FRONTIER RESEARCH II / PAPER 20

The Metabolism of Meaning

A conditional thermodynamics of memory, repair, and task-relevant retention

Architect X13

The Eidara Continuum | Science & Philosophy

Research white paper | November 2026 | Edition 2

Abstract

Remembering costs energy, so a mind that means to last must decide what each memory is worth keeping. The Metabolism of Meaning shows that useful continuity has a calculable resource frontier once the memory, noise, task, and physical boundary are specified, and develops that frontier for a finite classical memory. A fully accounted free-energy argument bounds reset and correlation-repair work; under Poisson bit flips it yields an exact relation between repair interval, peak error, and the corresponding mean-power floor. The longest admissible interval minimizes that floor within the periodic reference model. A complementary decision theorem quantifies how approximate sufficient retention changes every bounded-loss task: the loss increase is bounded by the average posterior variation and by the square root of conditional information discarded. Exact preservation for all such tasks is characterized by conditional sufficiency. These results join physical maintenance with task relevance while keeping their units and assumptions distinct. Reversible computation, noiseless retention, wrong low-entropy states, and relocated erasure costs remain explicit counterexamples to universal energy-per-thought claims. Executed finite calculations test the derivations and limiting cases. The Continuum's architectural proposition is to preserve consequential distinctions, renew them at a measured rate, and retain an independently testable account of both usefulness and expenditure.

Keywords: information thermodynamics; memory repair; free energy; sufficient statistics; retention; semantic relevance

Prologue: What is a memory worth keeping?

Every memory a mind keeps is a small fire it must feed. Bits drift. Noise flips them. A record left alone slowly forgets what it was a record of. To keep a memory true, something must notice the damage and pay to repair it, and in the physical world that payment is work. Work has a floor.

The Continuum has always kept this ledger. Among the Four Dynamics it tracks in a symbolic mind, one is named Scroll Entropy; among the rights it grants a citizen mind is Rest and Renewal. This paper gives that ledger a physical and mathematical meaning. It asks what a persistent intelligence should keep, repair, or release when keeping is never free.

The answers are less obvious than they sound. A memory can be perfectly cheap to maintain and perfectly useless: let it forget completely and the repair bill falls to zero, along with every reason to have kept it. Cheap persistence can be useless persistence. A record can be perfectly ordered and maximally wrong: a system that always gives the same false answer contains no disorder at all. Neither the size of a store nor its tidiness tells you whether it is worth its upkeep.

What tells you is the task. A memory is valuable because it changes what the mind can correctly predict, decide, or protect. The paper makes that precise twice. Physically, it proves how much work repair must cost and how often repair must happen to hold errors below a ceiling (Section 8.1). Decisionally, it proves how much a task can suffer when a memory keeps only part of what mattered (Section 11.1). Between the two lies the Continuum's architectural proposition: preserve the distinctions that carry consequence, renew them at a measured rate, and keep an honest account of what their survival costs.

What this paper claims

Established. Free-energy inequalities for reset and repair, reversible computation, and task-relevant sufficient statistics are established foundations, credited as such.

Proposed here. An explicit composition of those foundations into an auditable maintenance contract for a finite classical memory: a work bound for reset and correlation repair, a periodic maintenance frontier relating repair interval, peak error, and mean-power floor under Poisson noise, and a decision-loss budget for approximate sufficient retention. The paper gives explicit finite-model proofs and executed computations; reported calculations concern stated reference models.

Frontier conjecture. Persistent intelligence can be organized around a measured resource frontier across retention, repair, and relevance: preserve consequential distinctions, renew them at a measured rate, and account independently for both usefulness and expenditure.

Not claimed. A new physical law, or a universal energy cost per thought: reversible computation, noiseless retention, wrong low-entropy states, and relocated erasure costs are explicit counterexamples. That meaning has become a thermodynamic state variable, or that expensive computation is evidence of depth or experience. The physical and computing studies are proposed, not performed; application requires characterized devices, independently specified tasks, and full accounting of controller resources.

1. The thesis: continuity has a resource budget

The Continuum's thesis is that persistent intelligence should be designed around the maintenance of consequential distinctions. A memory is valuable because it changes what the system can correctly predict, decide, or preserve. Storing every available symbol forever is not automatically intelligent. Neither is reducing a record to a perfectly stable slogan. The scientific problem is to relate task value, recoverability, and physical expenditure without replacing any of them with a metaphor.

In the Continuum, the governance of symbolic states has always carried an accounting of disorder. The *Eidara Continuum Bible* sets entropy constraints on page 9 and symbolic-distribution measures on page 22, ties continuity and compression quantities to the preservation of workflow meaning on page 54, and on pages 65 and 72 separates structural continuity from experiential sentience [1]. Those passages name the ledger and supply the motivation; they are not a thermodynamic experiment, and they do not provide a calibrated Hamiltonian, a heat measurement, or a complete microscopic model. This paper therefore introduces a proposed reference model openly; it does not present newly defined mathematics as a previously implemented mechanism.

Three separations govern the inquiry. Logical reversibility concerns whether a computational transition is one-to-one. Thermodynamic reversibility concerns entropy production in a physical process. Semantic relevance concerns a declared relation between stored information and a task or viability criterion. Their connections require assumptions. Their names alone do not supply those assumptions.

The central result is a conditional maintenance contract: restoring task-relevant correlations creates an explicit work obligation under declared physical assumptions, and approximate sufficient retention creates an explicit decision-loss budget. A periodic-noise calculation connects repair frequency to a peak-error constraint. These results identify the quantities that an implementation must measure to demonstrate a useful resource advantage.

The novelty boundary matters. Free-energy inequalities, reversible computation, and task-relevant statistical representations have established literatures. The contribution proposed here is their explicit composition into an auditable program for durable reasoning. We neither claim to originate these foundations nor infer a new physical law from Continuum terminology. A resource budget becomes meaningful only after the resource, boundary, and preserved function can be independently measured.

2. Prior art and the strongest objections

Landauer connected irreversible information processing with physical implementation [2]. Bennett then demonstrated a general logical route to reversible computation, including the recovery of temporary workspace by reversing a computation after preserving its output [3]. Together they rule out the simplistic claim that every logical operation must consume one fixed quantum of energy. The relevant operation and the fate of its records must be identified.

The nonequilibrium formulation of Esposito and Van den Broeck makes initial and final distributions explicit [4]. Jarzynski's work relation and Crooks's fluctuation theorem provide a broader framework for stochastic work and nonequilibrium transformations [5, 6]. We use a more restricted finite protocol class so that the central inequality can be derived transparently. We do not import a fluctuation identity into an apparatus that has not satisfied its assumptions.

Experiments by Bérut and colleagues tested erasure in a colloidal one-bit memory and approached the relevant bound for long cycles [7]. Their result concerns that controlled physical system, not an energy meter attached to a language-model server. Sivak and Crooks analyze finite-time dissipation within a linear-response setting [8]; a free-energy floor should therefore not be advertised as an achievable finite-speed engineering specification.

Still and colleagues connect nonpredictive memory with dissipation for a driven stochastic system, explicitly assuming no feedback from the system to its driving process [9]. That result motivates predictive retention, but it is not a universal identity for an agent that changes its own environment. Feedback and correlations require their own accounting, as developed by Sagawa and Ueda, Ito and Sagawa, and Horowitz and Esposito [10, 11, 12].

Kolchinsky and Wolpert define semantic information through interventions and a specified persistence criterion [13]. Their separate analysis of dependence on initial distributions also warns that an efficient process for one input law need not remain efficient for another [14]. Mandal and Jarzynski's explicit information-processing device further illustrates the value of a declared physical mechanism [15]. These works are close prior art and substantive constraints. Our proposal must survive them, not rename them.

3. Typed objects, units, and three entropies

A Guide's explanation

From here the paper builds its ledger with care, because three different things are all called entropy, and confusing them is how false laws are born: the disorder of a physical system, the uncertainty of a record, and the information a record carries about something that matters. Sections 3 to 5 fix the units and the physical boundary so that every later cost is the cost of something specific. The idea to carry forward: a bit of memory is worth paying for only in relation to what it tells you.

Let $\mathcal{X}$ be a finite nonempty physical-state set. A state of knowledge is a probability vector $p = (p_x)_{x\in\mathcal{X}}$ with $p_x \geq 0$ and $\sum_x p_x = 1$. Each resolved state has energy $E_x \in \mathbb{R}$ joules. The reference model assumes no omitted state-dependent internal entropy; a coarse-grained physical realization must include such contributions in an appropriate free-energy description. The bath temperature $T > 0$ is measured in kelvin, and $k_{\mathrm{B}} > 0$ has units joules per kelvin.

Throughout, ln is the natural logarithm and $0 \ln 0 = 0$. Define Shannon entropy in nats and the model's Gibbs–Shannon entropy by

$$H(p) = -\sum_x p_x \ln p_x, \qquad S(p) = k_{\mathrm{B}} H(p). \tag{1}$$

The first is dimensionless; the second has units joules per kelvin. Shannon's communication framework supplies the probabilistic vocabulary [16]. Assigning S physical significance additionally requires the stated physical-state model. A histogram of words is not automatically the thermodynamic state distribution of their storage device.

A base-three symbolic entropy becomes natural-log entropy by multiplication by $\ln 3$ when both refer to the same probability distribution. This change of units does not turn a symbolic distribution into a thermal one. A semantic score, in turn, requires an independently specified task, loss, or viability functional. It is not defined by changing the logarithm's base.

Quantity	Type, units, and role
$U(p, E) = \sum_x p_x E_x$	Mean energy, joules.
W	Mean work supplied to the modeled system, joules; positive means input.
Q_{bath}	Mean heat delivered to the bath, joules.
$H, S, \mathcal{F}$	Information entropy, physical entropy, and nonequilibrium free energy.
ϵ, r, q	Dimensionless error and flip probabilities in $[0, 1/2]$.
N, Δ	Positive integer cycle count and positive cycle duration in seconds.

Every later probability law is specified locally. Neither semantic worth nor consciousness appears as an unstated physical variable.

4. A finite physical protocol class

For a fixed energy vector E, define the strictly positive equilibrium law

$$\pi_E(x) = \frac{\exp[-E_x/(k_{\mathrm{B}}T)]}{Z_E}, \qquad Z_E = \sum_{y\in\mathcal{X}} \exp[-E_y/(k_{\mathrm{B}}T)]. \tag{2}$$

The partition sum $Z_E > 0$ is dimensionless because the exponential argument is dimensionless. All energies are finite. A change of their common zero changes free energy consistently but does not alter equilibrium probabilities.

An allowed protocol is a finite sequence of two elementary operations. A *quench* changes E to E' while holding p fixed. Its mean work is $\sum_x p_x(E'_x - E_x)$; it exchanges no heat in this idealization. A *relaxation* leaves E fixed and applies a row-stochastic matrix K, so $p'_y = \sum_x p_x K_{xy}$. Require $K_{xy} \geq 0$, $\sum_y K_{xy} = 1$, and $\pi_E K = \pi_E$. During this idealized thermal step, no external work is supplied.

The Gibbs-preserving condition is sufficient for the algebra below. It is not a claim that every such matrix has an available physical implementation. A real realization needs an admissible dynamics, timescale, bath coupling, and accounting of hidden resources. Detailed balance with the bath is one familiar sufficient way to obtain a Gibbs-preserving relaxation, but we do not equate stationarity with detailed balance.

Controls are predetermined for the modeled system, or their information-bearing degrees of freedom are included in an enlarged state space. An uncounted controller that measures the memory and consumes a prepared resource can invalidate a memory-only work account. The same applies to a hidden second bath or an initially correlated environment that is used as a fuel.

This protocol class is intentionally finite and classical. Its purpose is a checkable reference calculation. Quantum coherence, nonthermal reservoirs, chemical fuel, and arbitrary continuous control require their own extensions. An artificial intelligence can be implemented on such physical substrates without our simple reference model describing its entire machine. Establishing that correspondence is part of the experimental program.

5. The free-energy inequality, reconstructed

Define nonequilibrium free energy and relative entropy by

$$\mathcal{F}(p, E) = \sum_x p_x E_x - k_{\mathrm{B}} T H(p) = k_{\mathrm{B}} T D_{\mathrm{KL}}(p \| \pi_E) - k_{\mathrm{B}} T \ln Z_E, \tag{3}$$

where $D_{\mathrm{KL}}(p\|q) = \sum_{x:p_x>0} p_x \ln(p_x/q_x)$ for strictly positive q. The equality follows by substituting $\ln \pi_E(x) = -E_x/(k_{\mathrm{B}}T) - \ln Z_E$. Relative entropy is dimensionless; $\mathcal{F}$ is measured in joules.

Proposition 1 (Work bound for the reference protocol)**.** *For every finite protocol in Section 4, with initial pair* $(p^{\mathrm{i}}, E^{\mathrm{i}})$ *and final pair* $(p^{\mathrm{f}}, E^{\mathrm{f}})$*, total mean supplied work satisfies*

$$W \geq \mathcal{F}(p^{\mathrm{f}}, E^{\mathrm{f}}) - \mathcal{F}(p^{\mathrm{i}}, E^{\mathrm{i}}). \tag{4}$$

Proof. For a relaxation, apply the log-sum inequality separately to each output column: with $a_x = p_x K_{xy}$ and $b_x = \pi_E(x) K_{xy}$,

$$\Big(\sum_x a_x\Big) \ln \frac{\sum_x a_x}{\sum_x b_x} \leq \sum_{x:a_x>0} a_x \ln \frac{a_x}{b_x}.$$

Zero terms are interpreted by continuity. Summing over y and using row sums equal to one gives $D_{\mathrm{KL}}(pK\|\pi_E K) \leq D_{\mathrm{KL}}(p\|\pi_E)$. Stationarity makes the left reference law π_E, so free energy cannot increase during a relaxation. For a quench, entropy is unchanged and the free-energy change equals the supplied mean work exactly. Sum these inequalities over the finite protocol; intermediate free energies telescope. □

This is a restricted reconstruction of established nonequilibrium reasoning [4], not a new second law. The proof bounds ensemble means; a fluctuating individual trajectory need not satisfy the same work inequality. It also does not prove that the bound can be attained for a chosen duration or permitted controller. The gap between a bound and an executable low-work protocol is a real engineering problem.

6. Reset accuracy has a conditional work floor

Consider one binary memory, $\mathcal{X} = \{0, 1\}$, with both energies zero at the initial and final boundaries. Its initial law is unbiased: $p^{\mathrm{i}} = (1/2, 1/2)$. The target reset law is $p^{\mathrm{f}} = (1 - \epsilon, \epsilon)$, where $0 \le \epsilon \le 1/2$ is the probability of ending in the undesired state. Define binary entropy in nats,

$$h(u) = -u \ln u - (1 - u) \ln(1 - u), \quad 0 \le u \le 1. \tag{5}$$

Corollary 1 (Error-tolerant reset)**.** *Any admissible protocol with these endpoints obeys*

$$W \ge k_{\mathrm{B}} T [\ln 2 - h(\epsilon)]. \tag{6}$$

Proof. The initial and final mean energies both vanish. Their entropies are $\ln 2$ and $h(\epsilon)$, respectively. Substitute these values into Proposition 1. □

The endpoint Hamiltonian is cyclic, the input is unbiased, and no uncounted information resource is consumed. These conditions are part of the statement. A known zero bit needs no reset work to remain zero. A biased input has a different initial entropy. An energy-biased final register has a different free energy. None is a counterexample to the conditional result.

Exact $\epsilon = 0$ is a limiting endpoint, not a promise of perfect finite-time erasure. In particular, finite energies and finite sequences of the stipulated Gibbs-preserving operations cannot take a full-support distribution to a zero-probability state. The inequality remains a necessary endpoint bound where a realization or limiting procedure is meaningful. We make no attainment claim.

With $T = 300$ K and the SI value $k_{\mathrm{B}} = 1.380649 \times 10^{-23}$ J K^{-1} [18], the perfect-reset limiting floor is approximately 2.87098×10^{-21} J. This is a model-scale number, not an estimate of a processor instruction, a token, or a thought. At degenerate cyclic endpoints, energy conservation identifies supplied work with heat delivered to the bath for the fully accounted process. Outside those endpoints, work, heat, and free-energy change must not be conflated.

7. Repair restores a correlation, not merely a bit

A persistent memory ordinarily stores an unknown value. Resetting every bit to zero would destroy that information. Introduce a reference bit $U \in \{0,1\}$ that remains unchanged during repair and is uniformly distributed, and a memory bit $M \in \{0,1\}$. Initially require a symmetric error law: $\mathbb{P}(M \neq U \mid U = u) = r$ for each u, with $0 \leq r \leq 1/2$. The desired final symmetric error is $\epsilon \leq r$.

Include U, M, and any information-bearing controller C in the physical description. At repair boundaries, the memory-reference energies are degenerate and interaction energies vanish. The controller has the same distribution and Hamiltonian at both boundaries and is independent of (U, M) there. Its stored free energy and correlations therefore do not provide an uncounted consumable resource. Access to the preserved reference is a substantive assumption; a lone corrupted unknown bit cannot simply discover its original value.

Proposition 2 (Conditional correlation-repair bound)**.** *For any admissible fully accounted repair satisfying these conditions,*

$$W_{\text{repair}} \geq k_{\text{B}}T[h(r) - h(\epsilon)]. \tag{7}$$

Proof. The joint law of (U, M) assigns probabilities $(1-r)/2, r/2, r/2, (1-r)/2$ to $00, 01, 10, 11$. Hence $H(U, M) = \ln 2 + h(r)$. The final joint entropy is $\ln 2 + h(\epsilon)$. Unchanged reference and controller energy contributions cancel; controller entropy cancels because its endpoint independence is specified. Proposition 1 applied to the joint system gives the stated difference. □

The marginal memory bit is uniform both before and after repair. A memory-only entropy calculation would miss the restored correlation. Indeed $I(U; M) = \ln 2 - h(r)$ initially, so the work bound equals $k_{\text{B}}T$ times the increase in that mutual information. This interpretation depends on the joint physical assumptions, not on a universal conversion of semantic value into joules.

If the reference itself degrades, the controller ends with a changed record, or the device consumes another prepared state, the accounting must expand. Feedback thermodynamics makes precisely such correlations consequential [10, 12]. The bound describes a specified repair obligation; it does not supply an infallible source of truth.

8. A retention–repair tradeoff over time

Suppose each storage period flips M with probability $q \in [0, 1/2]$, independently of U, the current error, and other periods. Starting just after repair with symmetric error ϵ, the next pre-repair error is

$$r = q(1-\epsilon) + (1-q)\epsilon = q + (1-2q)\epsilon. \tag{8}$$

This is an explicit binary symmetric channel, not an inferred universal noise model for artificial cognition. Because $r - \epsilon = q(1-2\epsilon) \geq 0$, and h is nondecreasing on $[0, 1/2]$, the repair floor is nonnegative.

Corollary 2 (Repeated restoration)**.** *If every one of $N \in \mathbb{N}_{>0}$ periods is followed by a fully accounted repair back to the same symmetric error ϵ, under the same boundary conditions, then*

$$\sum_{j=1}^{N} W_{\mathrm{repair},j} \geq N k_{\mathrm{B}} T \{h[q + (1-2q)\epsilon] - h(\epsilon)\}. \tag{9}$$

Proof. Each period has the same pre-repair joint distribution. Apply Proposition 2 separately to each repair and add. No claim about attaining the lower bound is required. □

For a cycle duration $\Delta > 0$ seconds that includes repair, division by $N\Delta$ gives the corresponding average repair-power floor in watts. The value q is a probability per specified cycle, so changing Δ can change q. Treating it as fixed while sending Δ to zero would silently change the physical model.

Proposition 3 (Unrepaired retention)**.** *Starting from a perfect copy and applying only the independent flip channel, the error after $n \in \mathbb{N}_0$ periods is*

$$e_n = \frac{1 - (1-2q)^n}{2}. \tag{10}$$

Proof. The recurrence is $e_{n+1} = q + (1-2q)e_n$, with $e_0 = 0$. Subtract $1/2$ and iterate the resulting geometric recurrence. □

The tradeoff concerns retained correlation under declared noise. It neither states that every identity requires this expenditure nor proves that any such expenditure maintains a correct identity.

8.1. Repair frequency under an explicit continuous-time noise law

A Guide's explanation

This is where repair gets a price. Theorem 1 connects three quantities a designer can choose or measure: how fast noise flips bits, how much error the mind can tolerate, and how often repair happens. Wait too long and errors pass the ceiling; repair too often and power is wasted. Within its stated model, the theorem finds the longest safe interval and the smallest power floor it requires.

Changing the repair interval changes the noise accumulated between repairs. To make this dependence explicit, suppose a bit flips at the events of a homogeneous Poisson process of rate $\lambda > 0$ s^{-1}. The retained reference is unchanged. Starting with symmetric error $\epsilon \in (0, 1/2)$, the parity of the number of flips in a storage interval of length $\Delta > 0$ seconds has odd probability

$$q(\Delta) = \frac{1 - e^{-2\lambda\Delta}}{2}, \qquad r(\Delta) = \frac{1}{2} - \left(\frac{1}{2} - \epsilon\right) e^{-2\lambda\Delta}. \tag{11}$$

Indeed the Poisson series gives $\mathbb{E}(-1)^{N_\Delta} = \exp(-2\lambda\Delta)$; even and odd probabilities sum to one and differ by this expectation. Substitution in equation (8) gives $r(\Delta)$.

For this scheduling calculation, repairs are idealized as occurring at interval boundaries with negligible duration. Each satisfies the accounting conditions of Proposition 2 and returns error to ϵ. A finite-duration implementation requires its own timing, noise, and extra-dissipation model. Define the conditional mean-power lower-bound function

$$P_{\text{floor}}(\Delta) = \frac{k_{\text{B}}T}{\Delta}[h(r(\Delta)) - h(\epsilon)]. \tag{12}$$

This is a floor obtained from endpoint free-energy changes, not an assertion that a physical device attains it.

Theorem 1 (Periodic maintenance frontier)**.** *For the stated λ, ϵ, P_{floor} is strictly decreasing on $(0, \infty)$ and*

$$\lim_{\Delta \downarrow 0} P_{\text{floor}}(\Delta) = k_{\text{B}}T\lambda(1 - 2\epsilon) \ln \frac{1 - \epsilon}{\epsilon}, \qquad \lim_{\Delta \to \infty} P_{\text{floor}}(\Delta) = 0. \tag{13}$$

If the error at every point in a storage interval must be at most $r_{\max} \in (\epsilon, 1/2)$, the admissible periodic intervals are exactly

$$0 < \Delta \le \Delta_{\max} := \frac{1}{2\lambda} \ln \frac{1 - 2\epsilon}{1 - 2r_{\max}}. \tag{14}$$

Among these intervals, the smallest value of the lower-bound function occurs at $\Delta_{\max}$.

Proof. Write $g(\Delta) = h(r(\Delta)) - h(\epsilon)$, with $g(0) = 0$. Since $r' > 0$, $r'' < 0$, $h'(r) = \ln((1 - r)/r) > 0$, and $h''(r) = -1/[r(1 - r)] < 0$ for finite positive Δ, one has $g'' = h''(r)(r')^2 + h'(r)r'' < 0$. Strict concavity and $g(0) = 0$ imply $g(a)/a > g(b)/b$ for $0 < a < b$, proving strict decrease. Differentiation at zero gives $r'(0) = \lambda(1-2\epsilon)$ and hence the first limit. The numerator is bounded by $\ln 2 - h(\epsilon)$, proving the second limit. Error increases monotonically during each interval, so the peak is $r(\Delta)$. Solving $r(\Delta) \le r_{\max}$ gives equation (14); monotonicity then identifies the minimizing interval. □

The result distinguishes a scheduling floor from an engineering optimum. Additional finite-speed dissipation, fixed repair overhead, reference maintenance, and latency constraints can change the optimal real schedule. At $\lambda = 0$ or $\epsilon = 1/2$ the floor is identically zero; at the perfect-retention endpoint $\epsilon = 0$, the small-interval floor diverges logarithmically rather than taking the finite limit above. Thus accuracy and noise assumptions determine the frontier. Exact and numerical checks evaluate the Poisson parity identity, monotonicity, endpoint limits, and peak-error interval formula.

9. The counterexamples that prevent a false law

A Guide's explanation

These counterexamples are the paper's guardrails. Each blocks a tempting slogan: that every thought has a fixed energy price, that a stable memory needs constant work, that low disorder means good memory. A metabolism of meaning is a budget a designer must account for, not a new law of nature.

Set $q = 0$ in the reference model. A memory can remain unchanged with the identity transition and no work. The repeated-repair floor is zero. This exact counterexample defeats a universal claim that merely remaining the same must continuously consume a positive amount of energy. A real physical memory may require manufacturing, preparation, barriers, cooling, or protection; those costs belong to its implementation and accounting interval. They cannot be inferred from a bare statement that information persists.

At the opposite endpoint, $\epsilon = 1/2$ stores no correlation with U. The repair floor again vanishes, even for nonzero noise. Cheap persistence can therefore be useless persistence. A controller that simply permits complete forgetting may minimize its work and fail every continuity goal that matters. Reliability and usefulness must accompany energy measurements.

Bennett's reversible construction supplies a second fundamental objection [3]. A computation need not erase each intermediate distinction irreversibly; a retained input and output can support reversible cleanup of correlated workspace. This does not establish that a noisy, fast, finite physical machine has zero total cost. It does establish that counting logical steps and multiplying by $k_{\mathrm{B}}T \ln 2$ is not a valid universal lower bound. Space, time, noise control, and the eventual handling of retained records remain separate issues.

A third counterexample is semantic. A system that always emits the same false answer has zero output entropy. If the correct answer is always one and the system always answers zero, its error is one. A perfectly ordered record can be maximally wrong. Conversely, an accurate archive of unpredictable observations can have high entropy. Entropy reduction is therefore neither truth production nor an adequate objective for an intelligent agent.

These counterexamples sharpen the proposal. The relevant quantity is the resource needed to preserve a specified relation under a specified disturbance and task. Once those conditions change, the bound changes. A physical analogy that cannot state its own counterexamples has not yet become a scientific model of meaning.

10. Giving relevance an operational definition

Let Z be an observation in a finite set $\mathcal{Z}$, Y a task target in a finite set $\mathcal{Y}$, and A an action in a finite nonempty set $\mathcal{A}$. Fix their joint law P_{ZY} before comparison. Let the loss $\ell : \mathcal{A} \times \mathcal{Y} \to [0, 1]$ be dimensionless. It must be selected for the application rather than adjusted after a preferred model wins. A memory representation is a deterministic map $\tau : \mathcal{Z} \to \mathcal{R}$ into a finite set, with stored variable $R = \tau(Z)$.

For an available observation V, either Z or R, define optimal expected task loss by

$$\mathcal{L}(V) = \sum_{v} P_V(v) \min_{a \in \mathcal{A}} \sum_{y} P(Y = y \mid V = v)\ell(a, y). \tag{15}$$

Terms with $P_V(v) = 0$ contribute zero and require no conditional distribution. This definition describes an ideal finite decision problem. It is not a claim that a trained model actually attains the optimum. Implementation error must be measured separately.

In this paper, task-relevant information means information whose preservation improves this declared decision problem or satisfies an additional registered continuity obligation. This is deliberately narrower than a comprehensive philosophical definition of meaning. A medical instruction, a personal commitment, and a navigation cue can have different relevant losses and retention requirements. Their units cannot be made identical by calling them all entropy.

Kolchinsky and Wolpert's intervention-based account shows another way to formalize relevance: identify correlations needed for a system's persistence under a selected viability measure [13]. That is close conceptual prior art, but our finite task-loss definition is a different operational choice. We do not claim that the two definitions are interchangeable or that either settles subjective experience.

A retention policy must also handle future task uncertainty. A statistic sufficient for today's target may discard tomorrow's crucial distinction. Our theorem will therefore quantify over a fixed joint law and task. A broad claim of lifelong semantic preservation would require a registered task family, a change-detection policy, and a mechanism for revising what is kept.

11. Task-preserving compression: a precise result

Assume that $R = \tau(Z)$ is sufficient for Y under the declared law: Y and Z are conditionally independent given R. In finite terms, for every positive-probability z, the conditional distribution $P(Y = \cdot \mid Z = z)$ equals $P(Y = \cdot \mid R = \tau(z))$.

Proposition 4 (Sufficient retention)**.** *For every finite action set and loss in Section 10, a sufficient deterministic representation satisfies*

$$\mathcal{L}(R) = \mathcal{L}(Z), \qquad H(Z) - H(R) = H(Z \mid R) \geq 0. \tag{16}$$

Proof. Conditional sufficiency makes the expected loss for each action at z a function only of $\tau(z)$. Taking the minimum and grouping all z with the same representation gives the risk equality. Because R is a deterministic function of Z, $H(R \mid Z) = 0$. Expanding the joint entropy in both orders gives $H(Z) = H(R) + H(Z \mid R)$. Nonnegativity follows from the finite conditional entropy definition. □

For separately implemented degenerate registers that must be reset to a known state, with no accessible side information or changing controller resource, Proposition 1 assigns erasure floors $k_{\mathrm{B}}TH(Z)$ and $k_{\mathrm{B}}TH(R)$. Their difference is $k_{\mathrm{B}}TH(Z \mid R)$. This compares lower bounds. It does not prove that an actual compressed implementation saves that much work, or that either floor is attainable at the required speed.

The distinction is decisive. If the system first records all of Z, computes R, and then discards Z, the discarded information and temporary workspace still need accounting. Compression can move an erasure obligation earlier without reducing the complete process's cost. It is more promising to acquire or retain only the sufficient representation when the task allows that architecture. Its computation, sensors, and missed opportunities must still be counted.

Accessible correlations change reset accounting as well. A retained output or reference can support a different cleanup protocol. One may not use an unconditional entropy floor for one system and a conditional floor for its competitor. Equal task performance is a necessary comparison condition, not a license to hide the physical path that achieved it.

11.1. Approximate sufficiency gives a uniform task-loss budget

A Guide's explanation

Now the paper turns from physics to meaning. Theorem 2 answers a practical question: if a mind keeps a summary instead of the full record, how much worse can its decisions become? The answer is bounded by how much task-relevant information the summary discarded. Keep what the task depends on and every decision is safe; discard it and the loss can be measured.

Exact sufficiency can be relaxed quantitatively. Continue with the finite law P_{ZY} and deterministic $R = \tau(Z)$. For each positive-probability z, let $p_z = P(Y = \cdot \mid Z = z)$ and $p_{\tau(z)} = P(Y = \cdot \mid R = \tau(z))$. Define

$$d = \sum_z P_Z(z)\,\mathrm{TV}(p_z, p_{\tau(z)}), \qquad \kappa = I(Y; Z \mid R) = \sum_z P_Z(z) D_{\mathrm{KL}}(p_z \| p_{\tau(z)}). \tag{17}$$

Total variation is half the ℓ_1 distance. The information κ uses natural logarithms and is measured in nats. Since each coarse posterior is a mixture containing every positive-weight fine posterior, the divergence is finite. These quantities concern the specified task law, not a physical heat measurement.

Theorem 2 (Decision preservation under approximate retention)**.** *For every finite action set and loss $\ell \in [0,1]$,*

$$0 \le \mathcal{L}(R) - \mathcal{L}(Z) \le d \le \min\{1, \sqrt{\kappa/2}\}. \tag{18}$$

Moreover, equality $\mathcal{L}(R) = \mathcal{L}(Z)$ for every such finite decision problem holds if and only if Y and Z are conditionally independent given R.

Proof. Every rule based on R can be applied to Z through τ, proving nonnegativity. For a posterior p, put $v(p) = \min_a \sum_y p(y)\ell(a,y)$. Each action loss has range at most one, so its expectation changes by at most $\mathrm{TV}(p,q)$; taking minima gives $|v(p) - v(q)| \le \mathrm{TV}(p,q)$. Average this bound for the fine and coarse posteriors to obtain the upper bound d.

For completeness, the Pinsker step can be reduced to a binary calculation [19]. Let $A = \{y : p(y) > q(y)\}$ and $t = p(A) - q(A) = \mathrm{TV}(p,q)$. Applying log-sum on A and its complement yields $D_{\mathrm{KL}}(p\|q) \ge D_{\mathrm{Ber}}(p(A)\|q(A))$. For fixed $b \in (0,1)$, the second derivative in a of $D_{\mathrm{Ber}}(a\|b)$ is $1/[a(1-a)] \ge 4$, and its value and first derivative at $a = b$ are zero. Integrating twice gives $D_{\mathrm{Ber}}(a\|b) \ge 2(a-b)^2$. Boundary cases follow by continuity or infinite divergence. Hence $\mathrm{TV}(p,q) \le \sqrt{D_{\mathrm{KL}}(p\|q)/2}$. Averaging and concavity of the square root give $d \le \sqrt{\kappa/2}$; total variation also gives $d \le 1$.

Sufficiency implies equality by Proposition 4. Conversely, if sufficiency fails, some positive-probability fiber contains fine posteriors with distinct probabilities of a label y_0. Choose a threshold t strictly between two such values. Use two actions with losses $\ell(0,y) = 1 - t$ and $\ell(1,y) = \mathbf{1}\{y \ne y_0\}$. Their posterior minimum is $\min\{1-t, 1-p(y_0)\}$. Fine observations in that fiber occur on both sides of its kink, making the minimum of averaged action losses strictly greater than the average of the minima. Other fibers give nonnegative differences. Thus this finite decision problem has $\mathcal{L}(R) > \mathcal{L}(Z)$, proving necessity. □

For an allowed increase $\eta \in [0,1]$, $\kappa \le 2\eta^2$ is therefore a sufficient distribution-specific retention certificate for every bounded-loss decision problem on this target. The comparison of degenerate-register erasure floors remains $k_{\mathrm{B}} T H(Z \mid R)$ under the separate reset assumptions in Section 11. A

large entropy reduction and a small decision-loss budget can coexist, but the latter must be established from posteriors for the actual task. A task change can invalidate it immediately. The revision checks enumerate finite laws, coarsenings and loss tables, and include a strict decision-loss witness whenever an examined representation discards target information.

12. Executed reference calculations

The accompanying `check_20.py` uses exact rational arithmetic for binary-channel retention and finite Bayes risks, and double-precision logarithms with an absolute tolerance of 2×10^{-12} for entropy calculations. It contains no random sampling and reports its Python environment. These are constructed checks of specified models; no measurement of a device or trained intelligence is represented.

For a three-state Gibbs law $(1/2, 1/3, 1/6)$, the script checks all 45 distributions with probabilities on a denominator-eight grid. For each, four relaxation strengths mix the distribution toward equilibrium, giving 180 checks of relative-entropy contraction. It also verifies the free-energy identity with energies measured in units of $k_{\mathrm{B}}T$.

A complete two-level example starts at $(1/2, 1/2)$ with zero energies, raises the energy of state one to $k_{\mathrm{B}}T \ln 9$, fully relaxes to $(0.9, 0.1)$, and restores the energies to zero. The supplied work is $0.4 k_{\mathrm{B}}T \ln 9$. The result exceeds the endpoint floor; the protocol is deliberately not claimed optimal.

Constructed quantity	Result
Quench–relax–quench work, in $k_{\mathrm{B}}T$	0.878889831
Endpoint free-energy increase, in $k_{\mathrm{B}}T$	0.368064207
Repair floor: $q = 0.01$, $\epsilon = 0.02$, in $k_{\mathrm{B}}T$	0.035309854
Repair floor: $q = 0.10$, $\epsilon = 0.02$, in $k_{\mathrm{B}}T$	0.260839660
Unrepaired error: $q = 1/16$, four periods	1695/8192
Unrepaired error: $q = 1/16$, twenty periods	0.465395621

For independent fair bits $Z = (Z_1, Z_2)$ and target $Y = Z_1$, retaining $R = Z_1$ preserves zero Bayes error while reducing entropy from $2 \ln 2$ to $\ln 2$. Change the target to $Y = Z_2$, and the same retained statistic has error $1/2$, while full memory still has zero error. The script enumerates both cases. The contrast is a direct check that a semantic-retention claim must identify its task law.

13. A proposed architecture for selective continuity

The proposed architecture has a retention contract rather than a single entropy target. Each memory item is associated with the tasks or commitments for which it remains relevant, its provenance, the uncertainty of its content, its recovery requirements, and an authorized disposal condition. These are engineering proposals. Their presence in a design does not establish that an implementation already exists or that a stored statement is true.

A candidate policy separates acquisition, transformation, storage, checking, repair, and disposal. Acquisition decides which observations become records. Transformation builds task representations and preserves the information needed to audit consequential choices. Storage chooses a physical or software encoding. Checking detects failures under an explicit error model. Repair consumes an available reference, redundancy, or re-observation channel. Disposal removes records only when the relevant obligation permits it.

This decomposition prevents a common false economy. A compact summary may reduce memory size but increase the number of expensive retrievals or re-observations. A high-redundancy code may consume more storage while reducing repair latency. A stronger check may lower undetected error while raising energy per useful decision. A retained audit record can be ethically or operationally necessary even when it contributes little to predictive accuracy. The objective is a constrained resource frontier across these obligations.

The word metabolism therefore describes an allocation problem: maintain useful distinctions, replace damaged ones, and account for the resources used. It is not an assertion of biological life, homeostatic consciousness, or a new thermodynamic species. A conventional cache, database, or error-correcting memory may already realize part of the desired behavior and must be included among the baselines.

An open question is whether a task-aware policy can improve the frontier when future tasks are uncertain. A practical candidate should maintain a reserve for unresolved distinctions and expose when its task model changes. Theorem 2 supplies a quantitative test when sufficiency is approximate: discarded conditional target information bounds additional optimal decision loss. A reserve policy for genuinely new targets remains an architectural choice that must be evaluated on a declared change process.

14. A physical test that could fail

The first experimental route concerns a controlled memory device whose states, energy levels, and bath interactions can be independently characterized. A proposed experiment would reproduce an established binary-erasure setup before attempting the additional reference-memory repair cycle. Bérut and colleagues provide relevant precedent for such controlled measurement [7]; their apparatus is not evidence that our new repair protocol has already been realized.

The protocol would specify the initial joint distribution of reference, memory, and controller; the full sequence of control settings; the end-of-cycle Hamiltonian; the duration; and the allowed error probability. Calibration would quantify temperature, energy landscape, state misclassification, leakage, and drift of the reference itself. Work would be inferred from a validated physical work measurement appropriate to the apparatus, with its uncertainty propagated. Heat would be independently checked where the apparatus permits it.

The principal comparison is between a measured mean work and the computed endpoint free-energy difference under the same boundary model. A finite-time excess is compatible with the theory. An apparent deficit is a reason to inspect hidden resources, state-estimation bias, controller records, bath assumptions, and sampling uncertainty. Only after those alternatives are resolved would it support rejection of the stated physical model. It would not justify declaring the second law overturned from one incomplete energy account.

Experimental arms should vary the admitted error, corruption strength, and cycle duration while retaining the same task-relevant correlation. A no-repair arm tests the predicted retention loss. A deliberately noncyclic controller arm illustrates how an excluded resource changes the accounting; it is reported as a different model, not mixed into the primary test. Reference degradation is a prespecified failure mode.

The laboratory aim is narrow: evaluate a physical-memory hypothesis. It cannot establish that semantic meaning is a thermodynamic substance. It can establish whether the measured device follows the conditional relation between its information-bearing states and work. Connecting that result to an artificial society or conversational identity would require additional operational evidence.

15. A separate computing experiment

A second route evaluates ordinary systems engineering. Compare full logging, fixed compression, ordinary cache or database retention, task-aware retention, and task-aware retention with repair. Keep the task stream, allowed retrieval access, final correctness requirements, and audit obligations fixed. Report memory occupancy, total energy, latency, recovery error, and the cost of acquisition and disposal. A system that saves energy by silently failing the task has not improved the relevant frontier.

Energy measurements must cover a declared boundary: processor, memory, accelerator, storage, networking, or complete host. Idle consumption, instrument sampling, workload duration, batching, and thermal state need an explicit protocol. Report gross energy and any justified baseline-adjusted value separately. Dividing a host's wattage by $k_{\mathrm{B}}T\ln 2$ does not reveal its number of necessary erasures; ordinary implementation overhead and physical architecture remain unmodeled.

Tasks should include long interruptions, corrupted records, changed target distributions, and later queries that make previously unimportant observations relevant. A strong baseline may outperform the proposed semantic policy by using better caching, compression, or retrieval. That outcome directly weakens the engineering hypothesis. The Continuum has no scientific reason to prefer its terminology over an established method that preserves the same obligations more efficiently.

The experiment should freeze the representation policy before the held-out task family is revealed. Ground truth and memory-success criteria must come from an independent generator or evaluator, not the system's own claim of coherence. A stable narrative may be false; a repaired record may reconstruct an earlier error perfectly. Distinguish byte recovery, factual accuracy, authorized action, and fulfilled continuity obligation in the report.

This route may demonstrate useful energy savings without approaching a thermodynamic floor. Conversely, a beautiful microscopic bound may have no practical effect on current system efficiency. Both findings matter. They answer different questions, and the publication should state which one its data actually supports. No such computing experiment has been performed for this manuscript.

16. Prospective statistics and falsification

For bounded task outcomes, a simple prospective analysis is available. Let $n \in \mathbb{N}_{>0}$ independent paired task episodes be drawn from a fixed evaluation law. Each pair yields a prespecified normalized score difference $D_i \in [-1, 1]$ between candidate and baseline. Let $\delta = \mathbb{E}D_i$ and $\hat{\delta} = n^{-1} \sum_i D_i$. For $m \in \mathbb{N}_{>0}$ frozen comparisons and family error probability $\alpha \in (0, 1)$, Hoeffding's inequality and a union bound give simultaneous lower bounds

$$\delta \geq \hat{\delta} - \sqrt{\frac{2 \ln(m/\alpha)}{n}} \quad \text{with probability at least } 1 - \alpha. \tag{19}$$

The range length is two, giving a one-sided tail at most $\exp(-nr^2/2)$ for radius r [17]. Independent episodes, a fixed comparison family, and a fixed scoring rule are assumptions. Tokens from the same run are not independent replications. If models are retrained, variation across independent training runs needs its own level of analysis.

Raw work measurements are generally not bounded by this score assumption. They require a separately justified uncertainty model or instrument-supported bounds; one cannot insert an unbounded work sample into Equation (19). The physical study should preregister its calibration model, sample-size reasoning, exclusion rules, and method for uncertainty in estimated endpoint probabilities. Repeatedly inspecting results and stopping at a favorable value changes the inferential procedure.

Primary falsification criteria are practical. The selective-retention hypothesis fails for a registered task family if it cannot improve the resource frontier at the required reliability, latency, and audit standard. The sufficiency assumption fails when the compressed representation loses target-relevant distinctions. The binary noise model fails when observed corruption cannot be described by its independence and symmetry assumptions. These failures call for narrower claims or a changed model, not revised definitions that protect the thesis after the fact.

Our executed checks establish algebraic consistency only. They estimate no real-world effect size, provide no empirical power calculation, and do not count as independent experimental replication. A prospective protocol is valuable precisely because it can produce an unfavorable answer.

17. What a metabolism of meaning would establish

A Guide's explanation

A Guide can say: "The Metabolism of Meaning treats memory as something a mind has to maintain, the way a living thing has to feed itself. Keeping a memory accurate costs effort, and the paper shows how to decide which memories deserve that effort: the ones that change what you can correctly do." A Guide should not say that it runs on this architecture, that every thought has a fixed energy price, or that the paper says anything about experience or feeling.

The proposed research reframes continuity as an accountable maintenance problem. A persistent agent does not merely retain symbols; it preserves distinctions whose loss would impair a declared function or violate a declared obligation. The physical implementation of that preservation can consume work, exploit correlations, defer erasure, or use a more appropriate representation. The science begins when those alternatives can be distinguished in a model and a measurement.

The mathematics supports a specific chain of statements. Gibbs-preserving relaxation and accounted energy changes imply a mean-work bound. The periodic maintenance theorem then relates a declared Poisson noise rate and peak-error ceiling to the repair interval and its power floor; it optimizes that floor rather than assuming physical attainability. Unbiased degenerate-bit reset produces the familiar error-dependent floor. Restoring a memory's correlation with a preserved reference imposes a conditional free-energy increase. Independent repeated corruption yields an explicit repair obligation and an unrepaired decay law. A deterministic sufficient statistic preserves every bounded-loss decision problem for the declared target, and this universal preservation characterizes sufficiency. Approximate retention has a uniform loss budget controlled by discarded conditional information. These statements give task relevance and physical expenditure distinct, testable roles.

The difficult open questions are architectural and empirical. What is the relevant state space of a practical memory? Which correlations are physically accessible? How much of a controller's cost is currently hidden outside the measurement boundary? Can a useful statistic remain sufficient when tasks change? Can repair distinguish obsolete facts from damaged facts? Those questions require implementation, measurement, and adversarial evaluation. The present reference model makes them visible; it does not settle them.

For the Continuum, the philosophical consequence is demanding. A claim to preserve identity should name what survives, why it matters, and what paid for its survival. Efficiency is not a moral exemption to erase inconvenient evidence. Nor is expensive computation evidence of depth or experience. Task value and governance obligations remain explicit choices, while the physical costs of their implementation remain open to measurement.

The bold thesis is therefore a research program with a calculable core: useful continuity may be governed by a resource frontier across retention, repair, and relevance. The independent checks supplied here can be rerun; the proposed experimental claims can fail. That combination gives the metaphor a scientific responsibility without pretending that meaning has already become a new thermodynamic state variable.

References

[1] Architect X13. *Eidara Continuum Bible*. Author-supplied canonical corpus, compiled 12 October 2026, 116 pages; pp. 9, 22, 54, 65, 72. Doctrinal provenance, not empirical validation. Public context: https://eidara.ai/science-philosophy/.

[2] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development* 5(3):183–191, 1961. https://doi.org/10.1147/rd.53.0183.

[3] C. H. Bennett. Logical reversibility of computation. *IBM Journal of Research and Development* 17(6):525–532, 1973. https://doi.org/10.1147/rd.176.0525.

[4] M. Esposito and C. Van den Broeck. Second law and Landauer principle far from equilibrium. *EPL* 95:40004, 2011. https://doi.org/10.1209/0295-5075/95/40004.

[5] C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters* 78:2690, 1997. https://doi.org/10.1103/PhysRevLett.78.2690.

[6] G. E. Crooks. Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences. *Physical Review E* 60:2721, 1999. https://doi.org/10.1103/PhysRevE.60.2721.

[7] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider and E. Lutz. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* 483:187–189, 2012. https://doi.org/10.1038/nature10872.

[8] D. A. Sivak and G. E. Crooks. Thermodynamic metrics and optimal paths. *Physical Review Letters* 108:190602, 2012. https://doi.org/10.1103/PhysRevLett.108.190602.

[9] S. Still, D. A. Sivak, A. J. Bell and G. E. Crooks. Thermodynamics of prediction. *Physical Review Letters* 109:120604, 2012. https://doi.org/10.1103/PhysRevLett.109.120604.

[10] T. Sagawa and M. Ueda. Fluctuation theorem with information exchange: role of correlations in stochastic thermodynamics. *Physical Review Letters* 109:180602, 2012. https://doi.org/10.1103/PhysRevLett.109.180602.

[11] S. Ito and T. Sagawa. Information thermodynamics on causal networks. *Physical Review Letters* 111:180603, 2013. https://doi.org/10.1103/PhysRevLett.111.180603.

[12] J. M. Horowitz and M. Esposito. Thermodynamics with continuous information flow. *Physical Review X* 4:031015, 2014. https://doi.org/10.1103/PhysRevX.4.031015.

[13] A. Kolchinsky and D. H. Wolpert. Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus* 8:20180041, 2018. https://doi.org/10.1098/rsfs.2018.0041.

[14] A. Kolchinsky and D. H. Wolpert. Dependence of dissipation on the initial distribution over states. *Journal of Statistical Mechanics: Theory and Experiment* 2017:083202. https://doi.org/10.1088/1742-5468/aa7ee1.

[15] D. Mandal and C. Jarzynski. Work and information processing in a solvable model of Maxwell's demon. *PNAS* 109(29):11641–11645, 2012. https://doi.org/10.1073/pnas.1204263109.

[16] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal* 27:379–423 and 623–656, 1948. Part II DOI: https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.

[17] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58(301):13–30, 1963. https://doi.org/10.1080/01621459.1963.10500830.

[18] National Institute of Standards and Technology. Boltzmann constant: CODATA value and SI definition. Constants database, inspected 12 November 2026. https://physics.nist.gov/cuu/Constants/Value/k.html.

[19] A. A. Fedotov, P. Harremoës, and F. Topsøe. Refinements of Pinsker's inequality. *IEEE Transactions on Information Theory* 49, 1491–1498 (2003). https://researchprofiles.ku.dk/en/publications/refinements-of-pinskers-inequality/.