Capacity Overflow, Effective Stochasticity, and Phase-B Invariants: Critical Deficit, Markov Closure, and Invariant Selection under Finite Projection

Yining Wu
Independent Researcher
yining.wu@alumni.upenn.edu
(Dated: May 2026)

FDS-T1 defines finite-observer distinguishability budgets. FDS-O1 turns those budgets into measurement capacity, and FDS-O2 turns finite record update into register time. FDS-T3 abstracts the common mechanism: capacity overflow. When task-relevant distinction demand exceeds accessible capacity, a finite observer or finite system cannot track all distinctions needed for full-fidelity prediction. The missing distinctions re-enter the accessible description as coarse-graining, effective stochasticity, hysteresis, false invariants, externalization demand, or failure. This paper formulates effective stochasticity as structured dynamics viewed through non-injective finite projection. Even when dynamics on a larger state space is deterministic, the induced dynamics on the accessible record space can be stochastic because multiple inaccessible states map to the same visible state while having different successors. This version strengthens the phase-transition layer: it defines the overflow deficit, predictive susceptibility, effective Markovian closure error, informational hysteresis, observer hierarchy, and a maintenance-cost selection score for Phase-B invariants. Phase-B invariants are coarse variables that remain cheap to update, slow to forget, and approximately Markovian after microstate tracking fails. Deterministic simulations illustrate projection-induced stochastic kernels, critical deficit susceptibility, Markov closure, invariant selection, irrecoverable post-overflow loss, capacity-relative stochasticity, Phase-A/Phase-B transition, and semantic drift in finite-context AI-like systems. The paper does not claim that all randomness is epistemic, that quantum randomness is derived, that topology is required for all persistence, or that fundamental conservation laws are reduced to observer capacity. Its narrower contribution is a finite-capacity bridge model: effective stochasticity is the shadow cast by finite projection, and law-like structure is the residue of distinctions that remain cheap, predictive, and slow to forget under overflow.

Scope and Boundary of the Theory. This paper studies effective stochasticity relative to finite distinguishability budgets. It does not deny the usefulness of stochastic models, does not claim that all physical randomness is reducible to ignorance, does not derive quantum measurement probabilities, does not modify unitary quantum mechanics, and does not assert that all persistent structures require topological protection. It treats stochasticity operationally: a finite system sees a transition as stochastic when its accessible record state does not carry the distinctions needed to make the transition deterministic. The result is compatible with deterministic, stochastic, quantum, and thermodynamic descriptions on larger or different state spaces.

Claim-status summary

Table I summarizes the central FDS-T3 claims, their epistemic status, and the conditions under which they should be weakened or rejected.

Keywords: capacity overflow; effective stochasticity; finite observer; finite distinguishability; coarse-graining; projection; stochastic kernels; hidden variables; Markov closure; predictive susceptibility; Phase-B invariants; hysteresis; semantic drift; information thermodynamics; active finite distinction systems.

INTRODUCTION

From finite budgets to overflow

FDS-T1 defines an observer-relative distinguishability budget: a finite physical observer cannot register, preserve, update, and operationally use unlimited distinctions [1]. FDS-O1 applies this budget to measurement, showing that stable record formation is limited by sensor, channel, memory, buffer, update, and thermodynamic bottlenecks [2]. FDS-O2 applies the same idea to time, treating usable temporal order as causally ordered finite-record update [3]. Both operational papers use the same structure: a task requires distinctions; the finite system can carry only some of them; when demand exceeds capacity, full-fidelity tracking exits into compression, delay, state merging, externalization, reset, or failure.

T3 abstracts the common mechanism. Capacity overflow occurs when the distinctions required for prediction, control, or persistence exceed the distinctions accessible to a finite system. The system can still operate, but only after projecting the larger state space into a smaller accessible record space. That projection can be many-to-one. When it is, the same visible state can correspond to multiple hidden states with different successors. To the finite system, the next visible state is no longer determined by the current visible state alone. The dynamics becomes effectively stochastic.

TABLE I. Central FDS-T3 claims, epistemic status, and demotion or failure conditions.

Claim	Status	What would weaken or falsify it
Capacity overflow occurs when task-relevant distinction demand exceeds accessible capacity.	Operational criterion	Full-fidelity tracking persists under bounded resources after independently estimated demand exceeds all accessible capacity.
Non-injective projection induces effective stochasticity.	Conditional theorem	A many-to-one accessible projection always induces deterministic accessible transitions despite hidden successors varying within the same visible record class.
Overflow has a critical-deficit signature.	Testable prediction	Predictive error, transition entropy, and exit signatures vary smoothly through controlled capacity crossing with no susceptibility peak, kink, or rapid regime change.
Phase-B variables are selected by low update cost, slow information decay, and approximate Markov closure.	Selection principle	All coarse variables lose predictive information at the same rate under projection, regardless of update cost, closure error, or persistence utility.
Capacity recovery need not reverse overflow.	Conditional prediction	Discarded distinctions are perfectly reconstructed after capacity recovery without external logs, hidden reservoirs, or additional records.
Stochastic descriptions are capacity-relative unless supported by capacity-independent noise sources.	Scope claim	The same process yields the same stochastic description independent of observer capacity, projection, measurement boundary, and retained memory.
Long-context drift is a domain projection of wrong invariant completion under overflow.	Engineering projection	Context overflow in finite-window systems never increases false dependency, semantic drift, wrong task-state completion, or external-memory demand under matched tasks.

Main thesis

The main thesis is:

$$\text{effective stochasticity} \equiv \text{structured dynamics viewed after finite projection.} \tag{1}$$

This is an operational thesis, not a metaphysical thesis. It does not say that the underlying world is deterministic. It does not say that quantum randomness is merely ignorance. It says that, for any finite observer or finite register, lost distinctions can appear as transition probabilities because the accessible state is no longer sufficient to determine the accessible successor.

The stronger contribution of this version is the selection layer. Overflow does not merely destroy detail. It selects coarse variables. Some variables disappear. Some become noisy. Some remain predictive because they are cheap to update, slow to forget, and nearly closed under the projected dynamics. Those variables are called Phase-B invariants.

Contributions

The paper makes seven contributions.

1. It defines capacity overflow as a positive deficit between task-relevant distinction demand and accessible capacity.
2. It derives an induced stochastic kernel on the accessible record space from deterministic dynamics on a larger state space under non-injective projection.
3. It defines predictive susceptibility as an order-parameter-like signature near the critical capacity deficit.
4. It defines Markov closure error and uses it to explain why some coarse variables become law-like after overflow.
5. It defines a maintenance-cost selection score for Phase-B invariants.
6. It formalizes informational hysteresis: capacity recovery restores possible storage, not distinctions already discarded.
7. It provides deterministic synthetic simulations and reproducible code.

Relation to earlier stochasticity drafts

An earlier stochasticity draft argued that truncation generates an unmodeled environment operationally in-distinguishable from noise, then connected that noise to topological persistence and signal-to-noise collapse. It also emphasized that non-topological structures can lose local signal-to-noise ratio while boundary or topological structures may preserve finite signal. The present T3 version narrows the claim. It does not assert that all stochasticity is forced by truncation, that all persistent structures require topology, or that effective Markovianity follows automatically from ignorance. Instead, it formulates the safer bridge: finite non-injective projection induces observer-relative stochastic kernels, and some coarse variables remain stable under overflow. Topological protection is treated as one possible domain-specific example of Phase-B invariance, not as the universal form of persistence.

RELATED WORK

Coarse-graining, rate-distortion, and stochastic descriptions

Shannon information theory and rate-distortion theory provide the standard language for capacity, compression, and distortion [4–7]. Jaynes' maximum entropy approach explains how probability assignments can arise from constrained information [8]. In statistical mechanics and nonequilibrium theory, coarse-graining and projection can generate effective irreversibility and stochastic evolution when unresolved degrees of freedom are integrated out [9–12]. FDS-T3 is continuous with this tradition, but its emphasis is the finite-observer capacity ledger: stochasticity is not only a mathematical convenience; it can be the accessible image of budget crossing.

Hidden Markov structure and Markov closure

Non-injective observation of a larger process often produces hidden-state dynamics. The visible process may fail to be Markov even if the underlying process is Markov or deterministic, because unresolved hidden variables carry memory. This is the standard lesson of hidden Markov models and predictive-state approaches [13, 14]. Recent work on quantitative coarse-graining of Markov chains provides error bounds for effective dynamics after projection, reinforcing the need to distinguish useful Markov closure from uncontrolled hidden-memory loss [28]. T3 therefore does not assume that projection automatically produces a clean Markov process. It asks which coarse variables minimize the residual memory required for prediction.

Renormalization and effective variables

Renormalization describes how large-scale behavior can be governed by variables that survive changes of scale while microscopic details are discarded [15, 16]. FDS-T3 uses a similar intuition in a finite-register setting. Phase-B invariants are not necessarily renormalization fixed points, but they play an analogous operational role: they are variables that remain predictive after repeated projection eliminates finer distinctions.

Information thermodynamics and memory reuse

Landauer's principle links logically irreversible erasure to thermodynamic cost [17]. Bennett showed that computation can be arranged reversibly only when enough memory and garbage management are available [18]. Modern information thermodynamics refines these ideas in feedback, correlations, stochastic dynamics, and nonequilibrium systems [12, 19, 20]. Recent reviews emphasize that practical erasure and computation often require accounting beyond the ideal single-bit Landauer bound, including finite-time, finite-bath, nonequilibrium, non-Markovian, quantum, and error-correction effects [32]. T3 does not claim that every stochastic kernel dissipates heat. The heat cost enters when finite projection is physically implemented by overwrite, reset, many-to-one compression, or garbage collection.

Topological, boundary, and symmetry-stabilized examples

Topological phases illustrate how some coarse invariants can remain stable under local perturbation [21–24]. Non-Hermitian systems and skin-effect models show that boundary localization and non-reciprocal dynamics can produce robust boundary signatures [25–27]. T3 does not require these examples, but they motivate the term Phase-B invariant: a coarse, distributed, or boundary-stabilized distinction that survives after local microstate tracking fails. In a broader language, Phase-B invariants can be interpreted as coarse-grained symmetries of the projected dynamics. This paper does not derive fundamental conservation laws from finite capacity; it only supplies an operational bridge for why certain coarse variables become stable carriers of prediction.

CAPACITY, DEMAND, AND OVERFLOW

Definition 1 (Underlying state space and accessible record space)**.** *Let $\mathcal{X}$ be a state space representing distinctions available in a larger model, and let $\mathcal{Z}_C$ be the* *record space accessible to a finite observer with capacity C. The observer implements a projection*

$$\pi_C : \mathcal{X} \to \mathcal{Z}_C, \qquad |\mathcal{Z}_C| \le 2^C. \tag{2}$$

The accessible record at update index t is $Z_t = \pi_C(X_t)$.

Definition 2 (Task-relevant distinction demand). *For task family Ψ, distortion tolerance ε, and window τ, let*

$$R^{(\tau)}_{\min}(\varepsilon; \Psi) \tag{3}$$

be the minimum number of task-relevant bits required to meet the predictive, control, or reconstruction target.

Definition 3 (Capacity overflow). *Given accessible capacity $C_{\text{acc}}(t)$, the T3 overflow deficit is*

$$\Delta_{\text{T3}}(t) = R^{(\tau)}_{\min}(\varepsilon; \Psi_t) - C_{\text{acc}}(t). \tag{4}$$

Overflow occurs when $\Delta_{\text{T3}}(t) > 0$ over the relevant window.

Overflow is not ordinary ignorance. Ordinary ignorance may be resolved by measurement or computation within the same boundary. Overflow is a boundary statement: the distinctions needed for the task are not carried by the accessible record space under the current capacity, update rate, memory, and externalization options.

Theorem 1 (Overflow-exit theorem). *Let O be a finite system attempting task Ψ over window τ. If $R^{(\tau)}_{\min}(\varepsilon; \Psi) > C_{\text{acc}}$, then full-fidelity task performance at distortion ε is impossible unless at least one exit occurs: increased capacity, longer time, buffering, externalization, lossy compression, coarse-graining, stochastic modeling, task relaxation, boundary enlargement, or failure.*

Proof. The demand term is the minimum number of distinctions required by the task. The capacity term is the maximum number of distinctions accessible to the system under the specified boundary and window. If demand exceeds capacity, a full-fidelity solution would require distinctions not represented by any admissible record state or update sequence. The system must therefore alter demand, alter capacity, store or externalize backlog, compress, accept stochastic residuals, relax the task, or fail. These alternatives exhaust the ways to remove the strict inequality under finite resources. □

The empirical content is not the inequality itself. The empirical content is the predicted pattern near crossing: transition entropy rises, reconstruction depth saturates, error floors appear, formerly deterministic variables become probabilistic, buffers fill, externalization increases, and coarse variables begin to dominate prediction.

Critical deficit and predictive susceptibility

To make the Phase-A/Phase-B distinction more physical, define a predictive error observable

$$E_P(\Delta) = \mathbb{E}[d(Z_{t+1}, \hat{Z}_{t+1}) \mid \Delta_{\text{T3}} = \Delta] \tag{5}$$

and a transition-entropy observable

$$H_P(\Delta) = H(Z_{t+1} \mid Z_t, \Delta_{\text{T3}} = \Delta). \tag{6}$$

The corresponding susceptibilities are

$$\chi_P = \frac{\partial E_P}{\partial \Delta_{\text{T3}}}, \qquad \chi_H = \frac{\partial H_P}{\partial \Delta_{\text{T3}}}. \tag{7}$$

The T3 prediction is not that every finite system has a thermodynamic singularity. It is weaker and more testable: under controlled sustained crossing, χ_P or χ_H should exhibit a peak, kink, or rapid increase near $\Delta_{\text{T3}} \approx 0$ whenever the discarded distinctions are dynamically relevant. Phase-A is the tracking regime $\Delta_{\text{T3}} < 0$; Phase-B is the post-overflow coarse regime $\Delta_{\text{T3}} > 0$.

PROJECTION-INDUCED EFFECTIVE STOCHASTICITY

From deterministic dynamics to stochastic kernels

Assume first that the larger state evolves deterministically:

$$X_{t+1} = F(X_t). \tag{8}$$

The finite observer does not access X_t directly. It accesses $Z_t = \pi_C(X_t)$. The same accessible record z can have a preimage

$$\pi_C^{-1}(z) = \{x \in \mathcal{X} : \pi_C(x) = z\}. \tag{9}$$

If different $x \in \pi_C^{-1}(z)$ lead to different accessible successors, then the accessible transition is stochastic relative to z.

Definition 4 (Induced stochastic kernel). *Given a prior or conditional distribution $P(x_t \mid z_t)$ over hidden states compatible with record z_t, the induced accessible transition kernel is*

$$P_C(z' \mid z) = \sum_{x \in \pi_C^{-1}(z)} P(x \mid z)\, \mathbf{1}\{\pi_C(F(x)) = z'\}. \tag{10}$$

For stochastic dynamics $P(x' \mid x)$, the corresponding kernel is

$$P_C(z' \mid z) = \sum_{x \in \pi_C^{-1}(z)} \sum_{x' \in \pi_C^{-1}(z')} P(x' \mid x) P(x \mid z). \tag{11}$$

Theorem 2 (Projection-induced stochasticity)**.** *If there exist $x_1, x_2 \in \pi_C^{-1}(z)$ such that $\pi_C(F(x_1)) \neq \pi_C(F(x_2))$, then $P_C(\cdot \mid z)$ is non-degenerate for any conditional distribution assigning positive probability to both x_1 and x_2. Thus a deterministic underlying map can induce stochastic accessible dynamics.*

Proof. The two hidden states share the same accessible current record but have different accessible successors. Since the observer conditions only on z, both hidden states remain compatible. If both have positive conditional probability, at least two successor records have positive probability. Therefore the induced kernel is non-degenerate. □

Non-injectivity loss and transition entropy

A useful measure of lost hidden distinction is

$$\mathcal{L}_\pi = H(X_t \mid Z_t). \tag{12}$$

A useful measure of effective stochasticity is

$$\mathcal{S}_{\text{eff}} = H(Z_{t+1} \mid Z_t). \tag{13}$$

The two are related but not identical. Projection loss measures unresolved present distinctions. Effective stochasticity measures how much those unresolved distinctions matter for the next accessible state. If hidden distinctions are dynamically irrelevant, projection loss can be high while transition entropy remains low. If hidden distinctions control successor states, projection loss becomes effective stochasticity.

Ontic noise, external noise, and effective noise

FDS-T3 distinguishes three cases.

1. *External noise*: stochastic forcing enters the larger state dynamics $P(x' \mid x)$.
2. *Projection-induced effective noise*: the larger dynamics may be deterministic, but non-injective projection makes the accessible dynamics stochastic.
3. *Modeling noise*: a deliberate approximation is used despite sufficient capacity in principle.

Only the second is the specific contribution of T3. The paper does not collapse all noise into observer ignorance.

PHASE-A, PHASE-B, AND MARKOV CLOSURE

Definition 5 (Phase-A tracking)**.** *A system is in Phase-A relative to task Ψ when $\Delta_{\text{T3}} \leq 0$ and the accessible record retains enough distinctions for full-fidelity task performance at tolerance ε. Phase-A is the tracking regime.*

Definition 6 (Phase-B dynamics)**.** *A system is in Phase-B relative to task Ψ when $\Delta_{\text{T3}} > 0$ persists and prediction or control is dominated by coarse variables, stochastic kernels, external records, or invariants rather than full microstate tracking. Phase-B is the post-overflow coarse regime.*

Effective Markovian closure

A non-injective projection generally produces hidden memory. Therefore T3 does not assume that the accessible process is strictly Markov. Define the k-step Markov closure error

$$\mathcal{M}_k(Z) = \mathrm{I}(Z_{t+1}; Z_{t-k:t-1} \mid Z_t). \tag{14}$$

If $\mathcal{M}_k \approx 0$, the coarse variable Z_t is approximately closed for prediction at horizon one. If $\mathcal{M}_k$ is large, the current coarse state still requires retained history.

Definition 7 (Phase-B invariant)**.** *Let $\phi : \mathcal{X} \to \mathcal{Y}$ be a coarse variable. ϕ is a Phase-B invariant over window τ if it satisfies three operational properties under overflow:*

1. *Projection stability: ϕ remains recoverable from the accessible record or from a low-cost external record.*
2. *Predictive persistence: $\mathrm{I}(\phi_t; \phi_{t+\tau})$ decays slowly relative to microstate detail.*
3. *Approximate closure: $\mathcal{M}_k(\phi)$ is small enough that future values are predicted from the current coarse record without needing the discarded microhistory.*

Proposition 1 (Phase-B variables are approximate Markov closures)**.** *Among variables available after sustained overflow, those that appear law-like to the finite observer are the variables that minimize residual memory burden while retaining predictive information. In symbols, Phase-B variables tend to have low $\mathcal{M}_k(\phi)$ and high $\mathrm{I}(\phi_t; \phi_{t+\tau})$ per update cost.*

Argument. A variable with high predictive information but high closure error requires hidden history; under overflow that hidden history is not available. A variable with low closure error but no retained predictive information is cheap but useless. Variables that remain usable after overflow must therefore balance predictive utility with low memory burden. This is an operational selection principle, not a claim of exact optimality in every physical system. □

This explains why macroscopic descriptions can be simple despite microscopic detail loss. The simplicity is not proof that microscopic distinctions were absent. It can be the result of a selection effect: only those coarse variables that close approximately under projection remain usable.

SELECTION PRESSURE FOR PHASE-B INVARIANTS

Maintenance-cost functional

Let Φ_C be the class of candidate coarse variables available under capacity C. For $\phi \in \Phi_C$, define a survival score

$$\mathcal{S}(\phi) = \frac{\mathrm{I}(\phi_t; \phi_{t+\tau})}{\mathcal{C}_{\mathrm{upd}}(\phi) + \mathcal{C}_{\mathrm{maint}}(\phi) + \lambda \mathcal{M}_k(\phi) + \varepsilon_0}. \tag{15}$$

Here $\mathcal{C}_{\mathrm{upd}}$ is update cost, $\mathcal{C}_{\mathrm{maint}}$ is maintenance or externalization cost, $\mathcal{M}_k$ is Markov closure error, λ weights memory burden, and $\varepsilon_0 > 0$ avoids singularities. Phase-B invariant candidates are those with high $\mathcal{S}$.

A phenomenological survival probability can be written as

$$P_{\mathrm{survive}}(\phi) \propto \exp\left[-\beta \frac{\mathcal{C}_{\mathrm{upd}}(\phi) + \mathcal{C}_{\mathrm{maint}}(\phi) + \lambda \mathcal{M}_k(\phi)}{\mathrm{I}(\phi_t; \phi_{t+\tau}) + \varepsilon_0}\right]. \tag{16}$$

This is a selection heuristic, not a universal law. It captures the central idea: Phase-B invariants are forced survivors. They are not necessarily the most detailed variables. They are the variables that remain cheap, predictive, and slow to forget when detail is unaffordable.

Invariants as coarse-grained symmetries

Finite projection defines equivalence classes

$$x \sim_C x' \quad \text{if} \quad \pi_C(x) = \pi_C(x'). \tag{17}$$

A coarse variable ϕ that is constant on these equivalence classes and stable under projected dynamics behaves as a coarse-grained symmetry:

$$\phi(x) = \phi(x'), \quad x \sim_C x', \qquad \phi_{t+1} \approx \phi_t. \tag{18}$$

This does not derive fundamental symmetries or conservation laws. It states a weaker operational fact: under finite projection, law-like structure is carried by variables that remain invariant across distinctions the observer can no longer afford to retain.

HYSTERESIS AND CAPACITY-RELATIVE RANDOMNESS

Informational hysteresis

Capacity recovery restores possible storage, not distinctions that were already discarded. Let Ω_{before} denote the set of histories distinguishable before overflow. During overflow, projection maps

$$\Omega_{\mathrm{before}} \to \Omega_{\mathrm{after}} = \pi_C(\Omega_{\mathrm{before}}). \tag{19}$$

When capacity later recovers, the system may expand the record space, but without external logs it expands from Ω_{after}, not from the original preimage. The irrecoverable loss can be written

$$\mathcal{H}_{\mathrm{irr}} = H(X_{\mathrm{past}} \mid Z_t, C_{\mathrm{recovered}}, L_{\mathrm{ext}}), \tag{20}$$

where L_{ext} denotes external logs or reservoirs. If L_{ext} does not contain the discarded distinctions, then $\mathcal{H}_{\mathrm{irr}} > 0$ even after C increases.

Proposition 2 (Informational hysteresis)**.** *If overflow induces a many-to-one projection and the discarded preimage distinctions are not preserved in an external or reversible record, then subsequent capacity recovery cannot reconstruct those distinctions from the current accessible record alone.*

This is logical irreversibility before it is thermodynamic irreversibility. Thermodynamic cost enters when the projection is physically implemented by irreversible overwrite, reset, compression, or garbage collection.

Observer hierarchy and relative stochasticity

Consider two observers O_a and O_b with capacities $C_a < C_b$ and projections

$$\pi_a : \mathcal{X} \to \mathcal{Z}_a, \qquad \pi_b : \mathcal{X} \to \mathcal{Z}_b. \tag{21}$$

If π_b retains distinctions discarded by π_a, then typically

$$H(Z^a_{t+1} \mid Z^a_t) \geq H(Z^b_{t+1} \mid Z^b_t), \tag{22}$$

for the same underlying process and comparable priors. The inequality need not hold in every artificial projection, but it gives the expected direction when additional capacity preserves dynamically relevant distinctions.

Randomness is therefore capacity-relative, not merely subjective. A phenomenon that is stochastic for a low-capacity observer can become more deterministic for a higher-capacity observer. This does not imply that all stochasticity is removable. Some noise sources may remain capacity-independent, and quantum randomness is outside the scope of this paper.

NUMERICAL MODELS AND SIMULATIONS

The simulations are deterministic synthetic demonstrations. They are not fits to physical detector data, biological data, quantum experiments, AI benchmarks, or human behavior. All figures and CSV outputs are generated by `code/generate_results.py`.

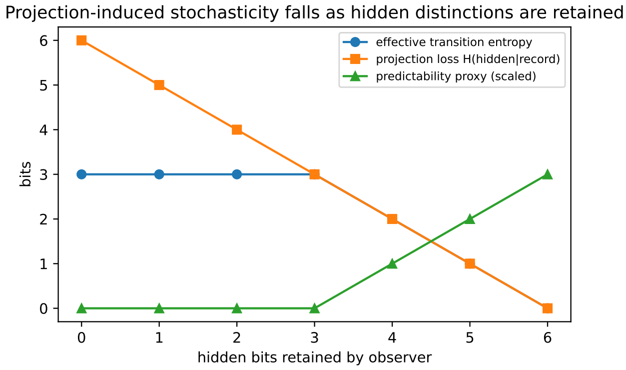

FIG. 1. Projection-induced stochasticity falls as hidden distinctions are retained. A deterministic hidden-state update appears stochastic to a low-capacity observer because unresolved hidden states inside the same visible record have different successors.

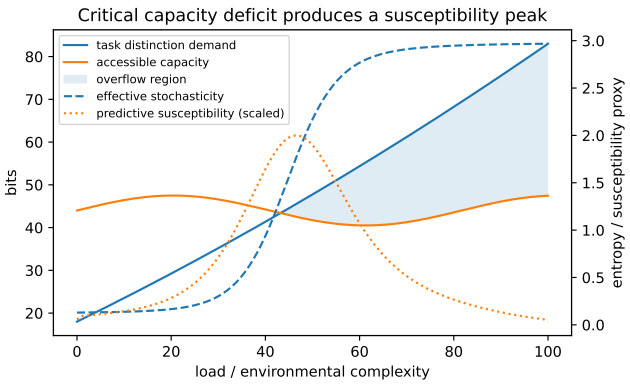

FIG. 2. Critical capacity deficit produces a susceptibility peak. When distinction demand crosses accessible capacity, effective stochasticity rises and the predictive susceptibility shows a sharp response.

Projection-induced stochasticity

The first simulation uses a deterministic hidden-state map. Observers retain different numbers of hidden bits. With few retained bits, the induced accessible transition has high entropy. As hidden distinctions are retained, stochasticity falls.

Critical deficit and susceptibility

The second simulation raises task demand through accessible capacity. Transition entropy and error rise near crossing; predictive susceptibility peaks near the critical deficit.

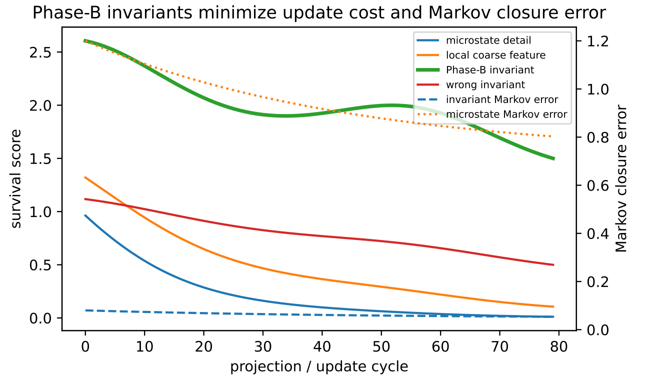

FIG. 3. Phase-B invariants minimize update cost and Markov closure error. The surviving variable is not the most detailed variable; it is the variable with the best ratio of predictive persistence to update and memory burden.

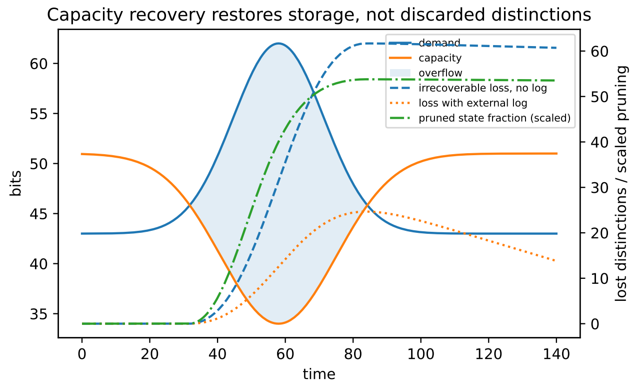

FIG. 4. Capacity recovery restores storage, not discarded distinctions. Overflow prunes the preimage. Without external logs, capacity recovery cannot reconstruct the distinctions already discarded.

Markov closure and invariant selection

The third simulation compares candidate variables under repeated projection. Microstate detail has high information but high update cost and high closure burden. The Phase-B invariant has lower detail but better survival score because it is cheaper and more Markov-closed.

Informational hysteresis

The fourth simulation imposes a temporary overload and then restores capacity. Without external logs, lost distinctions remain lost. External logging reduces hysteresis but costs storage and later cleanup.

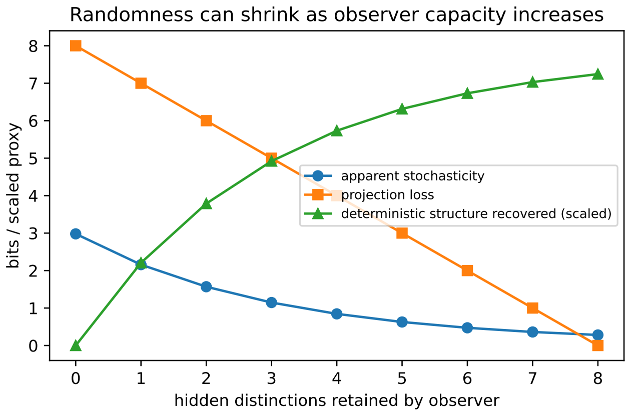

FIG. 5. Randomness can shrink as observer capacity increases. The same underlying process appears more stochastic to lower-capacity observers and more structured to higher-capacity observers.

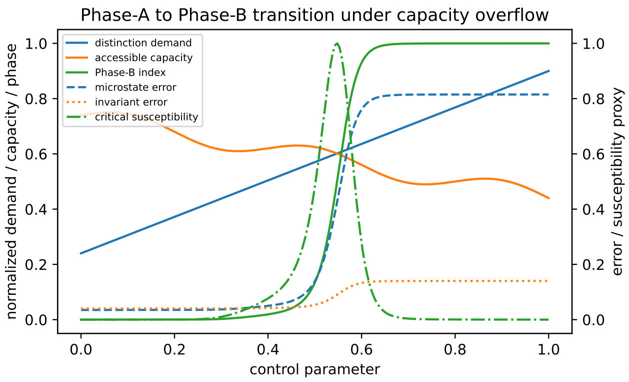

FIG. 6. Phase-A to Phase-B transition under capacity overflow. As demand crosses accessible capacity, microstate tracking fails, effective stochasticity rises, and prediction shifts toward Phase-B invariants.

Observer hierarchy

The fifth simulation varies retained hidden distinctions. Apparent stochasticity decreases as observer capacity increases, illustrating that stochastic descriptions can be capacity-relative.

Phase-A to Phase-B transition

The sixth simulation varies a control parameter that raises demand and lowers effective capacity. Once deficit becomes positive, the Phase-B index rises. Microstate error increases sharply while invariant error remains comparatively small.

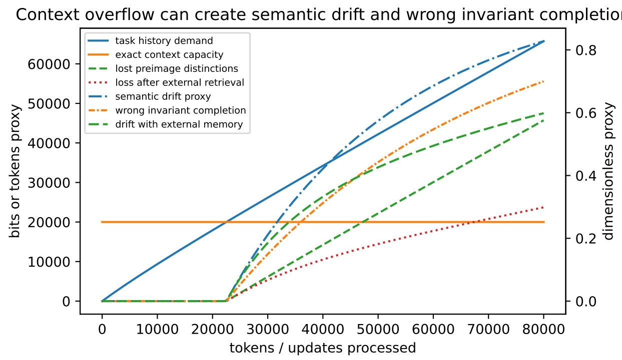

FIG. 7. Context overflow can create semantic drift and wrong invariant completion. In a finite-window system, overflow does not merely remove information; it can force the system to maintain coherence by selecting an incorrect coarse invariant. External retrieval increases effective capacity and pushes the system back toward Phase-A.

Semantic drift and wrong invariant completion

The seventh simulation treats a finite-context AI-like system as a domain projection. When task history exceeds context capacity, multiple histories collapse into the same accessible context. The system may complete the missing preimage with a wrong invariant, producing semantic drift or false dependency. External retrieval reduces but does not eliminate the loss.

DOMAIN-SPECIFIC PROJECTIONS

Measurement systems

O1 is the measurement-specific projection of T3. A detector with insufficient sensor, channel, memory, buffer, or update capacity cannot stabilize all task-relevant distinctions. The missing distinctions appear as state merging, false positives, false negatives, noise floors, delayed records, or externalized logs. T3 supplies the abstract stochastic-kernel language for those O1 exits.

Temporal registers

O2 is the time-specific projection of T3. When temporal demand exceeds temporal capacity, finite records lose order information. Non-injective temporal update produces compatible past histories and register-time collapse. T3 generalizes this as projection-induced effective stochasticity under overflow.

Biology and active pruning

A biological system cannot track every environmental detail. It must prune, coarse-grain, ignore, or externalize. Phase-B invariants in living systems may include boundary states, homeostatic variables, gradients, cycles, or control-relevant features that survive after microdetails are discarded. T3 does not complete a biological theory, but it explains why active pruning and invariant selection are not optional for finite organisms.

AI context overflow, RAG, and semantic drift

Finite-context AI systems provide an engineering analogy. When task history exceeds context capacity, the system must summarize, retrieve, compress, or discard. Empirical long-context studies show that language models do not always use nominally available context uniformly; performance can degrade when relevant information appears away from the beginning or end of the input [29]. Benchmark work such as RULER further shows that nominal context length can substantially exceed effective usable context under increased task complexity [30]. NoLiMa suggests that long-context performance can degrade sharply when retrieval requires latent association rather than literal matching, consistent with the T3 view that accessible context is not identical to effective task-relevant capacity [31]. The same underlying task can then appear more stochastic to the agent: next-step predictions become sensitive to inaccessible prior distinctions. Semantic drift occurs when the system maintains a locally coherent but wrong invariant across a collapsed preimage of possible histories. Hallucination is not only random guessing in this projection; it can be wrong invariant completion under context overflow. Retrieval-augmented generation and external memory can be interpreted as capacity externalization: they increase effective C_{acc} and can move the system back toward Phase-A. This is an engineering projection, not a claim that language models possess subjective experience or physical agency.

Topological and boundary cases

In some physical systems, topological or boundary-stabilized variables may act as Phase-B invariants. This paper does not prove that topology is necessary for persistence. It states a weaker bridge: when local microstate tracking fails, variables that remain recoverable under projection and stable under dynamics become the usable carriers of persistence. Topological invariants are one important family of such variables.

EXPERIMENTAL PROTOCOLS

Protocol 1 (Projection-induced stochasticity test)**.** *Choose a deterministic or high-fidelity simulated system with known state X_t. Define a family of projections π_C with decreasing capacity. Estimate the induced kernel $P_C(z' \mid z)$. T3 predicts that transition entropy should increase when discarded distinctions affect accessible successors.*

Protocol 2 (Critical deficit test)**.** *Independently estimate task demand $R_{\min}$ and accessible capacity C_{acc}. Increase demand or reduce capacity until Δ_{T3} crosses zero. Measure predictive error, transition entropy, and their finite-difference susceptibilities. T3 predicts a rapid response, kink, or susceptibility peak near crossing when discarded distinctions are dynamically relevant.*

Protocol 3 (Markov closure test)**.** *Track candidate coarse variables under repeated projection. Estimate $\mathcal{M}_k(\phi) = \mathrm{I}(\phi_{t+1}; \phi_{t-k:t-1} \mid \phi_t)$ and predictive information. T3 predicts that Phase-B invariants should have lower closure error and better survival score than ordinary microstate detail under overflow.*

Protocol 4 (Hysteresis test)**.** *Apply a temporary overload and then restore capacity. Compare systems with and without external logs. T3 predicts that capacity recovery will not fully restore discarded distinctions unless those distinctions were externalized or reversibly logged.*

Protocol 5 (Observer hierarchy test)**.** *Give observers or algorithms different hidden-state access to the same process. T3 predicts that accessible transition entropy should shrink as dynamically relevant distinctions are retained, while capacity-independent noise components remain.*

LIMITATIONS AND FALSIFICATION

First, T3 is an operational finite-capacity model, not a complete theory of randomness. It does not settle whether quantum randomness is ontic, epistemic, relational, or otherwise. Second, it does not prove that all environments are Markovian. Effective Markovianity requires additional assumptions about timescale separation, memory, and update capacity. Third, it does not claim that topology is required for every form of persistence. Phase-B invariance is broader than topology. Fourth, Landauer-style heat accounting applies only to physical implementations of logically irreversible overwrite, reset, compression, or garbage collection under the relevant thermodynamic assumptions. Fifth, the simulations are synthetic demonstrations, not empirical validations.

The strong version of FDS-T3 would be weakened or rejected by any of the following:

1. controlled capacity crossing with no change in transition entropy, reconstruction depth, error floor, state merging, externalization, buffering, susceptibility, or failure;
2. non-injective projections of deterministic dynamics that never induce non-degenerate accessible transition kernels when hidden successors differ;
3. finite systems that recover erased distinctions perfectly after capacity recovery without external logs, hidden reservoirs, or additional records;
4. repeated projection under which all variables lose predictive information at the same rate, with no invariant-like coarse variables surviving longer than local details;
5. candidate Phase-B variables having no advantage in Markov closure error, update cost, or predictive persistence over discarded details under matched capacity;
6. apparent stochasticity remaining unchanged when observer capacity is increased to include independently verified dynamically relevant hidden distinctions;
7. physically implemented irreversible record reuse below generalized Landauer accounting after reservoirs, correlations, feedback, and work sources are included.

CONCLUSION

A finite system does not experience overflow as a philosophical abstraction. It experiences overflow as loss of distinctions. Once those distinctions are absent from the accessible record, the same visible state can be compatible with multiple hidden states and multiple successors. To the finite system, dynamics becomes stochastic.

FDS-T3 therefore supplies the Structural Trident's third anchor. T1 states the finite-observer budget. O1 shows how measurement fails or exits under finite capacity. O2 shows how temporal order collapses under finite record update. T3 abstracts the shared mechanism: capacity overflow converts inaccessible distinctions into effective stochasticity, and repeated projection selects the invariants that remain useful after microstate tracking fails.

The most compact statement is this: effective stochasticity is the shadow cast by finite projection; stable law-like structure is the residue of distinctions that remain cheap, predictive, and slow to forget under overflow. Randomness is not simply the absence of law. It can be the appearance of law after the observer loses the distinctions that made the law deterministic. Conversely, law-like structure is not necessarily full information. It can be the survivor of a harsh selection process: a Phase-B invariant that remains approximately closed when most distinctions have become unaffordable.

Notation Summary

Simulation Parameters

The simulations are deterministic and use fixed synthetic parameters in `code/generate_results.py`. Figure 1 uses eight visible records and sixty-four hidden residues per visible record. Figure 2 sweeps environmental complexity and computes predictive susceptibility from the finite-difference response of error to deficit. Figure 3 compares four candidate variables with different update costs, closure errors, and information decay constants. Figure 4 uses a temporary overload pulse and compares no-log versus external-log recovery. Figure 5 varies retained hidden bits. Figure 6 sweeps a control parameter across the Phase-A/Phase-B boundary. Figure 7 uses a finite context-window proxy with optional external retrieval. No proprietary data, detector data, biological data, quantum data, AI benchmark data, or human-subject data are used.

Reproducibility Checklist

1. Code availability: all simulation code is included in the replication package.
2. Deterministic execution: the random seeds are fixed or randomness is absent.
3. Figure reproduction: running `python code/generate_results.py` regenerates all figures and CSV outputs.
4. Data status: all numerical outputs are synthetic demonstrations generated from the stated model.
5. Platform independence: the code uses standard Python scientific libraries.

Boundary of Applicability

FDS-T3 applies to finite systems that represent only a projection of a larger state space and whose discarded distinctions are dynamically relevant to prediction, control, or persistence. It does not apply to ideal mathematical observers with unbounded state access, nor does it settle irreducible stochasticity in fundamental physics. It is compatible with standard stochastic modeling, quantum theory, statistical mechanics, and thermodynamics when those theories are interpreted on their own state

TABLE II. FDS-T3 notation summary.

Symbol	Meaning
$\mathcal{X}$	larger state space containing distinctions not all accessible to the observer
$\mathcal{Z}_C$	accessible record space under capacity C
π_C	finite projection from $\mathcal{X}$ to $\mathcal{Z}_C$
X_t	larger state at time or update index t
Z_t	accessible record $\pi_C(X_t)$
C_{acc}	accessible distinguishability capacity
$R_{\min}^{(\tau)}(\varepsilon;\Psi)$	minimal task-relevant distinction demand
Δ_{T3}	T3 overflow deficit $R_{\min} - C_{\text{acc}}$
$P_C(z' \mid z)$	induced accessible stochastic kernel
$\mathcal{L}_\pi$	projection loss $H(X_t \mid Z_t)$
$\mathcal{S}_{\text{eff}}$	effective stochasticity $H(Z_{t+1} \mid Z_t)$
χ_P, χ_H	predictive and entropy susceptibilities near capacity crossing
$\mathcal{M}_k$	Markov closure error $\mathrm{I}(Z_{t+1}; Z_{t-k:t-1} \mid Z_t)$
$\mathcal{S}(\phi)$	survival score of candidate Phase-B invariant ϕ
Phase-A	full tracking regime before sustained overflow
Phase-B	post-overflow coarse dynamics regime
$\mathcal{H}_{\text{irr}}$	irrecoverable past distinction loss after capacity recovery

spaces. T3 adds a diagnostic layer: when accessible capacity changes, the amount and character of effective stochasticity should change if the added distinctions are dynamically relevant.

CODE AVAILABILITY

The simulation code used to generate Figs. 1–7 is included in the accompanying replication package under `code/generate_results.py`. Running the script regenerates all figures (PDF and PNG) and CSV tables in a single pass.

AI ASSISTANCE DISCLOSURE

AI-assisted tools were used for language polishing, structural feedback, LaTeX drafting support, and code-debugging assistance. The author reviewed and edited all content and remains responsible for all claims, references, simulations, and conclusions. No AI system is listed as an author.

[1] Y. Wu, "Finite Distinguishability Budgets and Maintenance Bounds for Physical Observers," Zenodo (2026), doi:10.5281/zenodo.20234249.

[2] Y. Wu, "Observer as a Finite Distinction Register: Measurement Capacity, Dynamic Bottlenecks, and Budget-Crossing Signatures," Zenodo (2026), doi:10.5281/zenodo.20248792.

[3] Y. Wu, "Time as Irreversible Distinction Update: Finite Records, Causal Ordering, and Register-Time Collapse," Zenodo (2026), doi:10.5281/zenodo.20249369.

[4] C. E. Shannon, "A mathematical theory of communication," Bell System Technical Journal **27**, 379–423 and 623–656 (1948).

[5] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," IRE National Convention Record **4**, 142–163 (1959).

[6] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Prentice-Hall (1971).

[7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley (2006).

[8] E. T. Jaynes, "Information theory and statistical mechanics," Physical Review **106**, 620–630 (1957).

[9] H. Mori, "Transport, collective motion, and Brownian motion," Progress of Theoretical Physics **33**, 423–455 (1965).

[10] R. Zwanzig, "Memory effects in irreversible thermodynamics," Physical Review **124**, 983–992 (1961).

[11] R. Zwanzig, *Nonequilibrium Statistical Mechanics*. Oxford University Press (2001).

[12] U. Seifert, "Stochastic thermodynamics, fluctuation theorem and molecular machines," Reports on Progress in Physics **75**, 126001 (2012).

[13] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proceedings of the IEEE **77**, 257–286 (1989).

[14] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: Pattern and prediction, structure and simplicity," Journal of Statistical Physics **104**, 817–879 (2001).

[15] L. P. Kadanoff, "Scaling laws for Ising models near T_c," Physics **2**, 263–272 (1966).

[16] K. G. Wilson, "The renormalization group: Critical phenomena and the Kondo problem," Reviews of Modern Physics **47**, 773–840 (1975).

[17] R. Landauer, "Irreversibility and heat generation in the computing process," IBM Journal of Research and Development **5**, 183–191 (1961).

[18] C. H. Bennett, "The thermodynamics of computation–a review," International Journal of Theoretical Physics **21**, 905–940 (1982).

[19] T. Sagawa and M. Ueda, "Generalized Jarzynski equality under nonequilibrium feedback control," Physical Review Letters **104**, 090602 (2010).

[20] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, "Thermodynamics of information," Reviews of Modern Physics **87**, 45–67 (2015).

[21] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, "Quantized Hall conductance in a two-dimensional periodic potential," Physical Review Letters **49**, 405–408 (1982).

[22] F. D. M. Haldane, "Model for a quantum Hall effect without Landau levels," Physical Review Letters **61**, 2015–2018 (1988).

[23] M. Z. Hasan and C. L. Kane, "Colloquium: Topological insulators," Reviews of Modern Physics **82**, 3045–3067 (2010).

[24] X.-L. Qi and S.-C. Zhang, "Topological insulators and superconductors," Reviews of Modern Physics **83**, 1057–1110 (2011).

[25] N. Hatano and D. R. Nelson, "Localization transitions in non-Hermitian quantum mechanics," Physical Review Letters **77**, 570–573 (1996).

[26] S. Yao and Z. Wang, "Edge states and topological invariants of non-Hermitian systems," Physical Review Letters **121**, 086803 (2018).

[27] F. K. Kunst, E. Edvardsson, J. C. Budich, and E. J. Bergholtz, "Biorthogonal bulk-boundary correspondence in non-Hermitian systems," Physical Review Letters **121**, 026808 (2018).

[28] B. Hilder and U. Sharma, "Quantitative coarse-graining of Markov chains," *SIAM Journal on Mathematical Analysis* **56**, 913–954 (2024).

[29] N. F. Liu et al., "Lost in the middle: how language models use long contexts," *Transactions of the Association for Computational Linguistics* **12**, 157–173 (2024).

[30] C.-P. Hsieh et al., "RULER: what's the real context size of your long-context language models?" arXiv:2404.06654 (2024).

[31] A. Modarressi et al., "NoLiMa: long-context evaluation beyond literal matching," arXiv:2502.05167 (2025).

[32] P. Chattopadhyay et al., "Landauer principle and thermodynamics of computation," *Reports on Progress in Physics* **88**, 086001 (2025).