathcal{I}_T) + \rho). \tag{39}$$

Therefore, maximizing $\lambda_{\min}(\mathcal{I}_T)$ increases a worst-direction lower bound on local posterior codelength contraction.

Proof. Let $\lambda_1, \dots, \lambda_{d_{\xi}}$ be the eigenvalues of \mathcal{I}_T . Then, $U_T^{\rho} = \frac{1}{2} \sum_i \log(\lambda_i + \rho)$. Since $\lambda_{\min} \leq \lambda_i \leq \lambda_{\max}$,

$$\frac{d_{\xi}}{2} \log(\lambda_{\min} + \rho) \leq U_T^{\rho} \leq \frac{d_{\xi}}{2} \log(\lambda_{\max} + \rho). \quad \square$$

Remark 11 (Status of the Fisher proxy). Proposition 4 is not a theorem about exact algorithmic entropy S_A or exact Kolmogorov complexity. Under a Laplace approximation, the posterior codelength contains a curvature term $\frac{1}{2} \log \det(\mathcal{I}_T^{\rho})$ plus lower-order constants, so Fisher information is used only as a local statistical codelength proxy. It can fail for multimodal posteriors, sensor aliasing, unobserved boundary parameters, severe model misspecification, or directions in which curvature is large but operationally irrelevant. In the scheduler, it is therefore paired with residual checks and NCD change rather than used as a standalone certificate.

Assumption 5 (Boundary-informative sensing). *There exists a neighborhood $\mathcal{N}_r(\partial\mathcal{F})$ of the constraint boundary and a constant $\sigma_b > 0$ such that the singular values of the Jacobian directions associated with the boundary parameters satisfy*

$$\sigma_{\min}(\partial_{\xi} h_{\xi}(x)) \geq \sigma_b \quad \text{for all } x \in \mathcal{N}_r(\partial\mathcal{F}).$$

Remark 12 (When safety and identifiability align). *Under Assumption 5, dwelling near the boundary can improve $\lambda_{\min}(\mathcal{L}_T)$ precisely in the directions that define the safe-set geometry. Without this assumption, the alignment can fail; there are perfectly safe problems whose boundaries are flat, hidden, or sensor-poor. Safety identifiability alignment is task-dependent rather than universal.*

5.2. Curriculum Scheduling via NCD and NID

Notations used in this section

- $\mathcal{F}_{\text{eff}}(\tau)$: the effective feasible set at curriculum time τ , defined in (29). Always written with the time argument; bare $\mathcal{F}(\kappa)$ refers to the abstract set parameterized by κ , and \mathcal{F}^t (without subscript eff) is not used.
- $\delta_{\text{safe}}(\tau) \in [0, 1]$: the admissible NCD threshold at step τ , derived from the Landauer budget B_{τ} and defined in (42).
- $L_Z(y | x) := L_Z(xy) - L_Z(x)$: practical conditional code length (not the algorithmic conditional $K(y | x)$); defined precisely in (40).
- Δ_t : total information change in bits at stage t (Proposition 5).
- Δk_t : number of materially changed model coordinates at stage t (Assumption 4). These two quantities are distinct: $\Delta_t = O(\Delta k_t \log d_{\xi} / \lambda_{*})$ under the warm-start assumptions.

Implementation and sensitivity.

We compute L_Z with a fixed, lossless compressor Z . Unless otherwise stated, Z is LZMA/XZ with preset 6 applied to a canonical byte string. Each environment is serialized as sorted-key JSON metadata, followed by ordered, quantized obstacle-boundary samples and, for grid tasks, row-major signed-distance samples. All strings are encoded as UTF-8, and all floating-point parameters are quantized before serialization. The practical rule uses

$$\text{NCD}_Z(x, y) = \frac{Z(xy) - \min\{Z(x), Z(y)\}}{\max\{Z(x), Z(y)\}},$$

and Appendix A reports sensitivity to gzip/zlib level 6 and to a parameter-vector encoding.

The core problem

A perennial learner does not encounter all scenarios at once; it moves through them in sequence. The order and pace of that sequence is the curriculum. A poorly ordered curriculum is susceptible to failure in two distinct ways. First, if consecutive tasks are too similar, the agent wastes update budget re-encoding information it already holds. Second, if they are too different, the agent cannot absorb the changes within one dissipation interval—the feasible set \mathcal{F}_{eff} would need to jump discontinuously, which the passivity speed limit (Proposition 3) forbids. Good curriculum design lives in the space between these two failure modes.

Compression distance as a curriculum metric

The key insight is that two tasks are informationally close if knowing one makes the other easy to describe. This is precisely what the normalized compression distance NCD_Z (Definition 14, (22)) captures, in the spirit of the similarity-metric and clustering-by-compression program of Li et al. [3], Cilibrasi and Vitányi [5]: if the concatenation $Z(\mathcal{E}_t \mathcal{E}_{t+1})$ is barely larger than $Z(\mathcal{E}_t)$ alone, then \mathcal{E}_{t+1} is almost entirely predictable from

\mathcal{E}_t —the transition carries little new information and is safe to absorb. Conversely, a large NCD_Z signals that the new task contains a structure the model has never seen, and the agent must allocate additional update budget before accepting the step.

This compression-based view has the following suggestive thermodynamic reading. Define the practical conditional code length

$$L_Z(y | x) := L_Z(xy) - L_Z(x), \tag{40}$$

the extra bits needed to describe y given x as a side-channel. On the visited serialized family from Assumption 3, $L_Z(\mathcal{E}_{t+1} | \mathcal{E}_t)$ locally approximates $K(\mathcal{E}_{t+1} | \mathcal{E}_t)$, so it serves as a practical overwrite proxy for the curriculum step. This does not make NCD_Z a universal learning-theoretic metric: it remains compressor- and representation-dependent and must be validated empirically against the transfer cost or safety violations in the task family.

The budget constraint

Let $\delta_{\text{safe}}(\tau)$ denote the fraction of the current code budget available in one curriculum interval, derived from the dissipation capacity $D_{\text{max}}(E_0)$ and the Landauer rate via Proposition 3. A curriculum step is admissible only if its information content fits within that budget. The ideal (uncomputable) condition uses the NID:

$$e(\mathcal{F}(\kappa_\tau), \mathcal{F}(\kappa_{\tau+1})) \leq \delta_{\text{safe}}(\tau), \tag{41}$$

and the practical condition replaces the NID with its computable NCD_Z surrogate (justified by Assumption 3):

$$\text{NCD}_Z(\mathcal{E}_t, \mathcal{E}_{t+1}) \leq \delta_{\text{safe}}(t). \tag{42}$$

Concretely, if the dissipation integral over one step can absorb at most B_t bits of erasure work, then $\delta_{\text{safe}}(t) = B_t/Z(\mathcal{E}_t)$, so the threshold shrinks as tasks grow more complex and grows as the agent builds dissipative capacity. A step that fails the screen is not forbidden; the scheduler instead inserts intermediate environments until the chain of steps is individually admissible.

Example 2 (NCD screening on the moving-obstacle curriculum). *Encode each environment as a structured description (dynamics, goal, sensor model, obstacle parameters). The compressed length $Z(\mathcal{E})$ is a proxy for the model’s minimum description length under that scenario. For a nearby transition $\mathcal{E}_A \rightarrow \mathcal{E}_B$, where the obstacle moves by two grid cells,*

$$Z(\mathcal{E}_A) = 222, \quad Z(\mathcal{E}_B) = 222, \quad Z(\mathcal{E}_A\mathcal{E}_B) = 232,$$

so $\text{NCD}_Z(\mathcal{E}_A, \mathcal{E}_B) = (232 - 222)/222 \approx 0.045$. *The small numerator (10 bytes) reflects the fact that once \mathcal{E}_A is known, \mathcal{E}_B requires almost no additional information to specify—only the new obstacle center.*

A more disruptive change $\mathcal{E}_A \rightarrow \mathcal{E}_C$ (obstacle doubles in size and relocates) yields $Z(\mathcal{E}_C) = 221, Z(\mathcal{E}_A\mathcal{E}_C) = 247$, so $\text{NCD}_Z(\mathcal{E}_A, \mathcal{E}_C) \approx 0.117$. Here, the numerator (25 bytes) reflects that \mathcal{E}_C contains substantial new geometry the compressor cannot predict from \mathcal{E}_A .

If the per-step Landauer budget corresponds to 18 new compressed bytes, then $\delta_{\text{safe}} = 18/222 \approx 0.081$. The first transition ($0.045 < 0.081$) is admitted; the second ($0.117 > 0.081$) is not, and the scheduler must find an intermediate scene $\mathcal{E}_{A'}$ with $\text{NCD}_Z(\mathcal{E}_A, \mathcal{E}_{A'}) \leq 0.081$ and $\text{NCD}_Z(\mathcal{E}_{A'}, \mathcal{E}_C) \leq 0.081$ before the full jump is attempted.

Curriculum Lifecycle

The four phases shown in Figure 4 describe not just geometric states but distinct information-theoretic regimes, each with its own dominant cost and its own trigger for the transition to the next phase.

1. Early phase (Phase 1 in Figure 4). The agent has limited knowledge of obstacle geometry, so it maintains a large, conservative \mathcal{F}_{eff} that includes a generous safety margin. The feasibility set is over-sized: the agent trades performance for safety by staying well away from the true boundary. Algorithmically, entropy \widehat{S}_A is high (many microstates consistent with the agent's current model), and NCD_Z between consecutive observations is small because nothing surprising is happening. Transition trigger: once the Fisher information $\lambda_{\min}(\mathcal{I}_T^0)$ rises above λ_* (Assumption 4), the boundary geometry is sufficiently identified and Phase 2 begins.
2. Learning phase (Phase 2). With reliable boundary estimates, the agent begins tightening \mathcal{F}_{eff} —increasing barrier steepness (the scale α in the log-barrier potential (31)) and raising boundary dissipation—so that \mathcal{F}_{eff} contracts toward the true obstacle contour. Each contraction step satisfies (42): the model description shortens as uncertainty collapses, so \widehat{S}_A falls. The Landauer cost per step is low because the agent is mostly deleting uncertainty (erasing microstates it has ruled out) rather than encoding a genuinely new structure. Transition trigger: \widehat{S}_A stops decreasing, signaling that further tightening requires new observations rather than inference.
3. Constraint shift (Phase 3). An external event changes the true constraint—in the double-integrator example, the obstacle moves (red arrow in Figure 4). This injects new information: the old model of obstacle geometry is partially wrong, so $\text{NCD}_Z(\mathcal{E}_t, \mathcal{E}_{t+1})$ spikes. The scheduler checks (42): if the spike exceeds δ_{safe} , the transition is broken into smaller steps. Even so, the agent must pay the Landauer cost $\widehat{\Delta K}_t \approx L_Z(\mathcal{F}_{\text{eff}}(t+1) \mid \mathcal{F}_{\text{eff}}(t))$ (using the conditional code notation (40)) for each bit of old geometry it overwrites, and the rate of deformation of \mathcal{F}_{eff} is bounded by Proposition 3. This is the phase in which the thermodynamic cost is highest, and the safety of curriculum ordering matters most.
4. Steady state (Phase 4). The agent has re-adapted to the shifted obstacle and now tracks slow constraint drift with the amortized cost $n_{\text{seq}} = \widetilde{O}(\Delta k_t / \lambda_*)$ measurements per update (Corollary 1), conditional on the warm-start assumptions of Proposition 5, where Δk_t is the number of parameters that have changed. The Casimir lock ensures that the invariants established in Phase 2 (momentum-budget conservation, obstacle-clearance structure) are not inadvertently erased during re-adaptation. \mathcal{F}_{eff} oscillates gently around the moving boundary rather than making discrete jumps. Return trigger: if NCD_Z spikes again above δ_{safe} , the lifecycle re-enters Phase 3.

Remark 13 (Why order matters even when all tasks are eventually visited). *A common objection is that if the agent will eventually see all environments anyway, curriculum ordering is merely a performance optimization. This misses the thermodynamic point: a badly ordered curriculum forces the agent into Phase 3 prematurely and repeatedly, incurring Landauer costs proportional to the total variation in the constraint trajectory rather than its net displacement. A well-ordered curriculum—one in which NCD_Z between consecutive steps stays below δ_{safe} —keeps the agent near Phase 2 or Phase 4, where the erasure cost is amortized over many steps rather than concentrated in sudden rewrites [50].*

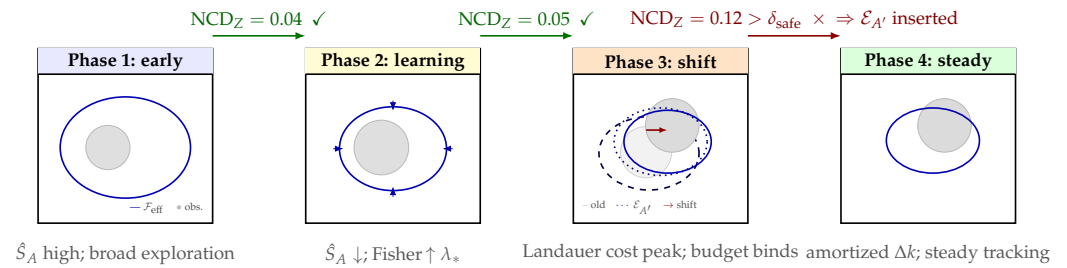


Figure 4. Schematic diagnostic of curriculum lifecycle discussed in Section 5.2. Each panel shows $\mathcal{F}_{\text{eff}}(\tau)$ (solid blue ellipse) and the obstacle (gray disc). Transition arrows above the panels: green \checkmark = step admitted ($\text{NCD}_Z \leq \delta_{\text{safe}}$); red \times = screened, intermediate $\mathcal{E}_{A'}$ inserted. Inward arrows in Phase 2 show \mathcal{F}_{eff} contracting as the boundary geometry is learned. In Phase 3, the dashed ellipse is the ghost of the old \mathcal{F}_{eff} ; the dotted ellipse is $\mathcal{E}_{A'}$; and the red arrow is the obstacle shift. Cost annotations below each panel indicate the dominant information-theoretic regime. Key invariant: the solid blue ellipse deforms continuously—it never jumps.

5.3. Sequential Information Budget

Proposition 5 (Sequential information budget). *Let Δ_t denote the conditional description change that must be learned at stage t , measured in bits by a practical code such as $L_Z(\xi_t | \xi_{t-1})$. Suppose each newly acquired measurement contributes at most b_{eff} bits of conditional mutual information about the changed component $I(\xi_t; Y_i | \xi_{t-1}, \mathcal{H}_{t-1}) \leq b_{\text{eff}}$ for each i . Then any sequential learner that resolves the changed description must use at least*

$$n_t \geq \frac{\Delta_t}{b_{\text{eff}}}. \tag{43}$$

If, in addition, Assumption 4 holds and the estimator is restricted to the active block S_t , then a warm-start local estimator can achieve

$$n_t = O\left(\frac{\Delta k_t \log d_{\xi} + \log(1/\delta)}{\lambda_*}\right) \tag{44}$$

for confidence level $1 - \delta$.

Proof. The lower bound follows from the following chain rule: $I(\xi_t; Y_{1:n_t} | \xi_{t-1}, \mathcal{H}_{t-1}) \leq n_t b_{\text{eff}}$. To resolve Δ_t bits of changed description, the left-hand side must be at least Δ_t , yielding (43).

Relationship between Δ_t and Δk_t . Δ_t is the total information content in bits that must be acquired; Δk_t is the number of model coordinates that have changed. Under the warm-start sparsity structure of Assumption 4, each of the Δk_t changed coordinates contributes at most $O(\log d_{\xi} / \lambda_*)$ bits of local identification cost, so $\Delta_t = O(\Delta k_t \log d_{\xi} / \lambda_*)$, and the lower bound (43) is matched (up to constants) by the upper bound (44). The upper bound is the standard local-identification scaling on the active block under the Fisher lower bound λ_* ; only Δk_t coordinates need to be relearned. \square

Corollary 1 (Comparison with a cubic cold-start baseline). *Suppose a chosen cold-start identification pipeline obeys $N_{\text{static}}(k, d_{\xi}) = \Theta(k^3 \log d_{\xi})$. Under the assumptions of Proposition 5, the warm-start sequential stage complexity obeys $N_{\text{seq}}(t) = O(\Delta k_t \log d_{\xi} / \lambda_*)$. Hence, the “ $k^3 \rightarrow \Delta k$ ” improvement is valid only relative to that cold-start baseline and only when $\Delta k_t \ll k_t^3$.*

Example 3 (Numerical instance of the conditional improvement). *Take $d_{\xi} = 256$, cold-start active complexity $k_t = 8$, changed block size $\Delta k_t = 2$, and Fisher lower bound $\lambda_* = 0.9$. A cold-start method has $N_{\text{static}} = 0.5 k_t^3 \log d_{\xi} \approx 1419$. A warm-start local estimator with*

$N_{\text{seq}} \approx 4\Delta k_t \log d_{\xi} / \lambda_*$ requires only about 49 measurements. If the drift rises to $\Delta k_t = 6$, the same formula yields about 148 measurements: the budget grows with drift rather than with total task size.

5.4. Where Differential Policy Optimization (DPO) Fits

Differential Policy Optimization (DPO) is an external transport-based optimal control solver whose convergence and regret analysis are obtained from [42,51]. Its attraction here is conceptual: pointwise transport updates align naturally with the reversible branch of the PH decomposition and with Pontryagin-style trajectory optimization [52]. This manuscript does *not* prove that DPO is the unique optimal steering law for the PH learner, nor does it rederive the external $\mathcal{O}(K^{5/6})$ regret bound.

Theorem 5 (DPO pointwise convergence [42]). *Let $G_{\theta_k}^{(j)}$ denote the learned transport operator at stage j and step k , and let $G^{*(j)}$ denote the optimal operator. Then,*

$$\mathbb{E} \|G_{\theta_k}^{(j)}(X) - G^{*(j)}(X)\| < jL^j \frac{\epsilon}{L - 1}, \tag{45}$$

where L is the Lipschitz constant and ϵ is the per-stage error.

Thermodynamic interpretation of the regret bound

The regret decomposes into:

- Inference regret: excess cost from suboptimal transport, i.e., $K^t(x) - K(x)$ from the speed–dissipation tradeoff (Remark 4).
- Safety regret: excess cost from unnecessary conservatism (staying too far from boundaries before learning constraint geometry).

As $K \rightarrow \infty$, both components decrease. Total excess Landauer dissipation per episode: $\mathcal{O}(K^{5/6})$.

Stochastic extension

The rough-path extension of DPO [51] handles the noise port $\Sigma \circ dW_t$ in the stochastic PH system (24).

State-constrained DPO for safety

Standard DPO applies Pontryagin without state constraints [53]. For safe perennial learning, the constraint $x(t) \in \mathcal{F}_{\text{eff}}(\tau)$ is incorporated via the PH barrier potential V_{barrier} (31) and boundary dissipation R_{margin} (33), converting the constrained problem into an unconstrained one with modified Hamiltonians. The co-state (adjoint) variable becomes the shadow price of constraint proximity.

Bilevel optimization

The coupled bilevel optimization governing the curriculum reads

$$\begin{aligned} \min_{\eta} \quad & \mathbb{E}_{\tau} [\mathcal{J}(\eta; \mathcal{E}(\tau))] \quad \text{s.t.} \quad \kappa(\tau) = C_{\psi}(\mathcal{H}_{\tau}), x(t) \in \mathcal{F}_{\text{eff}}(\tau) \forall t, \\ \max_{\psi} \quad & \mathbb{E}_{\tau} [\lambda_{\min}(\mathcal{I}_T(\xi; \mathcal{E}(\tau))) - \alpha \text{Cov}(\eta; \mathcal{E}(\tau))] \quad \text{s.t.} \quad \dot{\kappa}(\tau) \leq \kappa_{\max}(\tau), \end{aligned} \tag{46}$$

where $\dot{\kappa}(\tau) \leq \kappa_{\max}(\tau)$ is the passivity-derived speed limit (Proposition 3). Here, ψ denotes the curriculum-design parameters (the outer optimization variable governing how the curriculum sequence is shaped). This is distinct from $\psi_{\mathcal{F}_i}(x)$, which is the scalar safety-barrier/margin term appearing in the running cost (12); the two ψ -objects share a letter but play different roles and are always distinguished by their arguments.

Remark 14. [Algorithmic temperature and curriculum difficulty] Following Baez and Stay [54], the “algorithmic temperature” is the cost of doubling the number of programs. In the curriculum context, this maps to scenario difficulty: “hotter” scenarios (more stochastic, wider constraint sets) require more erasure work to compress. The curriculum implicitly schedules the temperature of the environments the demon faces [41].

6. Numerical Realization

The theoretical statements above are continuous-time. An implementation needs a discretization that does not destroy the passivity structure the theory relies on.

6.1. Discretization, Passivity Margin, and Stiffness Handling

We recommend a splitting integrator with three substeps:

1. Conservative step. Integrate $\dot{x} = J(x)\nabla H(x)$ using a symplectic method such as Störmer–Verlet or symplectic Euler [55].
2. Dissipative/barrier step. Integrate $\dot{x} = -R(x)\nabla H(x) + G(x)u$ with a discrete-gradient or average-vector-field method [56], so that a discrete passivity inequality is preserved at the substep level.
3. Stochastic step. Integrate the Stratonovich noise using a midpoint/Heun-type method compatible with the chain rule [47].

For the dissipative substep, a discrete-gradient update yields

$$H(x_{n+1}) - H(x_n) \leq h y_n^\top u_n - h \bar{\nabla} H^\top R(\bar{x}_n) \bar{\nabla} H, \quad (47)$$

where $\bar{\nabla} H$ is a discrete gradient and h is the time step. A Strang split preserves this inequality up to $O(h^3)$ local and $O(h^2)$ global error when the subflows are solved to matching order [49].

Remark 15 (Barrier stiffness). *The ideal log-barrier $-\alpha \log(g; \kappa)$ becomes stiff as $g \downarrow 0$. In numerics, we use the smoothed barrier $-\alpha \log(g + \epsilon)$ from Example 1, together with adaptive stepping or an implicit dissipative substep. The theory remains stated for the ideal barrier; the implementation uses its regularized counterpart.*

Remark 16 (Why Stratonovich rather than Itô?). *The Stratonovich formulation is adopted because it obeys the ordinary chain rule and respects coordinate changes on geometric state spaces, which is natural for PH systems. If one rewrites in Itô form, an additional drift correction $\frac{1}{2} \Sigma_j (\partial_x \Sigma_j) \Sigma_j$ appears [57]. Stratonovich models the continuous-time limit of rapidly varying physical perturbations more faithfully when PH geometry matters.*

6.2. Algorithmic Template

We provide Algorithm 1 as a general perennial PH learning algorithm under budgeted curriculum updates. The policy update is realized by a time integration step rather than a transition function and NCD is used to compute the budget and guide the learned model sequence do adapt or maintain its current state.

6.3. Validation Protocol

Constrained autonomy

The moving-obstacle double-integrator example tests whether the PH controller can reshape barriers, preserve controller-side invariants, and reuse previous information safely. Relevant metrics: violation rate, minimum distance to the obstacle boundary, energy ratio H/E_{\max} , and drift of locked invariants.

Algorithm 1 Perennial PH learning with code-budgeted curriculum updates

Input: prior model ζ_0 , initial environment encoding \mathcal{E}_0 , code proxy L_Z , dissipation budget schedule
for $t = 1, 2, \dots, T$ **do**
 encode proposed next environment \mathcal{E}_t and compute $\text{NCD}_Z(\mathcal{E}_{t-1}, \mathcal{E}_t)$
 if $\text{NCD}_Z(\mathcal{E}_{t-1}, \mathcal{E}_t) > \delta_{\text{safe}}(t)$ **then**
 replace \mathcal{E}_t by an intermediate environment satisfying the budget
 end if
 warm-start the solver/policy from ζ_{t-1} (DPO or another local transport solver)
 update barrier and dissipation parameters for $\mathcal{F}_{\text{eff}}(t)$
 integrate the PH dynamics by conservative/dissipative/stochastic splitting
 collect new measurements and update ζ_t
 record overwrite proxy $\widehat{\Delta K}_t = [L_Z(\zeta_t) - L_Z(\zeta_{t-1})]_+$
 monitor $\widehat{S}_A(t)$, $\lambda_{\min}(\mathcal{I}_t)$, and constraint-margin statistics
end for
Output: updated model sequence $\{\zeta_t\}_{t=0}^T$, safety and code-budget logs

Sequential data assimilation

The same framework applies when the “safe set” is a confidence region induced by sensors and physics constraints. Sensor dropout, new sensors, or changing observation operators generate the sequential-drift setting from Assumption 4; while ensemble Kalman and probabilistic transport formulations provide the practical baselines [58,59].

6.4. Numerical Case Study: NCD-Gated Moving-Obstacle Curriculum

We instantiate the running example as a planar double integrator with $q, p \in \mathbb{R}^2$, $\dot{q} = p$, and $\dot{p} = -\nabla_q V(q; \kappa) - D(q; \kappa)p$. The task is to move from $q_0 = (-1.15, -0.75)$ to $q_g = (1.15, 0.75)$ while avoiding a circular obstacle whose center and radius drift over curriculum time. The PH energy used by the controller is

$$H(q, p; \kappa) = \frac{1}{2} \|p\|^2 + \frac{\omega_g}{2} \|q - q_g\|^2 - \alpha \log(\|q - c\|^2 - r^2 + \varepsilon),$$

with anisotropic damping

$$D(q; \kappa) = \beta I + \frac{\beta_b}{(g(q; \kappa) + \varepsilon)^2} n(q)n(q)^\top, \quad g(q; \kappa) = \|q - c\|^2 - r^2, \quad n(q) = \frac{q - c}{\|q - c\|}.$$

The toy simulator in `curriculum_learning_simulation.py` uses $\alpha = 0.08$, $k_{\text{goal}} = 0.8$, $\beta = 0.8$, $\beta_b = 0.05$, $\varepsilon = 0.02$, $dt = 10^{-2}$, horizon 10, and acceleration clip 5.0. Figure 5 shows the NCD-gated curriculum from stages $\{A, B, A', C, D\}$, along with obstacle-avoidance stage trajectory and the stage obstacles.

Environment encoding and NCD gate

Each environment \mathcal{E}_t is serialized as a canonical byte string consisting of sorted-key JSON metadata, the dynamics label, goal, sensor model, obstacle center and radius, and ordered quantized samples of the obstacle boundary. The primary compressor is LZMA/XZ preset 6. In the toy run, the admissibility threshold is $\delta_{\text{safe}} = 0.245$. The direct shift $A \rightarrow C$ has $\text{NCD}_Z = 0.265$ and is rejected; the inserted chain $A \rightarrow B \rightarrow A' \rightarrow C \rightarrow D$ has admitted pairwise values of 0.236, 0.242, 0.237, and 0.210.

Diagnostics

We monitor four quantities: a practical algorithmic-entropy surrogate $\widehat{S}_A(t)$, a boundary-informative Fisher proxy, the NCD shock size between successive environments, and an

overwrite proxy proportional to the conditional compressed length. Decreasing $\hat{S}_A(t)$ together with an increasing Fisher score indicates that measurements are reducing uncertainty faster than the model description grows. Figure 6 plots the perennial-learning diagnostics during the curriculum for stages $\{A, B, A', C, D\}$ via four metrics for the toy obstacle-avoidance simulated example shown in Figure 5.

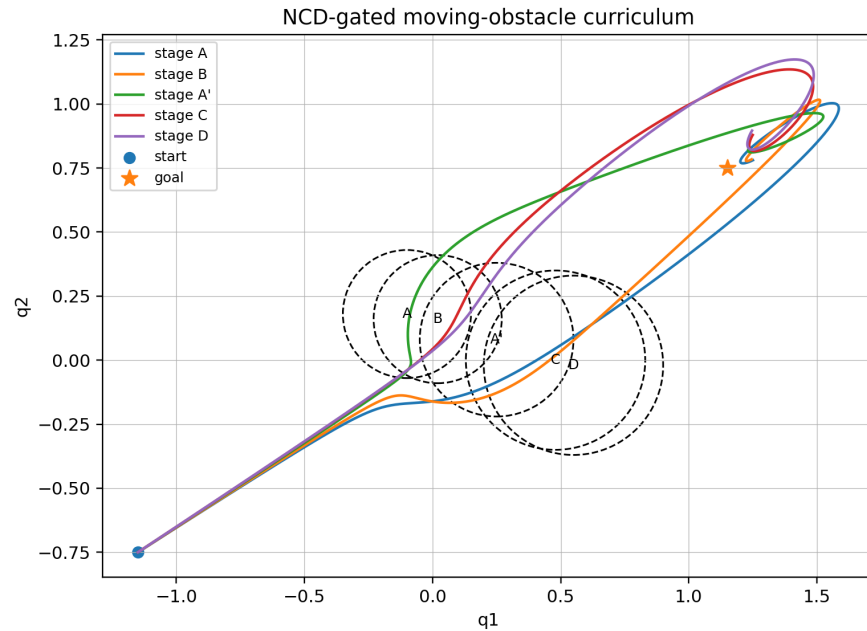


Figure 5. NCD-gated curriculum for the moving-obstacle double integrator. Each colored curve is a learned obstacle-avoidance stage trajectory from start to goal, and each dashed circle represents the corresponding obstacle at that stage.

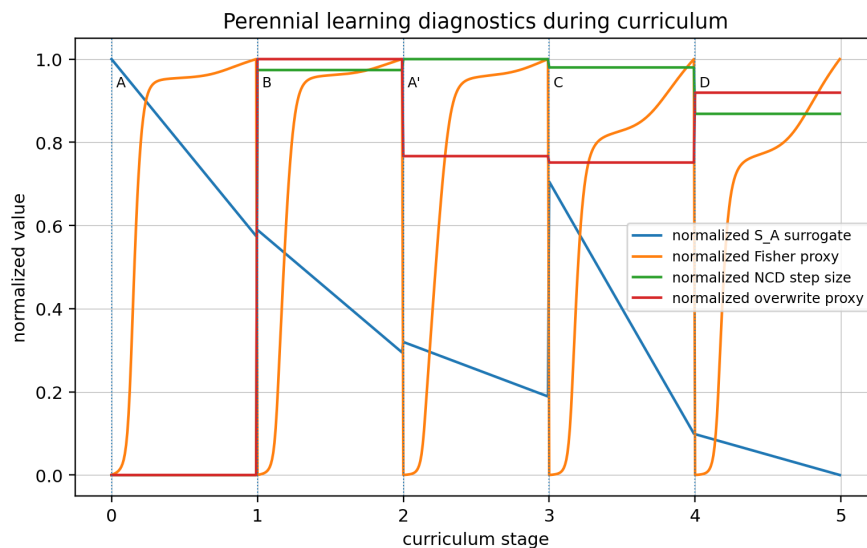


Figure 6. Perennial-learning diagnostics during the curriculum. The plot reports a normalized \hat{S}_A surrogate, a normalized Fisher proxy, a normalized NCD step size, and a normalized overwrite proxy over the five-stage curriculum.

Passivity stress test

To illustrate Proposition 3, we compare the NCD-gated update against a forced update in which the physical obstacle jumps from A to C while the controller barrier is still centered at A. Figure 7 plots the passivity speed-limit test of the toy example given above comparing four different modes. In the admitted curriculum the minimum clearance remains positive

(0.043 in normalized units). In the forced jump the minimum clearance is -0.234 , i.e., the true obstacle is penetrated before the lagged barrier can dissipate and redirect the trajectory. This should be read as a numerical illustration of the dissipation-bound interpretation, not as a universal benchmark.

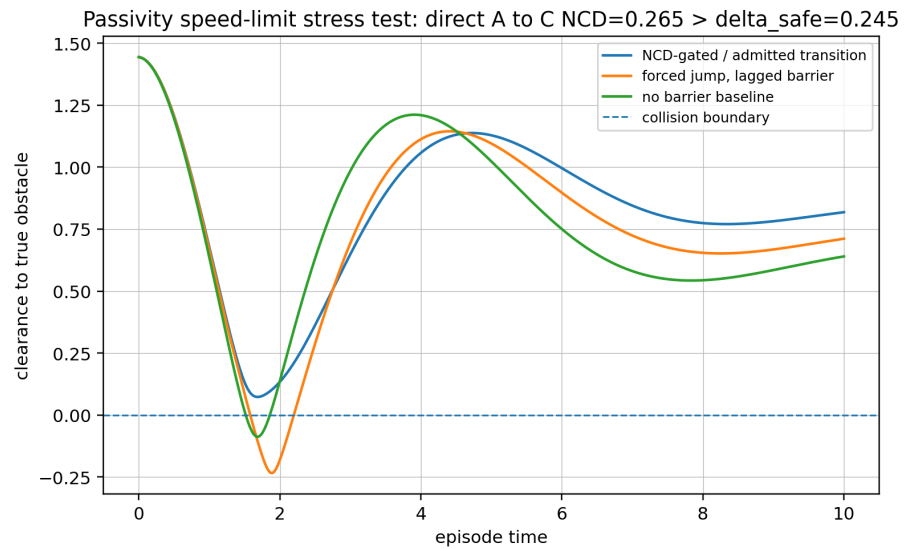


Figure 7. Passivity speed-limit stress test. The admitted transition maintains positive clearance, while the forced direct jump with a lagged barrier crosses the collision boundary.

Compact baseline summary

The toy simulator compares four modes: PH + NCD, forced jump with lagged barrier, no-barrier PH goal seeking, and a cold-start PH identification baseline. PH + NCD has zero violations in this run and uses the warm-start active-block sample budget. The forced jump violates the constraint in 6.2% of time steps; the no-barrier baseline violates in 3.4%; the cold-start baseline avoids collision but uses a much larger measurement budget under the chosen cubic cold-start pipeline. Figure 8 plots a toy benchmark summary for comparing the four modes. These values are reported as a reproducible toy case study, not as a claim of broad empirical superiority over continual-learning or safe-RL baselines.

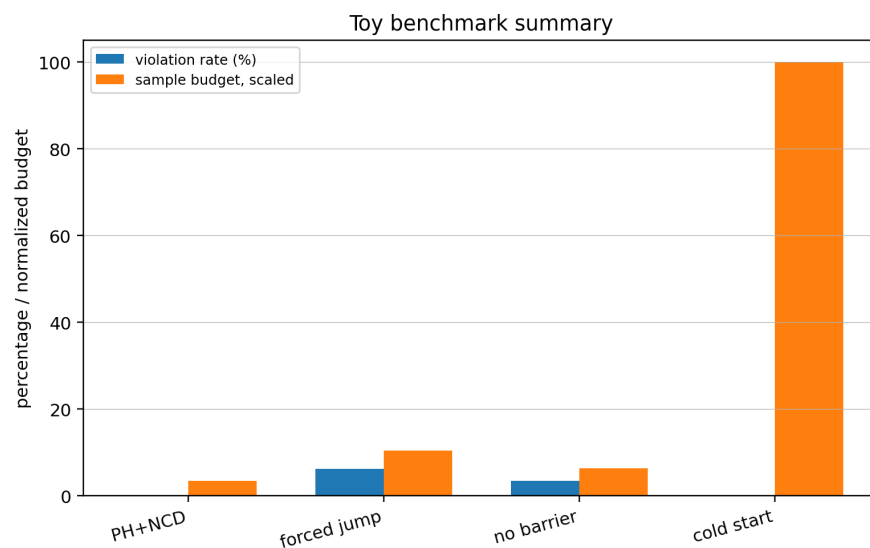


Figure 8. Toy benchmark summary. Sample budgets are scaled to the tallest bar; the exact numerical values are listed in Table 4.

Table 4. Quantitative metrics of a simulated moving-obstacle benchmark. We report the validation rate, minimum clearance of the path, mean path length (MPL), and number of measurements the program required for the computations.

Method	Violation %	Min. Clearance	MPL	Measurements	Comment
PH + NCD	0.00	0.043	3.771	49	Admitted chain $A \rightarrow B \rightarrow A' \rightarrow C \rightarrow D$
Forced jump	6.20	-0.234	3.921	148	Direct $A \rightarrow C$ exceeds δ_{safe} with lagged barrier
No barrier	3.40	-0.088	4.102	90	Goal controller without log-barrier
Cold start	0.00	0.063	4.073	1419	Cubic cold-start identification baseline

7. Integration and Validation

7.1. The Perennial Inference Engine: Full Assembly

We now assemble all components into a single operational cycle.

One inference step

1. Observation (port $G_{\xi} u_{\xi}$): new data arrives; the demon observes molecules.
2. Transport (DPO update via $J\nabla H$): the policy pushes the current posterior toward the updated posterior via the learned transport map. This is reversible (zero Landauer cost).
3. Dissipation (Landauer cost via R_{ξ}): outdated beliefs are overwritten; energy $k_B T \ln 2$ per erased bit flows through the dissipation port. This is irreversible.
4. Conservation (Casimir C): structural invariants—symmetries, sparsity patterns, discovered physical laws—are preserved automatically (no memory, no measurement, no Landauer cost).
5. Safety check (energy certificate): $H_{\xi}(x(t)) \leq E_{\text{max}}(\tau)$ is verified via passivity. If approaching the boundary, R_{margin} increases, slowing the trajectory.

Headline quantitative result

Under the assumptions of Proposition 5 and Corollary 1, and relative to a cold-start baseline requiring $n \propto k^3 \text{polylog } d$ measurements, the perennial PH + DPO engine achieves $n_{\text{seq}} = O(\Delta k_t \log d_{\xi} / \lambda_{\star})$ (equivalently, $\tilde{O}(\Delta k_t / \lambda_{\star})$) per step, conditional on the warm-start assumptions of Proposition 5, through the following mechanism:

$$\text{Casimir lock} \xrightarrow{\text{reduces phase-space volume}} \text{lower proposal variance} \xrightarrow{\text{DPO transport}} \text{proposal} \approx \text{posterior.}$$

Table 5 provides the key differences between a static learner and a perennial PH + DPO engine. This improvement is conditional on the warm-start assumptions and the chosen cold-start baseline.

Table 5. Static learner vs. perennial PH + DPO engine.

Component	Static Learner	Perennial PH + DPO
Proposal Q	Fixed product distribution	Learned transport μ_K via DPO
Search space	Discrete subset Ω_k	Casimir-locked phase space
Safety	Post-hoc constraint check	Energy certificate + barrier + dissipation
Scaling	$n = \Theta(k^3 \text{polylog } d)$ (global)	$n = \tilde{O}(\Delta k / \lambda_{\star})$ (warm-start, conditional)
Forgetting cost	Uncontrolled (catastrophic)	Budgeted by $\Omega_{\text{life}} \propto \Delta K$
Structure	None	Casimir invariants (zero Landauer cost)
Curriculum	None	NID/NCD-guided feasibility-set evolution

7.2. Numerical Validation Protocol and Broader Benchmark Agenda

Validation Class 1: Constrained autonomy and navigation

The toy moving-obstacle case study in Section 6.4 is the first validation class: it directly measures violation rate, minimum clearance, path length, overwrite proxy, Fisher score, and curriculum admissibility under actual simulated trajectories. A broader benchmark suite should extend this to multi-agent and multi-joint systems with progressively tightening safety/feasibility constraints [8]. The perennial agent must then handle multiple spatial scales simultaneously: local obstacle avoidance (1:1), neighborhood-level planning (1:20), district-level routing (1:400), and city-level strategy (1:80,000). Each scale has its own feasibility set $\mathcal{F}(\kappa)$, constraint complexity $K(\partial\mathcal{F})$, and identifiability challenges. The PH framework handles this via hierarchical energy shaping: Casimir invariants at the highest scale (road network topology, traffic laws), barrier potentials at intermediate scales (lane boundaries, intersection geometry), and dissipation shaping at the local scale (collision-avoidance margins).

Validation Class 2: Sequential inference and data assimilation

Changing observation operators (sensor failures, new sensors coming online) and tightening noise budgets [60,61] define the second validation class. The feasibility set then corresponds to posterior states consistent with the observation model; sensor failure expands the feasible posterior in an unexplored direction. In that regime, the relevant empirical baselines include replay-style continual-learning policies, EWC/GEM-style parameter protection, ensemble Kalman methods, and safe control and barrier-function baselines. We do not claim superiority over those families here; rather, the present paper supplies the PH bookkeeping and the toy case study that such a large benchmark should evaluate.

Safety-specific evaluation metrics

- (i) Constraint-violation frequency.
- (ii) Minimum distance to constraint boundary over trajectory.
- (iii) Energy-budget utilization $H(t)/E_{\max}$.
- (iv) Casimir-invariant drift $|C(x(t)) - C(x(0))|$.
- (v) Hausdorff distance between \mathcal{F}_{eff} and $\mathcal{F}(\kappa)$.

8. Discussion, Limitations, and Conclusions

We have developed a unified framework for perennial machine learning grounded in three classical pillars: Kolmogorov complexity (learning as compression), Landauer accounting (irreversible overwrite has a physical lower bound in concrete devices and a formal bookkeeping role in the abstract model), and port-Hamiltonian dynamics (the learning engine has structure). The Maxwell/Szilard analogy motivates the separation between reversible transport and irreversible overwrite, but the formal content of this paper consists of the explicit set of propositions and assumptions. The information-distance framework provides a computable geometry for measuring learning progress and designing safe curricula. Differential Policy Optimization supplies one transport law, imported from [42], with a regret bound interpretable as excess overwrite cost. The resulting perennial inference engine separates zero-cost reversible inference from costly irreversible forgetting by construction, enforces safety through energy certificates and Casimir invariants, and achieves amortized measurement complexity $n_{\text{seq}} = \tilde{O}(\Delta k \log d_{\tilde{\epsilon}}/\lambda_*)$ relative to a chosen cold-start baseline, conditional on warm-start initialization (Proposition 5 and Corollary 1).

8.1. Thermodynamic Learning by Lifting, Reduction, and PH Model Enrichment

A complementary thermodynamic interpretation of learning is complexity reduction through the recognition of a pattern in the PH phase portrait. In this view, the data are not only observations to be compressed but also a family of trajectories generated by a PH vector field. Learning occurs when this family is recognized as the projection of a lower-dimensional or more structured PH dynamics. This is common in nonequilibrium and rate thermodynamics: one often lifts the original variables to tangent, cotangent, or contact-type spaces, studies the lifted dynamics, and then identifies a reduced dynamics whose phase portrait captures the emergent pattern. The classical Lagrange multiplier construction can be read in this way: the constrained problem is lifted to an augmented space in which constraints become geometric variables, after which the reduced extremal dynamics reveals the effective law.

This perspective is consistent with the perennial-learning framework. The PH learner can either compress a phase portrait by reducing variables when a coherent pattern is discovered or enrich the model when the current coarse variables no longer close the dynamics. The latter is essential in thermodynamic modeling: a simple-fluid model may ignore microscopic configuration variables, whereas a complex-fluid model must promote molecular orientation, conformation, or internal stress into the state. In our notation, this enrichment may increase $K(\xi)$ or $L_Z(\xi)$ initially but can reduce the residual uncertainty enough to lower the total description length. Thus, reduction and enrichment are two sides of the same compression principle: reduce when the lifted phase portrait reveals a lower-order invariant and extend when omitted microscopic variables become macroscopically active. From the viewpoint of algorithmic statistics, this is also a model-selection question about which structured description best balances fit and complexity, in the spirit of Kolmogorov structure functions [4]. A full theory of automated PH lifting/reduction remains open.

8.2. Limitations

1. Proxy calibration is local and task-family dependent. Exact K is never optimized directly. Every implementable program statement uses L_Z or another explicit code proxy, and Assumption 3 is deliberately restricted to the finite serialized model family actually visited by the curriculum.
2. The Fisher codelength proxy is local. Proposition 4 is a Laplace/MDL-style curvature diagnostic; it can fail for multimodal, aliased, or strongly misspecified posteriors.
3. The $k^3 \rightarrow \Delta k$ comparison is conditional. Corollary 1 is relative to a chosen cold-start baseline and depends on warm start, active-block drift, and informative local curvature. It is not a universal law of perennial learning.
4. NCD admissibility is a screen, not a guarantee. Curriculum gating is compressor- and representation-dependent. It becomes a validated metric only after correlation with transfer cost, overwrite cost, or safety violations is measured in the task family.
5. Numerical stiffness and solver dependence remain practical constraints. Barriers near the boundary require regularization and careful time stepping, and DPO is one plausible transport solver rather than a theorem of necessity.

More broadly, the present framework also incorporates the algorithmic information-theoretic view of inductive learning and cognition: if learning is compression under structural constraints, then perennial adaptation can be considered a controlled form of inductive inference with memory and overwrite budgets [41,50].

8.3. Open Problems

1. Adaptive Casimir selection. Can the system learn which invariants to protect? If a constraint has not changed in N curriculum steps, $K(\kappa_{t+N} | \kappa_t) \approx 0$, does this suggest promotion to Casimir status?
2. Computable K proxies. Can we replace Kolmogorov complexity with practical compression-based surrogates (NCD, minimum description length) for the Landauer cost, Ω_{life} , and NID-based curriculum [3–5]?
3. The safety–identifiability frontier. Does the optimal boundary dissipation β minimize the total Landauer cost, balancing extra-measurement cost (poor identifiability at high β) against constraint-violation recovery cost (at low β)?
4. Algorithmic emergence detection. Can the agent detect phase transitions in its own learning dynamics—moments where S_A drops discontinuously, signaling qualitative discovery [44,62]?
5. Tighter regret under PH structure. Does the $\mathcal{O}(K^{5/6})$ bound tighten when DPO is applied to a system with known PH structure?
6. Quantum extension. How can we extend the framework to quantum inference, where K of pure quantum states replaces classical K ([32], §8.7)?

Funding: This work was additionally supported by gifts from the Peter O’Donnell Foundation (account #30212223) and the Jim Holland Foundation (account #30212279).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: Many thanks to Shubham Bhardwaj, Ryan Farell, Harsh Tirhekar, Luke McLellan, Minh Nguyen, Aditya Sai, and Yi Wang for staying the course with me as we continue to advance together using Port-Hamiltonian dynamics.

Conflicts of Interest: The author declare no conflicts of interest.

Appendix A. Reproducibility Details for the NCD-Gated Toy Case Study

Primary compressor

This manuscript uses LZMA/XZ preset 6 as the primary compressor for L_Z and NCD_Z in the toy study. Sensitivity checks use gzip/zlib level 6 and parameter-vector encoding.

Canonical serialization

Each environment encoding begins with sorted-key JSON metadata containing the dynamics label, mass, goal, sensor model, barrier type, damping type, grid size, obstacle center, obstacle radius, and stage label. This header is followed by repeated ordered text records for the obstacle geometry and ordered boundary samples. Strings are UTF-8 encoded, and floating-point values are quantized before serialization.

Toy simulation hyperparameters

The numerical study in Section 6.4 uses horizon 10.0, step size $dt = 0.01$, barrier weight $\alpha = 0.08$, goal-spring coefficient $k_{\text{goal}} = 0.8$, isotropic damping $\beta = 0.8$, margin damping $\beta_b = 0.05$, smoothing $\varepsilon = 0.02$, and acceleration clip 5.0. The curriculum scenes are

$$A = (-0.10, 0.18, 0.25), \quad B = (0.02, 0.16, 0.25), \quad A' = (0.25, 0.08, 0.30), \quad C = (0.48, 0.00, 0.35), \quad D = (0.55, -0.02, 0.35),$$

where each tuple has an obstacle center (c_1, c_2) and radius r .

Curriculum thresholds and outcomes

The toy run uses $\delta_{\text{safe}} = 0.245$. The admitted pairwise NCD values are 0.236, 0.242, 0.237, and 0.210, while the direct $A \rightarrow C$ jump has $\text{NCD}_Z = 0.265$ and is rejected by the

screen. The accompanying JSON file `toy_benchmark_results.json` records the exact table values used in Section 6.4.

References

- Li, M.; Vitányi, P.M. A new approach to formal language theory by Kolmogorov complexity. *SIAM J. Comput.* **1995**, *24*, 398–410. [[CrossRef](#)]
- Li, M.; Vitányi, P.M. Reversibility and adiabatic computation: Trading time and space for energy. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **1996**, *452*, 769–789.
- Li, M.; Chen, X.; Li, X.; Ma, B.; Vitányi, P.M.B. The Similarity Metric. *IEEE Trans. Inf. Theory* **2004**, *50*, 3250–3264. [[CrossRef](#)]
- Vereshchagin, N.K.; Vitányi, P.M. Kolmogorov’s structure functions and model selection. *IEEE Trans. Inf. Theory* **2004**, *50*, 3265–3290. [[CrossRef](#)]
- Cilibrasi, R.; Vitányi, P.M.B. Clustering by Compression. *IEEE Trans. Inf. Theory* **2005**, *51*, 1523–1545. [[CrossRef](#)]
- Parisi, G.I.; Kemker, R.; Part, J.L.; Kanan, C.; Wermter, S. Continual Lifelong Learning with Neural Networks: A Review. *Neural Netw.* **2019**, *113*, 54–71. [[CrossRef](#)]
- Thrun, S.; Mitchell, T.M. Lifelong Robot Learning. *Robot. Auton. Syst.* **1995**, *15*, 25–46. [[CrossRef](#)]
- Ellendula, A.S.; Wang, Y.; Nguyen, M.P.; Bajaj, C.L. GRL-SNAM: Geometric Reinforcement Learning with Differential Hamiltonians for Navigation and Mapping in Unknown Environments. In Proceedings of the Fourteenth International Conference on Learning Representations, Rio de Janeiro, Brazil, 23–26 April 2026.
- Kirkpatrick, J.; Pascanu, R.; Rabinowitz, N.; Veness, J.; Desjardins, G.; Rusu, A.A.; Milan, K.; Quan, J.; Ramalho, T.; Grabska-Barwinska, A.; et al. Overcoming Catastrophic Forgetting in Neural Networks. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 3521–3526. [[CrossRef](#)] [[PubMed](#)]
- Zenke, F.; Poole, B.; Ganguli, S. Continual Learning Through Synaptic Intelligence. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 3987–3995.
- Lopez-Paz, D.; Ranzato, M. Gradient Episodic Memory for Continual Learning. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017; pp. 6467–6476.
- Schwarz, J.; Luketina, J.; Czarnecki, W.M.; Grabska-Barwinska, A.; Teh, Y.W.; Pascanu, R.; Gretton, A. Progress & Compress: A Scalable Framework for Continual Learning. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 4528–4537.
- Reif, J.H. Complexity of the Mover’s Problem and Generalizations. In Proceedings of the 20th Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 29–31 October 1979; pp. 421–427. [[CrossRef](#)]
- Canny, J.F. *The Complexity of Robot Motion Planning*; MIT Press: Cambridge, MA, USA, 1988.
- Goldstein, H.; Poole, C.; Safko, J. *Classical Mechanics*, 3rd ed.; Addison-Wesley: San Francisco, CA, USA, 2001.
- Arnold, V.I. *Mathematical Methods of Classical Mechanics*, 2nd ed.; Springer: New York, NY, USA, 1989. [[CrossRef](#)]
- Still, S.; Sivak, D.A.; Bell, A.J.; Crooks, G.E. Thermodynamics of Prediction. *Phys. Rev. Lett.* **2012**, *109*, 120604. [[CrossRef](#)]
- Wolpert, D.H. Information Theory—The Bridge Connecting Bounded Rational Game Theory and Statistical Physics. In *Complex Engineered Systems; Understanding Complex Systems*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 262–290. [[CrossRef](#)]
- Parrondo, J.M.R.; Horowitz, J.M.; Sagawa, T. Thermodynamics of Information. *Nat. Phys.* **2015**, *11*, 131–139. [[CrossRef](#)]
- van der Schaft, A.; Jeltsema, D. Port-Hamiltonian Systems Theory: An Introductory Overview. *Found. Trends Syst. Control* **2014**, *1*, 173–378. [[CrossRef](#)]
- Todorov, E. Efficient Computation of Optimal Actions. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11478–11483. [[CrossRef](#)]
- Kappen, H.J. Path Integrals and Symmetry Breaking for Optimal Control Theory. *J. Stat. Mech. Theory Exp.* **2005**, *2005*, P11011. [[CrossRef](#)]
- Eysenbach, B.; Levine, S. Maximum Entropy RL (Provably) Solves Some Robust RL Problems. In Proceedings of the International Conference on Learning Representations, Virtual, 25 April 2022.
- Berkenkamp, F.; Turchetta, M.; Schoellig, A.P.; Krause, A. Safe Model-Based Reinforcement Learning with Stability Guarantees. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017; pp. 908–918.
- Fisac, J.F.; Akametalu, A.K.; Zeilinger, M.N.; Kaynama, S.; Gillula, J.; Tomlin, C.J. A General Safety Framework for Learning-Based Control in Uncertain Robotic Systems. *IEEE Trans. Autom. Control* **2019**, *64*, 2737–2752. [[CrossRef](#)]
- Desai, S.A.; Mattheakis, M.; Sondak, D.; Protopapas, P.; Roberts, S.J. Port-Hamiltonian Neural Networks for Learning Explicit Time-Dependent Dynamical Systems. *Phys. Rev. E* **2021**, *104*, 034312. [[CrossRef](#)]
- Sosanya, A.; Greydanus, S. Dissipative Hamiltonian Neural Networks: Learning Dissipative and Conservative Dynamics Separately. *arXiv* **2022**, arXiv:2201.10085. [[CrossRef](#)]

28. Zhong, Y.D.; Dey, B.; Chakraborty, A. Dissipative SymODEN: Encoding Hamiltonian Dynamics with Dissipation and Control into Deep Learning. *arXiv* **2020**, arXiv:2002.08860. [[CrossRef](#)]
29. Bajaj, C. Proving geometric algorithm non-solvability: An application of factoring polynomials. *J. Symb. Comput.* **1986**, *2*, 99–102. [[CrossRef](#)]
30. Bajaj, C. Geometric optimization and the polynomial hierarchy. *Theor. Comput. Sci.* **1987**, *54*, 87–102. [[CrossRef](#)]
31. Bajaj, C.; Li, M. Geometric optimization and D^P -completeness. *Discret. Comput. Geom.* **1989**, *4*, 3–13. [[CrossRef](#)]
32. Li, M.; Vitányi, P.M.B. *An Introduction to Kolmogorov Complexity and Its Applications*, 4th ed.; Springer: Cham, Switzerland, 2019. [[CrossRef](#)]
33. Vitányi, P.M. Conditional Kolmogorov complexity and universal probability. *Theor. Comput. Sci.* **2013**, *505*, 93–100. [[CrossRef](#)]
34. Grünwald, P.D. *The Minimum Description Length Principle*; MIT Press: Cambridge, MA, USA, 2007.
35. Landauer, R. Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.* **1961**, *5*, 183–191. [[CrossRef](#)]
36. Bennett, C.H. The Thermodynamics of Computation—A Review. *Int. J. Theor. Phys.* **1982**, *21*, 905–940. [[CrossRef](#)]
37. Fredkin, E.; Toffoli, T. Conservative Logic. *Int. J. Theor. Phys.* **1982**, *21*, 219–253. [[CrossRef](#)]
38. Boyd, A.B.; Mandal, D.; Riechers, P.M.; Crutchfield, J.P. Transient Dissipation and Structural Costs of Physical Information Transduction. *Phys. Rev. Lett.* **2017**, *118*, 220602. [[CrossRef](#)]
39. Morrison, P.J. A paradigm for joined Hamiltonian and dissipative systems. *Phys. D Nonlinear Phenom.* **1986**, *18*, 410–419. [[CrossRef](#)]
40. Bajaj, C. Computer Algebra Meets Hamiltonian Geometry. *Maple Trans.* **2026**, *6*, 24248. [[CrossRef](#)]
41. Li, M.; Vitányi, P.M.B. Inductive reasoning and Kolmogorov complexity. *J. Comput. Syst. Sci.* **1992**, *44*, 343–384. [[CrossRef](#)]
42. Nguyen, M.P.; Bajaj, C.L. A Differential and Pointwise Control Approach to Reinforcement Learning. In Proceedings of the Thirty-Ninth Annual Conference on Neural Information Processing Systems, San Diego, CA, USA, 2–7 December 2025.
43. Boyd, A.B.; Mandal, D.; Crutchfield, J.P. Thermodynamics of Modularity: Structural Costs Beyond the Landauer Bound. *Phys. Rev. X* **2018**, *8*, 031036. [[CrossRef](#)]
44. Crutchfield, J.P. The calculi of emergence: Computation, dynamics and induction. *Phys. D Nonlinear Phenom.* **1994**, *75*, 11–54. [[CrossRef](#)]
45. Szilard, L. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. *Z. Phys.* **1929**, *53*, 840–856. [[CrossRef](#)]
46. Boyd, A.B.; Mandal, D.; Crutchfield, J.P. Correlation-Powered Information Engines and the Thermodynamics of Self-Correction. *Phys. Rev. E* **2017**, *95*, 012152. [[CrossRef](#)]
47. Lázaro-Camí, J.A.; Ortega, J.P. Stochastic Hamiltonian Dynamical Systems. *Rep. Math. Phys.* **2008**, *61*, 65–122. [[CrossRef](#)]
48. Cordoni, F.; Di Persio, L.; Muradore, R. Stochastic port-Hamiltonian systems. *J. Nonlinear Sci.* **2022**, *32*, 91. [[CrossRef](#)]
49. Bhardwaj, S.; Bajaj, C. PHAST: Port-Hamiltonian Architecture for Structured Temporal Dynamics Forecasting. *arXiv* **2026**, arXiv:2602.17998. [[CrossRef](#)]
50. Hsu, A.S.; Chater, N.; Vitányi, P. Language Learning From Positive Evidence, Reconsidered: A Simplicity-Based Approach. *Top. Cogn. Sci.* **2013**, *5*, 35–55. [[CrossRef](#)]
51. Nguyen, M.P.; Bajaj, C.L. Stochastic Differential Policy Optimization: A Rough Path Approach to Reinforcement Learning. In Proceedings of the Workshop on the Theory of AI for Scientific Computing, Lyon, France, 30 June 2025.
52. Pontryagin, L.S.; Boltyanskii, V.G.; Gamkrelidze, R.V.; Mishchenko, E.F. *The Mathematical Theory of Optimal Processes*; Interscience: New York, NY, USA, 1962.
53. Bajaj, C.; Nguyen, M. Physics-informed neural networks via stochastic hamiltonian dynamics learning. In *Proceedings of the Intelligent Systems Conference*; Springer: Berlin/Heidelberg, Germany, 2024; pp. 182–197.
54. Baez, J.C.; Stay, M. Algorithmic Thermodynamics. *Math. Struct. Comput. Sci.* **2012**, *22*, 771–787. [[CrossRef](#)]
55. Hairer, E.; Lubich, C.; Wanner, G. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2006. [[CrossRef](#)]
56. Quispel, G.R.W.; McLaren, D.I. A New Class of Energy-Preserving Numerical Integration Methods. *J. Phys. A Math. Theor.* **2008**, *41*, 045206. [[CrossRef](#)]
57. McLennan, L.; Wang, Y.; Farrell, R.; Nguyen, M.; Bajaj, C. Learning Generalized Hamiltonian Dynamics with Stability from Noisy Trajectory Data. *arXiv* **2025**, arXiv:2509.07280. [[CrossRef](#)]
58. Evensen, G. *Data Assimilation: The Ensemble Kalman Filter*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009. [[CrossRef](#)]
59. Reich, S.; Cotter, C. *Probabilistic Forecasting and Bayesian Data Assimilation*; Cambridge University Press: Cambridge, UK, 2015. [[CrossRef](#)]
60. Ellendula, A.S.; Wang, Y.; Bajaj, C.L. Learning Material-Aware Hamiltonian Risk Fields for Safe Navigation. *arXiv* **2026**, arXiv:2605.07038. [[CrossRef](#)]

61. Wang, Y.; Bajaj, C. When Descent Is Too Stable: Event-Triggered Hamiltonian Learning to Optimize. *arXiv* **2026**, arXiv:2605.06868. [[CrossRef](#)]
62. Bédard, C.; Bergeron, D. An Algorithmic Approach to Emergence. *Entropy* **2022**, *24*, 985. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.