Preserving Declared Capabilities Across Representation Replacement: Executable Decoders, Adversarial Migration, and Trajectory Audits

Brian Theory
AI-assisted research draft

13 November 2026

Classification: technical synthesis note, version 3.0. The propositions are proved below, but their underlying coding, covering, coupling, and concentration principles are standard. No priority or external peer-review claim is made.

Abstract

An internal representation may be replaced without preserving its literal symbols. The question is which declared tests still receive the same answers. We give a finite contract for replacement between different memory spaces, separating preservation of an old predictive profile from accuracy about the world, preservation of interventions, and recovery of historical facts. For a fixed coarse migration and a trusted record that may inspect the old memory, we characterize the least record alphabet by per-fiber covering numbers of task profiles. The exact case reduces to counting distinct profiles within each fiber; a scalar Bernoulli specialization admits an exact partition algorithm and a nine-state example. For continuing controlled dynamics we prove the familiar product bound on adaptive transcript error under a uniform one-step simulation condition. Finite executable menus and adversarial coarse outputs give explicit covering and partition contracts. A visit-index confidence guarantee permits one evolving trajectory with observed cells, without resetting named states, provided a within-cell oscillation bound is certified. A protected-snapshot validator makes the actual-origin assumption explicit, and a sharp accepted-risk bound exposes rare-acceptance amplification. A conditional block-visitation premise supplies an audit deadline; stationary mass alone cannot. These additions are tested in an executable finite model, not a deployed secure system. Exact enumeration checks the small cases and supplies editable figures. The useful contribution is a self-contained operational specialization of established mathematics, not a new general theory of semantic continuity or personal identity.

1 Problem, baseline, and contribution boundary

An agent may reorganize a memory, replace a module, change a vocabulary, or compress a learned state. A useful preservation requirement should say which queries must remain answerable and under which future actions. The requirement need not preserve the name of an internal variable. Conversely, keeping a name or producing the same narrative does not recover information that a many-to-one map erased.

The nearest B-Theory baseline is Theory [14], the maintained 8 November 2026 manuscript. It studies endomorphisms of one finite Moore machine, robust registers, and trusted records depending on the replacement. Its Section 1 explicitly excludes replacement by a different machine. Its Section 4 already obtains a record optimum through functional characteristic graphs. We therefore make no new claim for the need to account for side information. Here the spaces may differ, the encoder sees the old memory, and the primary requirement is a declared prediction profile, possibly only

approximately preserved. These changed interfaces explain both the simpler exact counting formula and the different approximation problem.

Exact transformations between causal models already require agreement under a specified intervention map; Rubenstein et al. [13, Definition 3 and Lemma 5] provide a relevant antecedent. We do not identify arbitrary memory maps with their structural-equation transformations. For evolving systems, bounds from one-step discrepancies to path-law discrepancies are established in Abate et al. [1, Theorem 2]. Their total-variation norm ranges over $[0, 2]$; ours is half that norm and ranges over $[0, 1]$. Substituting $c_i = 2\delta_i$ and dividing their path bound by two gives $1 - \prod_i (1 - \delta_i)$ in our convention. Approximate bisimulation has a corresponding trace bound in Bian and Abate [4, Theorems 4–6]. These results make it especially important to avoid advertising a telescoping or coupling estimate as a new discovery.

The original core supplies four things: an explicit replacement interface; an exact finite record calculation with a scalar approximation algorithm; an adaptive test guarantee with all conditioning exposed; and verification limits for a learned translator. The manuscript is independently readable, but it is best viewed as a continuation of the recoverable-memory line. The extensions below remove selected idealizations under explicit replacement assumptions; they remain finite specializations, not a certified new research contribution.

2 Operational contract

Let M be a finite nonempty old-memory space and let $R \subseteq M$ be a nonempty declared set of reachable memories. These are internal states available to the migration procedure, not hidden world states. A finite nonempty family $\mathcal{T}$ names operational tests. Test τ has a finite outcome alphabet Ω_τ and old predictive distribution $P_m^\tau \in \Delta(\Omega_\tau)$ for $m \in R$. The phrase predictive distribution refers to the old model's prediction. Its empirical truth is a separate obligation.

A known deterministic migration $\phi : R \to M'$ supplies a coarse new state. During one copy operation an encoder may inspect m and write a trusted side record $z(m) \in [L] = \{1, \dots, L\}$. Afterwards the decoder sees only $(\phi(m), z(m))$ and the requested test τ. It returns a distribution $Q^\tau_{\phi(m),z(m)}$. The original memory and temporary buffers are erased before the query arrives. Encoder, decoder, test semantics, and prediction tables are fixed in advance. They do not receive the realized hidden target.

The objective is worst-case profile error over reachable memories, not average distortion, code length, or runtime:

$$\epsilon(\phi, z, Q) = \max_{m \in R} \max_{\tau \in \mathcal{T}} d_{\mathrm{TV}}\big(P_m^\tau, Q^\tau_{\phi(m),z(m)}\big), \qquad d_{\mathrm{TV}}(p, q) = \tfrac{1}{2} \sum_\omega |p(\omega) - q(\omega)|. \tag{1}$$

Fixed-length storage is $\lceil \log_2 L \rceil$ side bits. A one-symbol record needs zero bits. The coarse new state, immutable prediction tables, temporary copying workspace, and computation are additional resources. The theorems optimize only L; they do not establish low total implementation cost. There is no noise, corruption, or adversary on the trusted record.

Each test is separately queried. No joint probability law across distinct tests is demanded. If a common latent model must generate all reported distributions, allowable centers below must be restricted to that model's profile set. That restriction may increase the optimum.

Historical, predictive, and causal tests differ. A historical test can ask for a bit retained before replacement. A predictive test can ask for a future outcome under one fixed policy. A causal test can specify a real intervention and its response. Merely adding the word causal to a prediction query does not supply the intervention semantics or access rights. Sections 5 and 6 introduce their own additional contracts rather than silently inheriting them from (1).

3 Exact preservation and trusted records

Define the profile $\sigma(m) = (P_m^\tau)_{\tau\in\mathcal{T}}$. On the product of simplices use the distance

$$d(\sigma, \sigma') = \max_{\tau\in\mathcal{T}} d_{\mathrm{TV}}(P^\tau, P'^\tau).$$

Let $F_y = \{m \in R : \phi(m) = y\}$ for each attained coarse state y.

Proposition 3.1 (Exact factorization; elementary)**.** *With no side record, zero error is possible exactly when σ is constant on every F_y. With an L-symbol state-informed trusted record, the least possible alphabet size for zero error is*

$$L_0 = \max_{y\in\phi(R)} |\sigma(F_y)|. \tag{2}$$

Proof. For fixed y, two memories assigned the same record symbol receive the same decoded profile. At zero error they must have identical profiles. Therefore each distinct profile in F_y requires a different symbol, giving the lower bound. For the upper bound, enumerate the distinct profiles in each fiber with symbols $1, \ldots, |\sigma(F_y)|$ and let the decoder use y to interpret that fiber's enumeration. Symbols may be reused in different fibers because y is visible. All feasible decoder inputs have a defined profile; outputs on infeasible inputs may be arbitrary. The no-record assertion is $L_0 = 1$. □

This is a factorization and pigeonhole argument, not an original coding theorem. The information available to the encoder is decisive. An encoder that sees only the replacement map and not m cannot generally create this record. That is the reason the replacement-record problem in the existing B-Theory paper has a different characteristic graph.

Example 3.2 (A forgotten distinction and a preserved one)**.** Let $m = (u, v) \in \{0, 1\}^2$ and $\phi(m) = u \oplus v$. If the only test asks for $u \oplus v$, no record is needed. If tests can ask for either u or v, each coarse fiber contains two distinct deterministic profiles, so one trusted bit is necessary and sufficient: retain u and reconstruct $v = \phi(m) \oplus u$. If ϕ erases both bits, four symbols are required. The same internal map therefore preserves different amounts of knowledge under different declared tests.

Example 3.3 (No finite sample can recreate erased certainty)**.** Suppose two memories m_0, m_1 have the same complete post-migration interface but a later deterministic historical query requires outputs $0, 1$ respectively. The same answer distribution is used in both cases. Its two error probabilities sum to one, so worst-case error is at least $1/2$. Repeating that identical interface supplies no new information. This concerns erased evidence, not an inference algorithm's quality.

4 Approximate preservation as a finite covering problem

Definition 4.1. For a nonempty finite profile set S, let $N_\epsilon(S)$ be the smallest number of centers in $\prod_\tau \Delta(\Omega_\tau)$ whose closed d-balls of radius ϵ cover S. Centers need not belong to S.

Theorem 4.2 (Fixed-migration record frontier; standard covering specialization)**.** *For $0 \le \epsilon \le 1$, the least record alphabet achieving* (1) *at most ϵ is*

$$L_\epsilon = \max_{y\in\phi(R)} N_\epsilon(\sigma(F_y)). \tag{3}$$

Proof. Any decoder gives at most L centers for a fixed y. Assigning each memory its encoder symbol makes those centers cover $\sigma(F_y)$, establishing necessity. Conversely choose an optimal cover in each fiber, encode a memory by an index of a center covering its profile, and decode using the pair (y, z). The finite sets admit such a selection and symbols can be reused across fibers. This attains the maximum on the right. At $\epsilon = 0$, distinct points require distinct centers, recovering (2). □

If two memories are assigned the same final pair, their profile distance is at most 2ϵ by the triangle inequality. This necessary diameter condition is not sufficient for arbitrary probability simplices.

Example 4.3 (Why half the diameter can fail)**.** One coarse fiber contains three memories predicting point masses at outcomes $1, 2, 3$. All pairwise total variation distances are one. For any common decoder q, the error at point mass i is $1 - q_i$, and some $q_i \leq 1/3$. The smallest maximum error is therefore $2/3$, attained by the uniform distribution, rather than half the diameter $1/2$.

There is a complete elementary solution for one Bernoulli query. Write its success probabilities in a fiber as $p_1 \leq \cdots \leq p_n$. Total variation is absolute difference. A group from p_i to p_j has smallest enclosing radius $(p_j - p_i)/2$, attained at the midpoint. An optimal assignment to ordered scalar centers can be chosen with contiguous groups: nearest-center cells on the line are intervals, with boundary ties assigned consistently.

Proposition 4.4 (Scalar partition recurrence)**.** *Let $V_\ell(j)$ be the minimum possible maximum error when the first j ordered probabilities use at most ℓ centers. Set $V_0(0) = 0$, $V_0(j) = +\infty$ for $j > 0$, and $V_\ell(0) = 0$. Then*

$$V_\ell(j) = \min_{0 \leq i < j} \max \left\{ V_{\ell-1}(i), \frac{p_j - p_{i+1}}{2} \right\}. \tag{4}$$

Proof. For an optimal contiguous grouping, let its last nonempty group contain indices $i + 1, \ldots, j$. Its radius is the second term in the maximum; the prefix needs at most $\ell - 1$ groups and therefore incurs at least the first term. Minimizing over the split gives a lower bound. Conversely concatenate an optimal prefix grouping and a midpoint center for the last group. Every candidate in the minimum is attainable, proving equality. Repeated centers or unused symbols allow fewer than ℓ groups. □

After sorting, the recurrence uses $O(\ell n^2)$ comparisons and arithmetic operations. This is an algorithmic bound for explicit scalar data, not for general multi-query total-variation covering. Values represented as arbitrary exact reals do not give a finite-bit implementation; our computation uses rational probabilities throughout.

Example 4.5 (A nine-state migration frontier)**.** Two fibers have Bernoulli profiles

$$F_A : \quad 0,\ 1/10,\ 1/5,\ 4/5,\ 9/10,\ 1, \qquad F_B : \quad 1/4,\ 1/2,\ 3/4.$$

The exact minimum errors with $L = 1, \ldots, 6$ symbols are $1/2, 1/8, 1/10, 1/20, 1/20, 0$. For $L = 2$, the first fiber costs $1/10$ but the second costs $1/8$ and determines the uniform guarantee. A fifth symbol does not improve the best uniform error; it is the number of distinctions and their geometry, rather than record entropy, that matters.

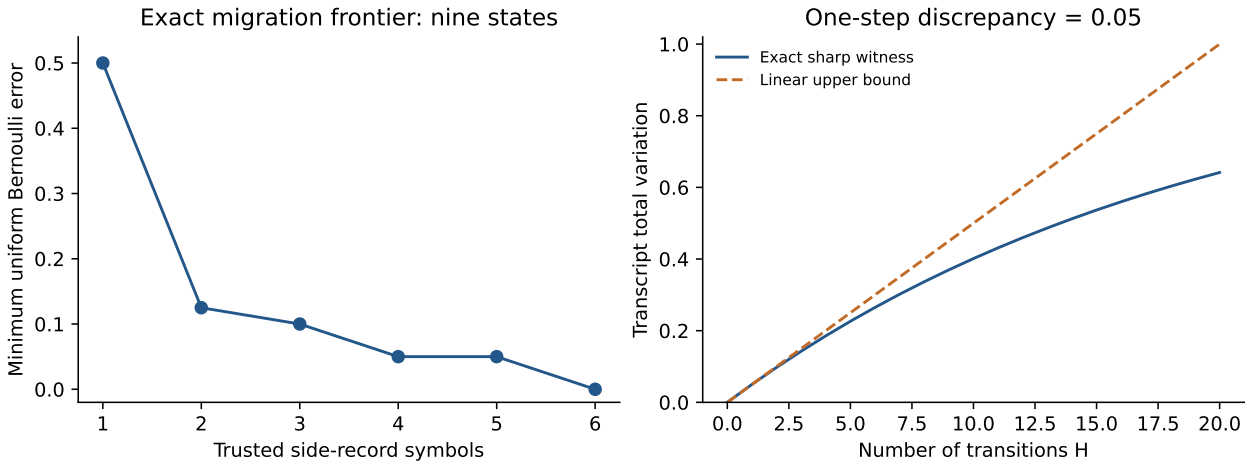


Figure 1: Original computations. Left: the exact scalar nine-state frontier. Right: a two-state process attains the standard product bound with one-step discrepancy 0.05. Curves are calculated from the models, not empirical data. Editable plotting source and exact rational data are included.

5 Repeated migrations and continuing actions

For a sequence of static migrations $\phi_i : M_i \to M_{i+1}$, suppose each stage has a profile σ_i in the same test space and $d(\sigma_i(m), \sigma_{i+1}(\phi_i(m))) \le \epsilon_i$ at every relevant reachable state. Repeated triangle inequalities give

$$d\big(\sigma_0(m), \sigma_k(\phi_{k-1} \cdots \phi_0(m))\big) \le \min\{1, \textstyle\sum_{i=0}^{k-1} \epsilon_i\}. \tag{5}$$

This elementary bound requires the same semantics at every stage. An empirical mean error under different training distributions cannot be substituted into it. Nor does agreement on one-step predictions, by itself, preserve a sequence of controlled interactions.

For the latter question, let the old and new systems have finite spaces M, M', common action alphabet A, transition kernels K_a, K'_a, and deterministic readouts $g : M \to O$, $g' : M' \to O$. A deterministic map $\phi : M \to M'$ connects them. Time starts with m_0 and $m'_0 = \phi(m_0)$. At time t, the observer sees $g(m_t)$, selects action a_t using the public readout/action history, and then the transition occurs. The new system uses the same policy on its own history. A policy may have private randomness, but the same independent random seed is used in a coupling. It cannot inspect either hidden memory. A horizon H means H actions and $H+1$ readouts.

Assume exact readout compatibility $g = g'\phi$ and, uniformly for every reachable m and permitted action,

$$d_{\mathrm{TV}}\big(\phi_{\#}K_a(m,\cdot), K'_a(\phi(m),\cdot)\big) \le \delta, \quad 0 \le \delta \le 1. \tag{6}$$

Here $\phi_{\#}$ denotes the distribution of the mapped old next state. The reachable domain must remain closed under every action allowed by the comparison, including actions an adaptive test might choose.

Theorem 5.1 (Adaptive transcript bound; coupling specialization)**.** *Under* (6)*, every common admissible policy π satisfies*

$$d_{\mathrm{TV}}\big(\mathcal{L}^{\pi}_{m}(O_0, A_0, \ldots, A_{H-1}, O_H), \mathcal{L}^{\pi}_{\phi(m)}(O'_0, A'_0, \ldots, A'_{H-1}, O'_H)\big) \le 1 - (1-\delta)^H. \tag{7}$$

Consequently any transcript loss taking values in $[0,1]$ changes in expectation by at most the same quantity.

Proof. Two finite distributions p, q can be coupled with agreement probability $\sum_x \min(p_x, q_x) = 1 - d_{\mathrm{TV}}(p, q)$: allocate this common mass to equal pairs, then couple the remaining normalized residual masses. Start with the exact migrated pair. As long as $m'_t = \phi(m_t)$ and past readouts agree, the readouts and the common policy action agree. Couple the mapped old next state with the new next state using the preceding construction. Conditional on every such paired history, (6) makes their agreement probability at least $1 - \delta$. Conditional on a mapped old next state, sample an old state from its conditional distribution; zero-mass cases are irrelevant and may be defined arbitrarily. This gives the correct old transition marginal and preserves the pairing whenever the two mapped states agree. Induction shows that no pairing failure occurs through H steps with probability at least $(1 - \delta)^H$. On that event the complete transcripts agree. For any coupling, the probability of unequal transcripts bounds total variation, giving (7). On equal transcripts the losses agree, and elsewhere their difference is at most one. □

The same argument permits step-dependent uniform bounds δ_t and replaces the right side by $1 - \prod_{t=0}^{H-1}(1 - \delta_t)$. This is the finite controlled form of the established kernel-perturbation principle, not a stronger universal result than the cited literature.

The bound is sharp without mixing assumptions. Let old and new state spaces be $\{0, 1\}$, let ϕ and readout be identity, and start at zero. The old system remains at zero; the new system jumps from zero to an absorbing one with probability δ on every step. The probability of a transcript containing one is exactly $1 - (1 - \delta)^H$. The example also shows why a small one-step mismatch does not imply a small unlimited-horizon guarantee.

6 What checking a learned translator can certify

Suppose a scalar Bernoulli test is available under each of N explicitly enumerated reachable old states. The old reference can be restored and independently sampled n times at each state. A candidate translator is frozen before evaluation and returns a declared probability q_m there. Let p_m be the actual reference probability and $\widehat{p}_m$ its sample mean. For $0 < \alpha < 1$, put

$$\eta = \sqrt{\frac{\log(2N/\alpha)}{2n}}, \qquad \widehat{\epsilon} = \max_m |\widehat{p}_m - q_m|.$$

Proposition 6.1 (Finite uniform verification; standard concentration)**.** *With probability at least* $1 - \alpha$, $\max_m |p_m - q_m| \le \widehat{\epsilon} + \eta$.

Proof. For a Bernoulli variable X, the logarithm of $\mathbb{E}e^{s(X - \mathbb{E}X)}$ has value and derivative zero at $s = 0$ and second derivative at most $1/4$; the latter is a variance under an exponentially tilted Bernoulli law. Integration gives the bound $s^2/8$ for its logarithm. Independence, Markov's inequality, and minimization at $s = 4\eta$ give $\mathbb{P}(\widehat{p}_m - p_m \ge \eta) \le e^{-2n\eta^2}$, and applying the same argument to $-X$ gives the lower tail. A union bound over N states therefore gives probability at most $2Ne^{-2n\eta^2} = \alpha$ of any failure. On the complementary event the triangle inequality completes the proof. This is the bounded-variable argument of Hoeffding [8] specialized to Bernoulli observations. □

The proposition requires Nn reference executions, a trusted reset into each state, stationary independent samples, and a credible reference interface. If both reference and replacement are sampled, allocate failure probabilities $\alpha/2$ to each family, apply the bound to both sets of means, and add their two confidence radii. The resulting joint confidence is at least $1 - \alpha$. If only a small deployment-weighted sample of states is tested, the conclusion is about that distribution unless a separate coverage guarantee is proved.

Example 6.2 (An unvisited state defeats uniform certification)**.** Consider a verifier with black-box observation access only, no source-code certificate or alternative interface revealing unvisited outputs, and a state $m_\star$ that it never inspects. Consider two candidate mechanisms whose answers coincide on every inspected state, with reference probability zero everywhere. At $m_\star$, candidate one returns probability zero and candidate two returns one. Their verification transcripts have identical laws, while their maximum errors are zero and one. No decision based only on those transcripts can distinguish a globally correct translator from the defective one. A certificate is therefore a resource with provenance, not a free consequence of checking a few examples.

7 Executable centers and adversarial coarse migration

The preceding product-simplex centers need not be realizable by the intended replacement. Fix, for every possible coarse output y, a finite nonempty menu $\mathcal{C}_y$ of profiles produced by admitted decoder artifacts. Each artifact returns exact rational probability tables for the declared finite tests under a specified evaluator and execution budget. Thus a menu entry is executable under that evaluator, not an arbitrary real-valued table asserted to be a program. Coherence across tests is supplied by menu construction when required. The menus, semantics and evaluator are trusted inputs. Determining whether arbitrary code implements them is a different verification problem. The decoder can select an entry independently for each (y, z); no shared code-size or installation budget couples these choices.

Proposition 7.1 (Finite executable-menu frontier)**.** *Let $N_\epsilon(S; \mathcal{C})$ be the minimum number of entries of $\mathcal{C}$ covering S to radius ϵ, with value $+\infty$ if impossible. Under the fixed migration contract the optimum is*

$$L_\epsilon^{\mathrm{exec}} = \max_y N_\epsilon(\sigma(F_y); \mathcal{C}_y).$$

Proof. For each y a decoder supplies at most L admitted entries, which must cover its fiber. Conversely choose a minimum menu cover in each fiber and index it with the trusted record, reusing symbols across fibers. When the old profiles and ϵ are also explicitly rational, exhaustive enumeration decides feasibility using exact rational comparisons. If one fiber has no cover no alphabet can repair it. □

This is finite set cover, not a new executable-semantics theorem. For one fiber with Bernoulli profiles $0, 1$, unrestricted centers give one-symbol error $1/2$; the menu $\{0, 1\}$ gives one-symbol error 1 and two-symbol error zero. A menu $\{0\}$ cannot reach error below one at any record size. A shared installation budget destroys the independent-fiber reduction: two singleton fibers requiring respectively the only artifacts 0 and 1 each have local cover number one, but a budget allowing only one installed artifact makes exact preservation impossible. Our theorem excludes that budget.

7.1 A precisely delimited untrusted component

Replace the deterministic coarse map by a finite nonempty allowed-output set $A(m) \subseteq Y$ for each old state. An untrusted migrator may output any $y \in A(m)$, even after seeing the record. A trusted encoder sees the actual m and writes $z(m)$ before that choice; it does not observe the chosen y. The record and the decoder's execution remain protected. The requirement is

$$\max_{m \in R} \max_{y \in A(m)} d(\sigma(m), Q_{y,z(m)}) \le \epsilon, \qquad Q_{y,z} \in \mathcal{C}_y.$$

Let $R_y = \{m : y \in A(m)\}$. A nonempty set $S \subseteq R$ is compatible when, for each y with $S \cap R_y \neq \varnothing$, some $c \in \mathcal{C}_y$ satisfies $d(\sigma(m), c) \le \epsilon$ for every $m \in S \cap R_y$.

Theorem 7.2 (Robust record as compatible partition)**.** *The least trusted record alphabet equals the minimum number of compatible blocks partitioning R, or $+\infty$ if no such partition exists.*

Proof. The preimages $z^{-1}(j)$ of a feasible encoder form a partition after empty blocks are removed. For fixed j, y, its installed decoder center witnesses compatibility of the block. Conversely, assign one record symbol to each compatible block and install, for every y, j, a witnessing center. Choose any menu entry on empty intersections. The construction handles every adversarial allowed y, so it has the claimed worst-case guarantee. □

Compatibility is hereditary. Equivalently, forbid each inclusion-minimal incompatible subset from being monochromatic. Singleton forbidden sets denote infeasibility, rather than an ordinary loop-free hypergraph coloring problem. At $\epsilon = 0$, when every required exact profile is available in every relevant menu, incompatibility reduces to an edge between m, n with unequal profiles and $A(m) \cap A(n) \neq \varnothing$. The optimum is the chromatic number of that confusability graph. This is the established zero-error side-information and functional-coding mechanism [16, 10, 6]; B14 already uses a characteristic graph for a different encoder interface. For singleton $A(m) = \{\phi(m)\}$ we recover the per-fiber menu theorem.

Pairwise compatibility does not suffice at positive tolerance. Take three point-mass profiles on three outcomes, a common allowed output, and a menu containing all three pairwise midpoint distributions. At $\epsilon = 1/2$ every pair has a center but the triple has none, since all three coordinates would have to be at least $1/2$. One symbol is impossible; two suffice. Hypergraph coding under maximal distortion already exploits this higher-order geometry: Basu et al. [2, Section 4, Definition 4] define compatibility by a common enclosing ball for each side-information value, and Basu et al. [3] develop the broader source-coding theory. Their support condition corresponds to $y \in A(m)$. Our one-shot finite-menu partition is an elementary operational specialization; neither the hypergraph idea nor its advantage over pairwise tests is claimed as new.

Proposition 7.3 (No origin binding, no state-specific guarantee)**.** *Suppose an adversary can replace the entire final interface, including all record and certificate fields, by the same accepted value v for every m. With no other trusted state-dependent information, any decoder's worst-case profile error is at least*

$$\inf_{q} \max_{m \in R} d(\sigma(m), q).$$

Proof. On that common accepted value the decoder has one profile q_v, independent of the actual old state. Taking the maximum over states and then the infimum over profiles proves the claim, also for a randomized decoder through its induced outcome law. Refusal to accept v avoids the premise but supplies no accepted migration for that interface. □

For two opposite deterministic bits the lower bound is $1/2$. A proof that a declared table has a property cannot authenticate the old input or installed artifact. A deployment must bind the certificate to both, or state an explicit soundness assumption. Translation validation already makes trusted checking of untrusted producers central [12, 15]. The present partition theorem models a protected encoder and decoder; it supplies neither a cryptographic protocol nor correctness of an arbitrary migrator.

7.2 Exact relationship to functional epsilon entropy

The compatibility family is not a new object. To compare contracts explicitly, choose any law $p(m, y)$ with support exactly $y \in A(m)$ and set $f(m, y) = \sigma(m)$. The common-center test in Basu et al. [2, Section 4, Definition 4] has the same support-indexed common-center form. Literal

equality of feasible sets requires matching both the distance and legal reconstruction set: the 2019 presentation uses Euclidean distance, whereas here the metric is maximum test-wise total variation and centers must be probability profiles. Menus further restrict witnesses. Their hyperedges are *compatible* sets, whereas our forbidden hyperedges are minimal *incompatible* sets. The expanded paper's corresponding locators are Section II-B, Definitions 1–4, and Section III-B, Theorem 3 [3]. This corrects earlier review notes that shifted these section numbers by one.

Proposition 7.4 (Cover, partition and forbidden-set equivalence)**.** *Let $\mathcal{F}$ be the hereditary family of compatible nonempty subsets of R. Its minimum cover size equals its minimum partition size and the weak chromatic number of the minimal-incompatible-set hypergraph. A forbidden singleton makes all three problems infeasible. Hence the robust alphabet is*

$$\min\left\{\sum_{S\in\mathcal{F}} x_S : \sum_{S\ni m} x_S \geq 1 \ (m\in R), \quad x_S\in\{0,1\}\right\}.$$

Proof. A partition is a cover. From a cover assign each state to one covering set; the resulting nonempty blocks remain compatible by heredity. A color class is compatible exactly when it contains no minimal incompatible subset. Weak coloring forbids a monochromatic forbidden set, not every repeated color within that set. The displayed integer program is the definition of minimum cover. □

Legal reconstruction centers matter independently of the hypergraph notation. For the three point masses, the ambient zero vector has half-ℓ_1 distance $1/2$ from each, yet it is not a probability law; a legal probability center requires radius $2/3$. An unrestricted ambient enclosing-ball computation would therefore be unsound for this decoder contract. This elementary reduction explains the overlap rather than claiming priority. The source-coding objective differs: Basu et al. [3, Section II-A.1] use vanishing average symbol-exceedance probability as block length grows and an information-rate optimization. Our guarantee is zero exceedance for every allowed pair in one copy operation, minimizing an integer alphabet. No asymptotic entropy value is substituted for that alphabet.

For example, let M be uniform on the three point-mass profiles above, with constant side information and all pairs compatible but not the triple. The one-shot optimum is two symbols, or one fixed-length bit. In contrast, for a random compatible set W containing M, $H(M \mid W)\leq 1$, so $I(M;W)\geq \log_2 3-1$. Choosing each pair containing M with probability $1/2$ attains equality: W is uniform on the three pairs and $M \mid W$ is uniform on its pair. Thus the auxiliary information optimum is $\log_2(3/2)\simeq 0.58496$ bits. This is an independent finite calculation showing distinct objectives, not fractional storage in a one-shot protocol. The contribution remains a migration interpretation and explicit restrictions, not a new functional compression theorem.

7.3 Binding a certificate to an actual captured origin

A locally executable validation contract is possible, but must name its trusted boundary. Let a trusted operation atomically capture the actual old state m at a declared time, retain its reference profile, and allocate a protected nonreused epoch e. The policy, reference table, tolerance, parser and evaluator are verifier-owned. An untrusted producer submits a claimed source, epoch and artifact. The checker rejects source/epoch mismatches and malformed artifacts; it copies proposed bytes into a protected immutable buffer, evaluates that buffer, and accepts only when its profile is within ϵ of the captured reference. Successful installation consumes the epoch. Queries execute that same retained buffer through the same evaluator, without reopening a mutable path or accepting an unrelated reported output.

Proposition 7.5 (Soundness under the protected capture contract)**.** *If capture is faithful, the trusted policy and evaluator are correct, epochs cannot be reused or rolled back, and the checked immutable buffer is the one executed, every accepted query preserves the captured profile to tolerance* ϵ*.*

Proof. The faithful capture identifies the actual profile to compare. Acceptance checks the interpreted profile of the retained bytes against that reference. Execution uses those same bytes and semantics, so the checked inequality is the required one. Consuming the protected epoch rejects later installation requests for that capture. A valid allowed artifact is accepted if submitted with matching source and current epoch before consumption; arbitrary producer behavior need not supply one. □

This is a finite validator with explicit premises, not cryptographic actual-origin authentication. Capturing a self-reported identifier would not satisfy the premise. If the real source changes during capture, an atomic or coherent snapshot mechanism is necessary. Protected storage, reference tables and artifact bytes are additional resources; this protocol does not optimize their combined cost. Stored-code integrity and evidence that the intended code actually produced an output are distinct established obligations; VAPE treats the latter with a concrete hardware/software security model [7, Section 4]. RFC 9334 separates evidence appraisal and freshness and recognizes a race after evidence generation [5, Section 10]. Our single-process model tests the protocol logic, while assuming host isolation and correct measurement.

Theorem 7.6 (Accepted-risk accounting; elementary and sharp)**.** *Fix a probability law for source, implementation randomness and adversarial strategy. Let* A *be acceptance,* B *failure of the binding assumptions, and* $D \in [0,1]$ *the actual profile discrepancy on an accepted execution. Assume* $D \le \epsilon$ *on* $A \cap B^c$ *and* $\mathbb{P}(A \cap B) \le \rho$*, where* $0 \le \epsilon, \rho \le 1$*. If* $a = \mathbb{P}(A) > 0$*, then*

$$\mathbb{E}[D \mid A] \le \epsilon + (1-\epsilon)\min\{1, \rho/a\}.$$

Also $\mathbb{P}(A \cap \{D > \epsilon\}) \le \rho$*.*

Proof. Put $b = \mathbb{P}(A \cap B) \le \min(a, \rho)$. Splitting the accepted expectation across B gives $\mathbb{E}[\mathbf{1}_A D] \le \epsilon(a-b) + b$; divide by a. The exceedance event is contained in $A \cap B$. Equality is attainable by setting $b = \min(a, \rho)$, $D = 1$ on the bad accepted event, and $D = \epsilon$ on its good complement. No independence assumption is used. □

If $a = 0$, accepted conditional risk is undefined; the joint-risk statement remains meaningful. For $\epsilon = 0$ and $A = B$ of probability $0 < \rho < 1$, every accepted run may have discrepancy one, refuting the careless conditional bound $\epsilon + \rho$. A certified $a \ge a_{\min} > 0$ allows substitution of $a_{\min}$ in the denominator. This theorem neither establishes ρ from software tests nor certifies an acceptance floor. Uniform adversarial claims require these premises for every allowed strategy. Statistical-certificate failure must also be included in the bad event or bounded separately.

8 Verification along one trajectory without named-state resets

A weaker observation interface can replace resets only by adding explicit structure. Let R be partitioned into K nonempty cells B_c. At trial t a trustworthy observable label $C_t = c$ identifies the current cell, but need not reveal the actual state $M_t \in B_c$. Both are fixed before the outcome $X_t \in \{0,1\}$ is drawn. An analyst's pre-outcome filtration may contain the hidden state, while the verifier observes only C_t, X_t. Assume fixed reference probabilities p_m and

$$\mathbb{E}[X_t \mid \mathcal{F}_{t-1}] = p_{M_t}, \quad C_t, M_t \text{ are } \mathcal{F}_{t-1}\text{-measurable}, \qquad \sup_{m,m' \in B_c} |p_m - p_{m'}| \le \beta_c.$$

The bounds β_c are independently certified structural inputs, not estimates inferred from seeing only one member of a cell. The outcome X_t is $\mathcal{F}_t$-measurable. The trajectory and revisits may be adaptive. There is no reset, independence of visits, recurrence, mixing-time bound, or guaranteed finite completion. The frozen replacement returns q_c for the scalar query in every state of cell c.

Let $\widehat{p}_{c,n}$ be the mean of the first n outcomes observed in cell c, defined only if those visits occur. For $0 < \alpha < 1$ put

$$r_{c,n} = \sqrt{\frac{\log(2Kn(n+1)/\alpha)}{2n}}.$$

Theorem 8.1 (Visit-index certificate with certified cell oscillation)**.** *With probability at least $1-\alpha$, simultaneously for all cells and every completed positive visit count n,*

$$\max_{m \in B_c} |p_m - q_c| \le \min\{1, |\widehat{p}_{c,n} - q_c| + r_{c,n} + \beta_c\}.$$

Taking the maximum of these bounds over all cells, assigning bound one to an unvisited cell, gives a time-uniform certificate over R. It remains valid at an adaptively chosen stopping time.

Proof. For fixed c define $I_t = \mathbf{1}\{C_t = c\}$, $N_t = \sum_{s \le t} I_s$ and $S_t = \sum_{s \le t} I_s(X_s - p_{M_s})$. Conditional Hoeffding's inequality makes $Z_t = \exp(\lambda S_t - \lambda^2 N_t/8)$ a nonnegative supermartingale starting at one. Stop at the nth visit, possibly infinite. Bounded stopping followed by Fatou's lemma gives $\mathbb{E}[Z_{\tau_n}\mathbf{1}\{\tau_n < \infty\}] \le 1$. Thus for each $a > 0$, optimizing $\lambda = 4a$ bounds the upper tail event $\{\tau_n < \infty, S_{\tau_n}/n \ge a\}$ by e^{-2na^2}; use $-\lambda$ for the lower tail. With $a = r_{c,n}$ the two-tail failure probability is at most $\alpha/[Kn(n+1)]$. Sum over c and $n \ge 1$, using $\sum_n 1/[n(n+1)] = 1$. On the resulting event the empirical mean differs from the average of visited probabilities by at most $r_{c,n}$. Every p_m in B_c is within β_c of that average, and the triangle inequality proves the assertion. No conditioning on future visit counts is used. □

This is elementary error spending and martingale concentration, not a new confidence-sequence principle. Sharper time-uniform bounds are established by Howard et al. [9]; our $\sqrt{\log n/n}$ radius is deliberately conservative. Singleton cells recover observed named states without resets. Non-singleton cells permit hidden identities at the price of a certified β_c term. Multiple queries require pre-outcome query selection and a separate allocation of the failure budget; arbitrary new interventions are not covered here.

Example 8.2 (Why the cell certificate cannot be learned for free)**.** A label identifies one cell containing m_0, m_1, but the trajectory remains forever at m_0. Two reference systems have $p_{m_0} = 0$ and respectively $p_{m_1} = 0$ or 1. With $q_c = 0$ their audit transcripts are identical, while uniform error is zero or one. Without a certified oscillation bound or access to m_1, no amount of repetition distinguishes them. Likewise no finite universal completion time follows without a quantitative visitation assumption.

The value of the theorem is its explicit trade: reset access is replaced by predictable labels, conditional stationarity and certified within-cell control. It does not certify unknown latent-state models from raw passive observations.

8.1 A quantitative visitation premise and a deadline

Fix integers $b, J, n \ge 1$. Let G_j be the full analysis history at the end of block j, containing observations $(j-1)b+1, \dots, jb$. Write $H_{c,j}$ for the indicator that block j visits cell c. Assume, for every cell and block,

$$H_{c,j} \text{ is } G_j\text{-measurable}, \qquad \mathbb{P}(H_{c,j} = 1 \mid G_{j-1}) \ge \kappa_c > 0 \quad \text{almost surely.}$$

This conditional reachability bound must come from a specified process/policy or a separate valid certificate. Stationary mass and observed visit fractions are insufficient. Blocks need not be independent. With the earlier predictable labels and $G_j = \mathcal{F}_{jb}$, the first label of a block is already known; for $K > 1$, a positive all-cell bound cannot hold with $b = 1$. Longer blocks can include fresh transitions and make the premise meaningful.

Theorem 8.3 (Availability deadline under conditional block hits)**.** *Let $N_c(T)$ count visits to cell c by $T = Jb$ observations. Then*

$$\mathbb{P}\{\exists c : N_c(Jb) < n\} \le \Gamma_{J,n} := \min\left\{1, \sum_{c=1}^{K} \mathbb{P}\{\mathrm{Bin}(J, \kappa_c) < n\}\right\}.$$

Proof. For one cell put $S_j = \sum_{i \le j} H_{c,i}$ and $f_r(s) = \mathbb{P}\{s + \mathrm{Bin}(r, \kappa_c) < n\}$. This is nonincreasing in s and satisfies $f_{r+1}(s) = (1-\kappa_c)f_r(s) + \kappa_c f_r(s+1)$. Given G_{j-1}, the hit probability $p_j \ge \kappa_c$ implies

$$\mathbb{E}[f_{J-j}(S_j) \mid G_{j-1}] \le (1-\kappa_c)f_{J-j}(S_{j-1}) + \kappa_c f_{J-j}(S_{j-1}+1) = f_{J-j+1}(S_{j-1}).$$

Iterating yields $\mathbb{P}(S_J < n) \le f_J(0)$. Since each hit block supplies at least one visit, $N_c(Jb) \ge S_J$. A union bound over cells proves the claim. This does not assert independence of hits or of cells. □

The domination argument is standard and also underlies portfolio D09's progress bound; the present use connects block reachability to audit completion. General hitting-time theory distinguishes stationary mass from quantitative access [11, Chapters 10–11].

At the deadline retain each cell's *first* n observations and its Theorem 8.1 upper bound $U_{c,n}$. If $\Gamma_{J,n} \le \gamma$, then all first-n certificates are available and valid with probability at least $1 - \alpha - \gamma$. Validity alone still has probability at least $1 - \alpha$: γ measures completion failure, not a second source of false certification. Accept only when all K cells have their required observations and every cell bound is at most a declared tolerance ϵ; otherwise abstain. Time-uniform validity bounds the probability of ever accepting a truly over-tolerance fixed candidate by α, regardless of coverage. Availability does not force acceptance.

A sufficient acceptance margin is explicit. Put $e_c = \max_{m \in B_c} |p_m - q_c|$. On the confidence event the first-n mean is within $e_c + r_{c,n}$ of q_c, hence

$$U_{c,n} \le \min\{1, e_c + 2r_{c,n} + \beta_c\}.$$

If the right-hand untruncated expression is at most ϵ for every cell, the first-n policy accepts by Jb with probability at least $1 - \alpha - \gamma$. This is a sufficient slack condition, not a guarantee for all candidates whose true error merely equals the tolerance.

Example 8.4 (Stationary probability does not give a uniform deadline)**.** The two-cell chain with switching probability $s \in (0,1)$ has stationary probabilities $(1/2, 1/2)$. From cell zero, the probability of never visiting cell one in the first T observed states is $(1-s)^{T-1}$; from stationarity it is $\frac{1}{2}(1-s)^{T-1}$. As $s \downarrow 0$, neither expression supplies an arbitrarily high-confidence uniform deadline despite unchanged stationary mass. A rate floor is a substantive assumption.

8.2 A fully specified hidden-state test case

Let $M_t = (C_t, H_t) \in \{0,1\}^2$. Only C_t is observed. After each outcome, the cell flips independently with probability κ, and the hidden bit flips independently with probability $1/2$. The four outcome probabilities are $(1/5, 3/10, 7/10, 4/5)$ in lexicographic state order; the candidate reports $q_0 =$

$1/4, q_1 = 3/4$. Thus $\beta_c = 1/10$ and $e_c = 1/20$ are exact structural facts of this model. Two-observation blocks give a conditional cell-hit floor κ: the first state either supplies the hit, or the intervening flip does. No reference state is reset or exposed to the verifier.

For $K = 2$, $\alpha = 0.025$ and $n = 5000$, $r_{c,n} \simeq 0.04702$, so $e_c + 2r_{c,n} + \beta_c \simeq 0.24404 < \epsilon = 0.25$. Choose $J = \lceil 2n/\kappa \rceil$ and $b = 2$. For a binomial variable of mean $\mu = J\kappa \geq 2n$, Markov's inequality applied to $e^{-(\log 2)X}$ gives

$$\mathbb{P}(X < n) \leq \exp(\mu(\log 2)/2 - \mu/2) \leq e^{-\mu/8}.$$

Here $\mathbb{E}e^{-(\log 2)X} = (1 - \kappa/2)^J \leq e^{-\mu/2}$ and $(1 - \log 2)/2 > 1/8$. Consequently $\Gamma_{J,n} \leq 2e^{-J\kappa/8}$ and the sufficient deadline is $2J$. At $\kappa = 1/4$ it is 80,000 observations; at $\kappa = 1/50$ it is 1,000,000. These are conservative guarantees, not observed stopping times. Report the failure bound through its logarithm, at most $\log 2 - 1250$, rather than rounding it to mathematical zero.

The supplied seeded trajectory audit uses the latest-count bounds of Theorem 8.1 and retains all trials, including timeouts. The conservative deadline above instead concerns a retained-first-n policy; these are different valid policies, and their times are compared descriptively, not as identical procedures. A negative control assigns an unjustified $\beta = 0$ to a cell with an unvisited state of probability one and an observed state of probability zero. Repeated zeros would then produce a small reported bound although actual uniform error is one. Visits do not certify the structural diameter assumption.

9 Reproducible evidence

The script `code/verify_and_plot.py` is deterministic and uses exact rational arithmetic for the mathematical checks. It compares the scalar dynamic program with enumeration of every set partition of the six-state and three-state fibers for all six record capacities: twelve independent comparisons. It also enumerates all 256 binary two-test tables on four old states with a fixed two-fiber migration and checks feasibility for one and two record symbols, giving 512 checks. Finally it explicitly enumerates binary paths for horizons zero through eight in the sharp dynamic witness, and verifies normalization and the product formula.

The new script `code/verify_extensions.py` independently compares compatible partitions with direct decoder-table enumeration on 5,832 three-state binary-profile instances, allowing all nonempty two-output adversarial sets and nonempty binary center menus at radii $0, 1/2, 1$. It executes a finite rational-table decoder language for the three-point counterexample and independently recovers the original scalar frontier by greedy interval covering. A 64-step fair-coin confidence diagnostic uses exact path masses but floating-point logarithmic boundaries; it is not a proof of the general adaptive theorem.

The original CSV files retain exact fractions; floating-point arithmetic is used for figures and the explicitly labeled confidence-boundary diagnostic. Those original scripts are deterministic and require no seed. The new validation below includes explicitly seeded random trajectories. These finite calculations check implementations and illustrate proofs; they neither prove general covering complexity nor establish historical priority.

9.1 Protocol fault injection and no-reset trajectory validation

The additional `code/validate_binding_visitation.py` tests a single-process protected-runtime model. Its allowed producer inputs are source claim, epoch and artifact bytes; the tolerance belongs to trusted setup. The rational-PMF language admits at most 256 bytes, exactly the opcode and

probability fields, and unsigned integer or fraction encodings with at most nine digits in each integer and a positive denominator. Duplicate keys, scientific notation, decimal aliases, policy overrides and unsupported opcodes are rejected. This syntactic bound makes interpreter arithmetic finite and bounded; it is not a measured runtime or host-isolation guarantee.

All 672 source/claim/epoch/grid-probability combinations match the declared predicate: 12 accepted, 660 rejected. Twelve consumed-epoch replay attempts are rejected. Four post-check proposal mutations leave the installed immutable buffer unchanged; a deliberately vulnerable re-read comparator changes its answer in all four. Twelve malformed/unsupported/oversized artifacts are rejected. A checker that trusts only a producer's claimed source accepts a zero-error claimed profile with actual error 3/5, a retained negative control. The accepted-risk formula is attained on 4,830 rational grid cases; this tests the extremal arithmetic, while the proof handles general joint laws.

For visitation, 27 exact two-cell occupancy distributions satisfy the block-tail bound. A separately written reviewer script checks 640 exact inequalities over 128 adapted Bernoulli policy trees. Neither calculation is a substitute for the all-parameter proof. Three sets of 100 seeded four-state trajectories start at $(0, 0)$ and run until the latest-count certificate accepts or 20,000 observations are reached. Table 1 retains the slow-regime failures. The candidate's true uniform error is 0.05 in every regime; this positive-model study measures completion, not empirical false-acceptance calibration.

κ	Completed/100	Median if complete	95% Wilson interval	First-n deadline
0.25	100	993	$[0.963, 1.000]$	80,000
0.02	100	1,148.5	$[0.963, 1.000]$	1,000,000
0.0001	80	6,785	$[0.711, 0.867]$	200,000,000

Table 1: Synthetic no-reset audits; time unit is one observation, cap is 20,000. Intervals describe completion fractions across independently seeded runs, not proof confidence or physical deployment reliability. Medians exclude timeouts. The sufficient deadline concerns a different, conservative first-n policy.

Seeds and every trial are recorded. Floating-point arithmetic is used for random thresholds, confidence radii and descriptive intervals; rational arithmetic is used for the finite protocol, occupancy and risk checks. The observed fast completion cannot remove the quantified visitation assumption. For switching probability 10^{-6}, the exact no-visit probability from cell zero in 1,000 observations is $(999999/1000000)^{999} \simeq 0.9990015$, despite stationary mass 1/2. Likewise the retained false-β control shows that apparent precision can coexist with actual uniform error one. These are synthetic mathematical and software-model validations, not measured secure migrations or hardware attacks.

10 Interpretation, portfolio distinctness, and next research

Being. At a given time a memory supports a declared set of operational tests. The profile records that capability, without making a metaphysical claim about what the system is. A capability can persist when the implementation changes if the corresponding replacement contract holds.

Becoming. A migration changes internal organization. It can discard irrelevant distinctions while preserving a chosen task family, or destroy a distinction needed by a later task. The covering frontier makes one price of preserving such distinctions explicit. Task families that expand after deletion need a larger archive or an admitted failure mode.

The MIND. A system can use migration certificates to decide which old predictive or historical claims remain warranted. This is an operational form of tracking the limits of its own information. Neither profile preservation nor a trusted side record establishes subjective experience, a unique self-boundary, or personal identity.

The central distinction from D04 is that this note starts with an old capability and asks what survives replacement; D04 chooses which new distinctions to learn. D07 optimizes average predictive risk under a retained bit budget and a noisy acquisition channel. Its source-coding bound should not be repeated here as a second novelty claim. The present uniform, fixed-migration, multi-test contract and the adaptive transition theorem are different problems, so an immediate mathematical merge is unnecessary. Editorially, D06 belongs naturally as a continuation of the existing recoverable-memory manuscript, with the scalar example possibly an appendix.

Unresolved proposals, not theorems. The executable-menu and adversarial-output results close only finite declared contracts. Open problems include deriving menus from general programs, shared code and runtime budgets, authenticating a real old state and installed decoder, and obtaining cell oscillation or visitation guarantees for an actual measured system. The four-state synthetic model supplies such premises analytically but does not validate them in deployment. D07 owns prior-weighted compression; D08 owns paid solver verification and its learning value. This note does not duplicate those claims. No theorem here turns small training error into arbitrary intervention preservation.

Readiness. This is a complete technical-note draft with proved elementary results, finite reproducible examples, and an internal review record. Its significance lies in contract clarification and integration. The current primary-source search identifies substantial overlap with zero-error coding, translation validation and confidence sequences. It does not certify novelty, and no theorem here warrants the claim that a new general continuity theory has been discovered. Version 3 distinguishes conditional guarantees, protocol-model tests and empirical deployment evidence; only the first two are supplied.

References

[1] Alessandro Abate, Frank Redig, and Ilya Tkachev. On the effect of perturbation of conditional probabilities in total variation. *Statistics & Probability Letters*, 88:1–8, 2014. doi: 10.1016/j.spl.2014.01.009. URL `https://www.cs.ox.ac.uk/people/alessandro.abate/publications/ART14.pdf`.

[2] Sourya Basu, Daewon Seo, and Lav R. Varshney. Functional epsilon entropy, 2019. URL `https://arxiv.org/abs/1910.08276`. Version 1, 18 October 2019; introduction and Section 4 inspected.

[3] Sourya Basu, Daewon Seo, and Lav R. Varshney. Hypergraph-based source codes for function computation under maximal distortion, 2022. URL `https://arxiv.org/abs/2204.02586`. Version 2, 28 December 2022; Sections II-A.1, II-B and III-B inspected in v3; prior reviewer section locators corrected.

[4] Gaoang Bian and Alessandro Abate. On the relationship between bisimulation and trace equivalence in an approximate probabilistic context, 2017. URL `https://www.cs.ox.ac.uk/people/alessandro.abate/publications/bcBA17.pdf`. Author-hosted 17-page conference manuscript; Theorems 4–6. Extended version arXiv:1701.04547.

[5] Henk Birkholz, Dave Thaler, Michael Richardson, Ned Smith, and Wei Pan. Remote attestation procedures (rats) architecture. RFC 9334, 2023. URL https://www.rfc-editor.org/rfc/rfc9334.html. Informational RFC; architecture and freshness sections inspected.

[6] Nicolas Charpenay, Mael Le Treust, and Aline Roumy. Side information design in zero-error coding for computing. *Entropy*, 26(4):338, 2024. doi: 10.3390/e26040338.

[7] Ivan de Oliveira Nunes, Karim Eldefrawy, Norrathep Rattanavipanon, and Gene Tsudik. A verified architecture for proofs of execution on remote devices under full software compromise, 2020. URL https://arxiv.org/abs/1908.02444. Version 2; Section 4 security definitions inspected; no complete hardware-proof audit.

[8] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830. URL https://www.cs.rpi.edu/academics/courses/spring06/random/hoefding.pdf.

[9] Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. doi: 10.1214/20-AOS1991. URL https://arxiv.org/abs/1810.08240.

[10] Prashant Koulgi, Ertem Tuncel, Shankar L. Regunathan, and Kenneth Rose. On zero-error source coding with decoder side information. *IEEE Transactions on Information Theory*, 49(1):99–111, 2003. doi: 10.1109/TIT.2002.806154. URL https://web.ece.ucsb.edu/publications/rose/pubs/pub40-T-IT1-03.pdf.

[11] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, second edition, 2017. URL https://pages.uoregon.edu/dlevin/MARKOV/markovmixing.pdf. With contributions by Elizabeth L. Wilmer; Chapters 10–11 on hitting and cover times.

[12] George C. Necula. Translation validation for an optimizing compiler. In *Proceedings of PLDI*, pages 83–94, 2000. doi: 10.1145/349299.349314.

[13] Paul K. Rubenstein, Sebastian Weichwald, Stephan Bongers, Joris M. Mooij, Dominik Janzing, Moritz Grosse-Wentrup, and Bernhard Schölkopf. Causal consistency of structural equation models. In *Proceedings of the 33rd Conference on Uncertainty in Artificial Intelligence*, volume R15 of *Proceedings of Machine Learning Research*, pages 151–160. PMLR, 2017. URL https://proceedings.mlr.press/r15/rubenstein-17a.html. Author-hosted 15-page reading copy with appendix inspected; PMLR metadata reissued 4 October 2026.

[14] Brian Theory. Recoverable memory under finite-state endomorphic replacement: Universal permutation quotients and trusted replacement records, 2026. URL https://www.b-theory.com/manuscripts-html/recoverable-memory/. Reviewed working manuscript, 8 November 2026; site reading copy 1. Internal AI review, not external peer review. Accessed 13 November 2026.

[15] Jean-Baptiste Tristan and Xavier Leroy. Formal verification of translation validators: A case study on instruction scheduling optimizations. In *Proceedings of POPL*, pages 17–27, 2008. doi: 10.1145/1328438.1328444.

[16] Hans S. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, 1976. doi: 10.1109/TIT.1976.1055607. Publisher abstract inspected; graph mechanism also checked in Koulgi et al.