

# Six Birds: No-Go Theorems for Audited Emergence

Ioannis Tsiokos 

March 23, 2026

## Abstract

Six Birds theory studies finite systems observed through deterministic coarse-grainings (“lenses”) using a small audited vocabulary of primitives. This paper proves eight no-go theorems that constrain when common emergence narratives can be made honest in a finite, checkable setting. For a finite Markov system and horizon  $T$ , define an *arrow audit* as the Kullback–Leibler divergence between the observed path law and its time reversal. We prove a deterministic data-processing inequality: observation cannot increase arrow. As a consequence, any reversible system started in stationarity has zero observed arrow, even when the observation supports nontrivial protocol structure. Further results give graph-theoretic obstructions for force-like antisymmetric drives (exactness on forests and zero cycle sums), a finite variational identity for best macro-kernels and closure deficit, and a bounded-interface saturation theorem that prevents unbounded laddering. Together these theorems delimit what can and cannot be certified as emergent from audited primitives in finite models.

**Keywords:** coarse-graining; finite Markov chains; data processing; exact graph forms; closure deficit; bounded interfaces.

**Contribution summary.** This paper contributes an audited finite semantics for eight no-go fronts, a unified theorem framework linking arrow, graph exactness, closure, objecthood, and ladder obstructions, and a technical appendix structure that isolates routine derivations from the main proof flow.

## 1 Introduction

Many “emergent” claims in finite-state modeling take the form: after coarse observation, one can attribute a directed arrow of time, a force-like drive, a stable macro-law, or an agent-like persistence to a system that was specified only microscopically. This paper isolates a family of finite *no-go* statements showing that several such attributions are impossible unless specific structural primitives are present. The point is not metaphysical; it is bookkeeping. We work entirely with finite objects (finite sets, kernels, deterministic maps) and with audits that can be evaluated exactly.

**Six primitives (grammar, not section structure).** We use the following six primitives as a vocabulary for what may be added to a finite audited package. They are recorded per witness and used only to state contrasts; they are not treated as mutually exclusive or exhaustive.

1. *Rewrite*: state updates that change which microstate is represented (e.g. deterministic rewrites or state relabelings that alter future conditionals).

2. *Gating*: deletion or suppression of transitions (support-level constraints).
3. *Holonomy / protocol structure*: hidden cyclic or protocol coordinates whose evolution can be invisible to the observation.
4. *Staging / refinement*: state-splitting or refinement that introduces additional internal degrees of freedom.
5. *Packaging*: deterministic endomaps or simplex operators that “compress” state or distributions, typically studied via idempotence and contraction.
6. *Drive*: explicit breaking of detailed balance / nonreversibility, producing nonzero affinities.

**Honest vs. proxy macro-quantities.** Coarse quantities are defined by deterministic pushforward of the micro path law along an observation map. In particular, for a lens  $f : X \rightarrow Y$  we define the observed process  $Y_t = f(X_t)$  and the observed path law as the pushforward of the micro path law. When a first-order macro Markov chain on  $Y$  is introduced, it is used as an auxiliary approximation rather than as the theorem object.

**Main results (eight fronts).** All results are stated for finite systems and deterministic lenses. The theorem statements are collected in Section 3, and the later theorem-front sections refer back to those labels. In one line each, the fronts are:

- *Arrow/DPI*: deterministic observation cannot increase forward–reverse path-space KL; equality holds for identity observation. (Theorem 3.1)
- *Protocol trap*: holonomy can produce apparent protocol structure without honest entropy production, under a stationary initialization scope. (Theorem 3.3)
- *Force implies cycles*: nontrivial force-like structure requires non-forest support (positive cycle rank). (Theorem 3.5)
- *Null force under exactness*: if the log-ratio 1-form is exact (potential), cycle affinities vanish. (Theorem 3.7)
- *Closure deficit*: coarse closure deficit is a conditional-information obstruction characterized by a finite variational best macro kernel. (Theorem 3.9)
- *Objecthood (bounded)*: strict contraction implies at most one fixed distribution; the theorem front addresses uniqueness and separation in the contractive total-variation regime. (Theorem 3.11)
- *Idempotence no-ladder*: idempotent packaging stabilizes after one step (no iterate ladder). (Theorem 3.13)
- *Bounded-interface no-ladder*: with a fixed finite interface, definability is bounded by  $2^{|\text{im}(f)|}$  and no unbounded predicate ladder can arise without leaving the interface scope. (Theorem 3.15)

**Relation to prior Six Birds papers.** The present paper is a consolidation into a single no-go package with a uniform finite semantics and an auditable witness suite. The historical Six Birds line developed semantics for packaging, coarse-graining, directionality, protocol, closure, objecthood, and observable taxonomy; here we restrict to the finite setting and package the obstruction layer. The cited papers are used only as background and historical context; see [19–24].

**Paper organization.** Section 2 fixes notation and defines the audits. Section 3 collects the formal theorem statements in one place. Sections 4–8 develop the eight fronts in order: arrow/protocol, graph/affinity, closure, objecthood, and bounded-interface laddering. The appendices record routine finite calculations and derivations that support the main proofs. Supplementary material records provenance, computational audit artifacts, and formalization status; it is not required for the statements and proofs in the main text.

## 2 Formal setup and audit definitions

### 2.1 Finite probability notation

Let  $X$  be a finite set. Write  $\Delta(X)$  for the simplex of probability distributions on  $X$ . For  $\mu \in \Delta(X)$  and  $A \subseteq X$ , write  $\mu(A) = \sum_{x \in A} \mu(x)$ . All random variables in the paper are finite-valued unless stated otherwise.

### 2.2 Markov kernels and stationarity

A *Markov kernel* on  $X$  is a function  $K : X \times X \rightarrow [0, 1]$  such that for each  $x \in X$ ,

$$\sum_{x' \in X} K(x, x') = 1.$$

Given  $\mu \in \Delta(X)$  we write  $(\mu K)(x') = \sum_{x \in X} \mu(x) K(x, x')$ . A distribution  $\pi \in \Delta(X)$  is *stationary* if  $\pi K = \pi$ . (For finite  $X$  such a  $\pi$  always exists.) These are standard finite Markov-chain notions; see [12, 15].

### 2.3 Finite-horizon path laws

Fix an initial law  $\mu \in \Delta(X)$  and a kernel  $K$  on  $X$ . For an integer horizon  $T \geq 0$ , the *path space* is  $X^{T+1}$ . The forward path law  $P_T^{\mu, K} \in \Delta(X^{T+1})$  is

$$P_T^{\mu, K}(x_0, \dots, x_T) = \mu(x_0) \prod_{t=0}^{T-1} K(x_t, x_{t+1}).$$

Write  $\text{rev}(x_0, \dots, x_T) = (x_T, \dots, x_0)$  for path reversal, and define the reversed law  $P_T^{\mu, K, \text{rev}}$  on  $X^{T+1}$  by

$$P_T^{\mu, K, \text{rev}}(\gamma) = P_T^{\mu, K}(\text{rev}(\gamma)).$$

This definition treats reversal as an involution on the same finite space. In stationary reversible settings, this is the standard time-reversal construction for finite chains; see [9]. When  $P$  denotes the micro kernel, we also write the finite-horizon law as  $\mathbf{P}_{\mu, P}^{(T)}$ .

### 2.4 Deterministic lenses and pushforward

A *deterministic lens* (coarse map) is a function  $f : X \rightarrow Y$  between finite sets.<sup>1</sup> We also write  $\ell : \mathbf{X} \rightarrow \mathbf{Y}$  when denoting such a deterministic lens in theorem statements. The pushforward of  $\mu \in \Delta(X)$  is the distribution  $f\#\mu \in \Delta(Y)$  defined by

$$(f\#\mu)(y) = \sum_{x: f(x)=y} \mu(x).$$

---

<sup>1</sup>The term “lens” originates in the view-update and bidirectional transformation literature; here we use only the forward deterministic observation map [8].

For path space, define  $f^{(T)} : X^{T+1} \rightarrow Y^{T+1}$  by  $f^{(T)}(x_0, \dots, x_T) = (f(x_0), \dots, f(x_T))$ . The *honest observed path law* is

$$P_T^{\mu, K, f} = (f^{(T)})_{\#} P_T^{\mu, K} \in \Delta(Y^{T+1}).$$

We similarly define the reversed observed law  $P_T^{\mu, K, f, \text{rev}}(\eta) = P_T^{\mu, K, f}(\text{rev}(\eta))$ . Coarse observation can differ from the corresponding micro path-space asymmetry under aggregation [7].

## 2.5 Finite KL divergence and arrow audits

For distributions  $P, Q \in \Delta(\Omega)$  on a finite set  $\Omega$ , define the Kullback–Leibler divergence

$$D_{\text{KL}}(P\|Q) = \sum_{\omega \in \Omega} P(\omega) \log \frac{P(\omega)}{Q(\omega)},$$

with the conventions  $0 \log(0/q) = 0$  and  $p \log(p/0) = +\infty$  for  $p > 0$ . In the theorem statements we either assume mutual absolute continuity on the relevant supports or allow the value  $+\infty$ . For finite relative entropy and its standard identities, see [3, 11].

**Definition 2.1** (Micro and observed arrow at horizon  $T$ ). The *micro arrow* at horizon  $T$  is

$$\mathcal{A}_T(\mu, K) = D_{\text{KL}}(P_T^{\mu, K} \| P_T^{\mu, K, \text{rev}}).$$

For a lens  $f : X \rightarrow Y$ , the *observed arrow* is

$$\mathcal{A}_T(\mu, K; f) = D_{\text{KL}}(P_T^{\mu, K, f} \| P_T^{\mu, K, f, \text{rev}}).$$

These forward-versus-reversed path-space KL quantities are the finite relative-entropy objects underlying standard entropy-production and irreversibility diagnostics [14, 17].

**Deterministic DPI (informal statement).** Because  $f^{(T)}$  is deterministic, data processing implies

$$\mathcal{A}_T(\mu, K; f) \leq \mathcal{A}_T(\mu, K),$$

with equality for the identity lens. This is the deterministic specialization of the standard data-processing inequality for relative entropy; see [3].

## 2.6 Support graph, cycle rank, and affinity

The *directed support graph* of  $K$  has vertex set  $X$  and an edge  $x \rightarrow x'$  whenever  $K(x, x') > 0$ . Let  $G$  denote the underlying undirected support graph on  $X$  obtained by forgetting orientation and retaining an undirected edge  $\{x, x'\}$  when either direction has positive probability. For graph-theoretic statements we also write a generic finite graph as  $G = (V, E)$ , and a *forest* means that every connected component is a tree.

The *cycle rank* (first Betti number) of  $G$  is

$$\beta_1(G) = |E(G)| - |V(G)| + c(G),$$

where  $c(G)$  is the number of connected components. Forests satisfy  $\beta_1(G) = 0$ .

On a bidirected support edge (both  $K(x, x') > 0$  and  $K(x', x) > 0$ ), define the log-ratio 1-form

$$a(x, x') = \log \frac{K(x, x')}{K(x', x)}.$$

In graph notation we may equivalently write the same antisymmetric edge label as  $a(u, v)$  on oriented edges  $(u, v)$ . For a closed walk  $\gamma = (x_0 \rightarrow x_1 \rightarrow \dots \rightarrow x_m = x_0)$  on bidirected support, its *cycle sum* is  $\sum_{i=0}^{m-1} a(x_i, x_{i+1})$ . In graph notation we may also write a closed walk as  $v_0, \dots, v_k = v_0$  and the corresponding cycle sum as  $\sum_{i=0}^{k-1} a(v_i, v_{i+1})$ . We say  $a$  is *exact* if there exists a potential  $\phi : X \rightarrow \mathbb{R}$  such that  $a(x, x') = \phi(x') - \phi(x)$  on every bidirected support edge. Exactness forces every closed-walk sum to vanish. The cycle-rank identity is standard graph theory [1, 5], while the exactness and closed-walk viewpoint is the discrete 1-form perspective used in Hodge-theoretic treatments [13].

The *max cycle affinity* is the maximum absolute value of the cycle sum over a fixed cycle basis (equivalently, it vanishes iff all cycle sums vanish). In Markov-network thermodynamics, such cycle sums are the basic affinity objects [16]. The graph/affinity fronts use only these finite notions.

## 2.7 Closure deficit and best macro kernel

Fix a deterministic packaging map  $\Pi : X \rightarrow Y$  and consider the stationary initialization  $X_0 \sim \pi$  where  $\pi$  is stationary for  $K$ . Let  $Y_t = \Pi(X_t)$ . For an integer lag  $\tau \geq 0$ , define the *closure deficit*

$$\text{CD}_\tau(K; \Pi) = I(X_0; Y_\tau | Y_0),$$

the conditional mutual information under the stationary law. Equivalently,

$$\text{CD}_\tau(K; \Pi) = \mathbb{E} \left[ \log \frac{\mathbb{P}(Y_\tau | X_0, Y_0)}{\mathbb{P}(Y_\tau | Y_0)} \right].$$

A *macro kernel* on  $Y$  is any Markov kernel  $L : Y \times Y \rightarrow [0, 1]$ . The *best macro kernel at lag  $\tau$*  is the rowwise conditional law

$$K_\tau^*(y, y') = \mathbb{P}(Y_\tau = y' | Y_0 = y),$$

well-defined under stationarity. Exact preservation of a Markov macro-process under aggregation is the classical lumpability question [2, 10], and the variational minimizer below is a finite KL-projection viewpoint [4]. The closure front packages  $\text{CD}_\tau(K; \Pi)$  as the variational gap between the true packaged future laws and the best macro kernel; the exact finite identity used in the theorem statement is recorded in Section 6.

## 2.8 Packaging operators, idempotence, and bounded interfaces

A *deterministic packaging map* is an endomap  $e : X \rightarrow X$ . It is *idempotent* if  $e \circ e = e$ . Idempotence implies one-step stabilization of iterates and is used by the idempotence no-ladder front.

More generally, a *stochastic packaging operator* can be taken as a kernel  $P$  on  $X$  acting on  $\Delta(X)$  by  $\mu \mapsto \mu P$ . For objecthood statements on  $\Delta(Y)$  we write the induced affine operator as  $T_P(\nu) := \nu P$ . When contraction is discussed, we use total variation distance

$$\text{TV}(\mu, \nu) = \frac{1}{2} \sum_{x \in X} |\mu(x) - \nu(x)|.$$

The Dobrushin contraction coefficient of a kernel  $P$  is

$$\lambda(P) = \sup_{\mu \neq \nu} \frac{\text{TV}(\mu P, \nu P)}{\text{TV}(\mu, \nu)}.$$

This coefficient is standard in finite-state contraction theory; see [6, 18]. Strict contraction ( $\lambda(P) < 1$ ) implies uniqueness of a fixed distribution for  $P$ .

## 2.9 Definable predicates and the $2^{|\text{im}(g)|}$ bound

A *predicate* on  $X$  is a function  $p : X \rightarrow \{0, 1\}$ . Given a lens  $f : X \rightarrow Y$  (an “interface”), a predicate  $p$  is *f-definable* if there exists  $g : Y \rightarrow \{0, 1\}$  with  $p = g \circ f$ . Equivalently,  $p$  is the indicator of a union of fibers of  $f$ , and we write

$$\text{Def}(f) := \{g \circ f : X \rightarrow \{0, 1\} \mid g : Y \rightarrow \{0, 1\}\}.$$

Hence the number of  $f$ -definable predicates is exactly

$$|\text{Def}(f)| = 2^{|\text{im}(f)|}.$$

This finite counting fact is the combinatorial backbone of the bounded-interface no-ladder front.

## 2.10 Convention: what counts as evidence

Throughout, theorems are stated for finite objects and refer to the audits defined above. Fitted macro Markov models on  $Y$  and thresholded graph variants are auxiliary constructions rather than theorem objects. The objecthood front is formulated entirely within the stated contraction setup.

## 3 Main Results

We state the eight no-go theorems. Each theorem is followed by a brief remark clarifying scope.

**Theorem 3.1** (NG\_ARROW\_DPI). *Let  $\mu \in \Delta(\mathsf{X})$ , let  $P : \mathsf{X} \times \mathsf{X} \rightarrow [0, 1]$  be row-stochastic, and let  $T \geq 1$ . Define the micro path law on  $\mathsf{X}^{T+1}$  by*

$$\mathbf{P}_{\mu, P}^{(T)}(x_{0:T}) = \mu(x_0) \prod_{t=0}^{T-1} P(x_t, x_{t+1}).$$

*Let  $\text{rev}_T(x_{0:T}) = (x_T, \dots, x_0)$  and let  $\ell : \mathsf{X} \rightarrow \mathsf{Y}$  be deterministic with pointwise extension  $\ell^{(T)}(x_{0:T}) = (\ell(x_0), \dots, \ell(x_T))$ . Then*

$$\text{KL}\left(\left(\ell^{(T)}\right)_{\#} \mathbf{P}_{\mu, P}^{(T)} \parallel \left(\text{rev}_T\right)_{\#} \left(\ell^{(T)}\right)_{\#} \mathbf{P}_{\mu, P}^{(T)}\right) \leq \text{KL}\left(\mathbf{P}_{\mu, P}^{(T)} \parallel \left(\text{rev}_T\right)_{\#} \mathbf{P}_{\mu, P}^{(T)}\right).$$

**Remark 3.2. Interpretation.** This theorem rules out any increase of forward–reverse path-space KL under an honest deterministic observation. In the finite audited setting, coarse observation cannot manufacture a larger arrow than the micro path law already carries.

**Scope / escape routes.** The obstruction is evaded only by leaving the theorem object: stochastic observation, fitted macro chains, or other proxy reconstructions are outside this claim. Real microscopic nonreversibility can still produce positive arrow; the theorem states that deterministic coarse-graining does not amplify it.

**Theorem 3.3** (NG\_PROTOCOL\_TRAP). *Let  $\tilde{\mathsf{X}}$  be finite, let  $\tilde{P} : \tilde{\mathsf{X}} \times \tilde{\mathsf{X}} \rightarrow [0, 1]$  be row-stochastic, and let  $\tilde{\pi}$  be a stationary distribution satisfying detailed balance  $\tilde{\pi}(x)\tilde{P}(x, x') = \tilde{\pi}(x')\tilde{P}(x', x)$  for all  $x, x' \in \tilde{\mathsf{X}}$ . Then for every horizon  $T \geq 1$ ,*

$$\left(\text{rev}_T\right)_{\#} \mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)} = \mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)} \quad \Rightarrow \quad \text{KL}\left(\mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)} \parallel \left(\text{rev}_T\right)_{\#} \mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)}\right) = 0.$$

*Moreover, for every deterministic observation  $\ell : \tilde{\mathsf{X}} \rightarrow \mathsf{Y}$ ,*

$$\text{KL}\left(\left(\ell^{(T)}\right)_{\#} \mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)} \parallel \left(\text{rev}_T\right)_{\#} \left(\ell^{(T)}\right)_{\#} \mathbf{P}_{\tilde{\pi}, \tilde{P}}^{(T)}\right) = 0.$$

**Remark 3.4. Interpretation.** This theorem rules out honest observed arrow for a stationary reversible lift, even when hidden protocol coordinates are present. Hidden protocol structure by itself does not create entropy production in the audited stationary regime.

**Scope / escape routes.** The scope changes once stationarity or reversibility is dropped. Nonstationary initialization, external scheduling, or genuinely driven lifted dynamics fall outside the theorem hypotheses and can reopen positive observed arrow.

**Theorem 3.5 (NG\_FORCE\_FOREST).** *Let  $G = (V, E)$  be a finite undirected forest. Let  $a$  be an antisymmetric function on oriented edges, so that  $a(u, v) = -a(v, u)$  and  $a(u, v) = 0$  whenever  $\{u, v\} \notin E$ . Then there exists a potential  $\phi : V \rightarrow \mathbb{R}$  such that for every oriented edge  $(u, v)$ ,*

$$a(u, v) = \phi(v) - \phi(u).$$

**Remark 3.6. Interpretation.** This theorem rules out force-like cycle obstruction on forest support. On a finite forest, every antisymmetric edge field is potential-generated, so there is no cycle-supported drive to sustain a nontrivial affinity class.

**Scope / escape routes.** The obstruction is evaded by allowing non-forest support or by changing the exact support notion. Positive cycle rank reopens the possibility of nontrivial circulation, while thresholded proxy graphs are outside the theorem object.

**Theorem 3.7 (NG\_FORCE\_NULL).** *Assume  $a(u, v) = \phi(v) - \phi(u)$  on oriented edges of a finite graph. Then for every closed walk  $v_0, \dots, v_k = v_0$ ,*

$$\sum_{i=0}^{k-1} a(v_i, v_{i+1}) = 0.$$

**Remark 3.8. Interpretation.** This theorem rules out cycle affinity when the log-ratio 1-form is exact. Exact ratio-form structure collapses every closed-walk sum, so the force-like class is null in the precise bidirected-support sense of the theorem.

**Scope / escape routes.** The obstruction is evaded by breaking exactness or by replacing the exact support with a thresholded or regularized proxy graph. The theorem is about exact bidirected support and exact ratio-form data, not about diagnostic pictures.

**Theorem 3.9 (NG\_MACRO\_CLOSURE\_DEFICIT).** *Let  $(\mathsf{X}, P)$  be a finite Markov chain, let  $\mu \in \Delta(\mathsf{X})$ , let  $\Pi : \mathsf{X} \rightarrow \mathsf{Y}$  be deterministic, and fix  $\tau \geq 1$ . For each  $x \in \mathsf{X}$ , define the packaged future law*

$$p_x^{(\tau)}(y') := \Pr_{\mu, P}[\Pi(X_{t+\tau}) = y' \mid X_t = x].$$

For a candidate macro kernel  $K : \mathsf{Y} \times \mathsf{Y} \rightarrow [0, 1]$ , define

$$\mathcal{L}_\tau(K) := \sum_{x \in \mathsf{X}} \mu(x) \text{KL}(p_x^{(\tau)} \mid K_{\Pi(x)}).$$

Define the closure deficit by

$$\text{CD}_\tau(\Pi) := I(X_t; \Pi(X_{t+\tau}) \mid \Pi(X_t)).$$

Then the kernel

$$K_y^\star(y') := \Pr_{\mu, P}[\Pi(X_{t+\tau}) = y' \mid \Pi(X_t) = y]$$

satisfies

$$\text{CD}_\tau(\Pi) = \min_K \mathcal{L}_\tau(K) = \mathcal{L}_\tau(K^\star).$$

Moreover, if there exist  $x, x' \in \mathsf{X}$  with  $\Pi(x) = \Pi(x')$ ,  $\mu(x) > 0$ ,  $\mu(x') > 0$ , and  $p_x^{(\tau)} \neq p_{x'}^{(\tau)}$ , then  $\text{CD}_\tau(\Pi) > 0$ .

**Remark 3.10. Interpretation.** This theorem identifies closure deficit as the exact obstruction to a closed macro law at the chosen lag. Positive conditional-information gap means that no single macro kernel captures the packaged future laws without loss.

**Scope / escape routes.** The obstruction is evaded only by changing the package, lag, or underlying dynamics, or by accepting proxy fitted kernels as diagnostics rather than theorem objects. A fitted macro chain may still be useful operationally, but it is not theorem evidence for closure.

**Theorem 3.11 (NG\_OBJECT\_CONTRACTIVE).** *Let  $P$  be a stochastic matrix on finite  $Y$ . Let  $T_P : \Delta(Y) \rightarrow \Delta(Y)$  be the operator  $\nu \mapsto \nu P$ , and let  $\lambda(P)$  be the Dobrushin contraction coefficient in total variation. Assume  $\lambda(P) < 1$ . If  $\nu, \nu' \in \Delta(Y)$  satisfy the  $\varepsilon$ -stability bounds*

$$\text{TV}(\nu, \nu P) \leq \varepsilon, \quad \text{TV}(\nu', \nu' P) \leq \varepsilon,$$

then

$$\text{TV}(\nu, \nu') \leq \frac{2\varepsilon}{1 - \lambda(P)}.$$

In particular, if  $\varepsilon = 0$ , the stationary distribution is unique.

**Remark 3.12. Interpretation.** This theorem bounds epsilon-stable packaged states in a strict-contraction regime and yields uniqueness in the exact fixed-point case. Within the admissible TV/Dobrushin hypotheses, multiple well-separated robust objects cannot coexist at that scale.

**Scope / escape routes.** The obstruction is evaded by moving to noncontractive regimes or by using clustering heuristics that are outside theorem evidence. The claim is deliberately bounded to the contraction setup and does not attempt a general objecthood theory.

**Theorem 3.13 (NG\_LADDER\_IDEM).** *Let  $e : S \rightarrow S$  satisfy  $e \circ e = e$ . Then for all integers  $n \geq 1$ ,*

$$e^{\circ n} = e.$$

**Remark 3.14. Interpretation.** This theorem rules out any iterate ladder from an idempotent packaging map. Once the operator is applied once, later iterates add no new structure because the map has already stabilized.

**Scope / escape routes.** The obstruction is evaded only by changing the operator class or by changing the package itself. Non-idempotent updates, interface growth, or theory changes lie outside the theorem hypotheses.

**Theorem 3.15 (NG\_LADDER\_BOUNDED\_INTERFACE).** *Let  $f : X \rightarrow Y$  be deterministic with finite image  $\text{im}(f)$ . Define*

$$\text{Def}(f) := \{g \circ f : g : Y \rightarrow \{0, 1\}\}.$$

Then

$$|\text{Def}(f)| = 2^{|\text{im}(f)|}.$$

Consequently, if  $|\text{im}(f)| < \infty$ , there is no infinite strictly increasing sequence of pairwise distinct  $f$ -definable predicates.

**Remark 3.16. Interpretation.** This theorem rules out an infinite strictly increasing ladder of predicates once the interface is fixed and finite. The finite image of the lens gives a finite algebra of definable predicates, so internal ladder growth must stabilize.

**Scope / escape routes.** The obstruction is evaded only by changing the theory: lens growth, domain growth, or package change can restart definable expansion. With a fixed finite interface, no unbounded internal ladder remains.

**Remark 3.17** (Supplementary apparatus). Supplementary material collects the reproducibility map, including computational audit artifacts and a formalization-status table. The statements and proofs in the main text do not depend on those materials.

## 4 Arrow audits and protocol traps

This section proves Theorem 3.1 (deterministic arrow cannot increase under observation) and Theorem 3.3 (reversible dynamics at stationarity has zero observed arrow). All objects and notation are as in Section 2. The forward-versus-reversed path-space KL used here is the standard relative-entropy object behind entropy-production and irreversibility diagnostics [14, 17]. Under coarse observation such dissipation can be hidden, which is why the honest pushforward convention matters [7]. The underlying relative-entropy monotonicity is the standard data-processing inequality, and the stationary reversal calculation is standard for reversible chains [3, 9].

### 4.1 Deterministic data processing for finite KL

**Lemma 4.1** (Deterministic DPI for KL). *Let  $\Omega$  be a finite set and let  $g : \Omega \rightarrow \Omega'$  be deterministic. For any probability laws  $P, Q \in \Delta(\Omega)$ ,*

$$D_{\text{KL}}(g_{\#}P \parallel g_{\#}Q) \leq D_{\text{KL}}(P \parallel Q).$$

*Proof.* See Appendix A. □

### 4.2 Reversal commutes with pointwise observation

Fix a horizon  $T \geq 0$ . Let  $\text{rev} : X^{T+1} \rightarrow X^{T+1}$  denote the coordinate reversal map  $\text{rev}(x_0, \dots, x_T) := (x_T, \dots, x_0)$ . Let  $f^{(T)} : X^{T+1} \rightarrow Y^{T+1}$  denote the pointwise application  $f^{(T)}(x_0, \dots, x_T) := (f(x_0), \dots, f(x_T))$ .

**Lemma 4.2** (Commutation of reversal and observation). *For any law  $P \in \Delta(X^{T+1})$ ,*

$$f_{\#}^{(T)}(\text{rev}_{\#}P) = \text{rev}_{\#}(f_{\#}^{(T)}P).$$

*In particular, for the path law  $P_T^{\mu, K}$  from Section 2,*

$$P_T^{\mu, K, f, \text{rev}} = \text{rev}_{\#}P_T^{\mu, K, f} = f_{\#}^{(T)}(P_T^{\mu, K, \text{rev}}).$$

*Proof.* See Appendix A. □

### 4.3 Proof of the arrow DPI theorem

*Proof of Theorem 3.1.* Fix  $T \geq 0$  and abbreviate  $P := P_T^{\mu, K} \in \Delta(X^{T+1})$  and  $Q := P_T^{\mu, K, \text{rev}} = \text{rev}_{\#}P$ . Apply Lemma 4.1 to  $\Omega = X^{T+1}$ ,  $\Omega' = Y^{T+1}$  and  $g = f^{(T)}$ :

$$D_{\text{KL}}(f_{\#}^{(T)}P \parallel f_{\#}^{(T)}Q) \leq D_{\text{KL}}(P \parallel Q).$$

By Lemma 4.2,  $f_{\#}^{(T)}P = P_T^{\mu, K, f}$  and  $f_{\#}^{(T)}Q = P_T^{\mu, K, f, \text{rev}}$ . Therefore the left-hand side equals  $\mathcal{A}_T(\mu, K; f)$  and the right-hand side equals  $\mathcal{A}_T(\mu, K)$ , proving  $\mathcal{A}_T(\mu, K; f) \leq \mathcal{A}_T(\mu, K)$  as claimed. □

#### 4.4 Reversibility at stationarity and the protocol trap

**Lemma 4.3** (Detailed balance implies path-law reversal invariance). *Assume  $\pi \in \Delta(X)$  is stationary for  $K$  and satisfies detailed balance:*

$$\pi(x) K(x, x') = \pi(x') K(x', x) \quad \text{for all } x, x' \in X.$$

Then for every  $T \geq 0$ ,

$$P_T^{\pi, K} = P_T^{\pi, K, \text{rev}}.$$

*Proof.* See Appendix A. □

*Proof of Theorem 3.3.* Lemma 4.3 implies  $P_T^{\pi, K} = P_T^{\pi, K, \text{rev}}$  for every  $T \geq 0$ , hence  $\mathcal{A}_T(\pi, K) = D_{\text{KL}}(P_T^{\pi, K} \parallel P_T^{\pi, K, \text{rev}}) = 0$ . Applying Theorem 3.1 with initial law  $\pi$  gives  $\mathcal{A}_T(\pi, K; f) \leq \mathcal{A}_T(\pi, K) = 0$ , hence  $\mathcal{A}_T(\pi, K; f) = 0$ . □

**Remark 4.4** (What is ruled out, and what is not). Theorem 3.3 is a statement about an *arrow audit*: it rules out certifying irreversibility of the observed process via  $D_{\text{KL}}(P_T^{\pi, K, f} \parallel P_T^{\pi, K, f, \text{rev}})$  when the underlying system is reversible and started in the stationary law. The observed variable  $f(x_t)$  may still carry a nontrivial ordering or protocol structure while having zero arrow under stationary reversible dynamics. To obtain a nonzero arrow audit one must leave this scope, for example by nonstationary initialization, explicit nonequilibrium forcing, or by switching to a different witness.

## 5 Graph structure and affinity

This section isolates the graph-theoretic content behind the affinity and force audits. We work with finite graphs and antisymmetric edge labellings (discrete 1-forms). The two main results proved here are Theorem 3.5 and Theorem 3.7 from Section 3.

### 5.1 Graphs, antisymmetric 1-forms, and cycle sums

Let  $G = (V, E)$  be a finite undirected graph. Write

$$\vec{E} = \{(u, v) \in V \times V : \{u, v\} \in E\}$$

for the set of oriented edges.

**Definition 5.1** (Antisymmetric edge labelling / 1-form). An *antisymmetric edge labelling* on  $G$  is a function  $a : \vec{E} \rightarrow \mathbb{R}$  such that  $a(v, u) = -a(u, v)$  for all  $(u, v) \in \vec{E}$ .

**Definition 5.2** (Exactness and potentials). An antisymmetric labelling  $a$  is *exact* if there exists a function  $\phi : V \rightarrow \mathbb{R}$  such that for every  $(u, v) \in \vec{E}$ ,

$$a(u, v) = \phi(v) - \phi(u).$$

Any such  $\phi$  is called a *potential* for  $a$ .

**Definition 5.3** (Closed-walk sum / cycle affinity). Let  $w = (v_0, v_1, \dots, v_k)$  be a finite walk with  $v_k = v_0$  and  $\{v_i, v_{i+1}\} \in E$  for all  $i$ . The *closed-walk sum* of  $a$  along  $w$  is

$$S_a(w) = \sum_{i=0}^{k-1} a(v_i, v_{i+1}).$$

When  $a$  is a log-ratio 1-form,  $S_a(w)$  is the (log) cycle affinity. This is the standard cycle-affinity language for finite Markov networks [16].

**Definition 5.4** (Cycle rank). Let  $c(G)$  be the number of connected components of  $G$ . The *cycle rank* (first Betti number) of  $G$  is

$$\beta_1(G) = |E| - |V| + c(G).$$

In particular,  $\beta_1(G) = 0$  if and only if every component of  $G$  is a tree. These cycle-space facts are standard in graph theory; see [1, 5].

## 5.2 Forest exactness

We now prove that on a tree (hence on each component of a forest) every antisymmetric edge labelling is potential-generated. The potential interpretation of exact antisymmetric edge labellings is also standard in this finite graph setting; see [5, 13].

**Lemma 5.5** (Rooted parent map on a tree). *Let  $G = (V, E)$  be a finite tree and fix a root  $r \in V$ . For each vertex  $v \neq r$  there exists a unique neighbor  $\text{par}(v)$  of  $v$  lying on the unique simple path from  $r$  to  $v$ .*

*Proof.* See Appendix B. □

*Proof of Theorem 3.5.* Let  $G = (V, E)$  be a finite tree with root  $r$ , and let  $a : \vec{E} \rightarrow \mathbb{R}$  be antisymmetric. Define  $\phi : V \rightarrow \mathbb{R}$  by  $\phi(r) = 0$  and, for  $v \neq r$ ,

$$\phi(v) = \phi(\text{par}(v)) + a(\text{par}(v), v),$$

where  $\text{par}(v)$  is given by Lemma 5.5. The verification that  $a(u, v) = \phi(v) - \phi(u)$  on every oriented edge is recorded in Appendix B. For a forest, apply the same construction componentwise by choosing a root in each component. □

## 5.3 Exactness forces vanishing closed-walk sums

*Proof of Theorem 3.7.* Assume  $a$  is exact:  $a(u, v) = \phi(v) - \phi(u)$  for some  $\phi : V \rightarrow \mathbb{R}$ . Let  $w = (v_0, v_1, \dots, v_k)$  be any closed walk with  $v_k = v_0$ . Then

$$S_a(w) = \sum_{i=0}^{k-1} (\phi(v_{i+1}) - \phi(v_i)) = \phi(v_k) - \phi(v_0) = 0.$$

The telescoping calculation is written out in Appendix B. □

**Remark 5.6** (Connection to log-ratio affinity). If  $K$  is a Markov kernel on a finite state space with bidirected support graph  $G$ , a common affinity 1-form is the log-ratio

$$a_K(u, v) = \log \frac{K(u, v)}{K(v, u)} \quad \text{on each bidirected edge.}$$

This is antisymmetric on  $\vec{E}$ . Theorems 3.5 and 3.7 show that (i) on forest support every such antisymmetric form is potential-generated, and (ii) any potential-generated form has zero closed-walk sums, so every cycle affinity vanishes.

**Remark 5.7** (Thresholding and regularization). The theorems above are statements about the *exact* graph object  $(V, E)$  and the *exact* antisymmetric data  $a$  supported on  $\vec{E}$ . Graph thresholding or regularization changes  $E$  and can change  $\beta_1(G)$ ; such modifications lie outside the theorem hypotheses and define different graph objects.

## 6 Closure deficit and a variational obstruction

This section proves Theorem 3.9. The setting is finite throughout. Let  $X$  be a finite state space, let  $Y$  be a finite macro space, and let  $(X_t)_{t \in \mathbb{Z}}$  be a stationary Markov chain on  $X$  with stationary law  $\mu$ . Fix a deterministic packaging map  $\Pi : X \rightarrow Y$  and define the macro process by  $Y_t := \Pi(X_t)$ . For an integer lag  $\tau \geq 1$  and a microstate  $x \in X$ , write

$$p_x^{(\tau)}(\cdot) := \text{Law}(Y_{t+\tau} \mid X_t = x) \in \Delta(Y).$$

For a macrostate  $y \in Y$ , write

$$\bar{p}_y^{(\tau)}(\cdot) := \text{Law}(Y_{t+\tau} \mid Y_t = y) \in \Delta(Y).$$

A macro kernel is a function  $K : Y \rightarrow \Delta(Y)$ , equivalently a row-stochastic array  $K(y, y')$  on  $Y \times Y$ . The variational objective at lag  $\tau$  is

$$\mathcal{L}_\tau(K) := \sum_{x \in X} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel K(\Pi(x), \cdot)\right).$$

The KL decompositions and conditional-mutual-information identities used below are standard finite information-theoretic calculations, and the variational minimizer is a finite KL-projection instance [3, 4].

**Lemma 6.1** (KL mixture decomposition). *Let  $(w_i)_{i \in I}$  be nonnegative weights on a finite index set  $I$ , not all zero, and let  $(P_i)_{i \in I}$  be probability laws on a finite set  $\Omega$ . Define*

$$W := \sum_{i \in I} w_i, \quad R := \frac{1}{W} \sum_{i \in I} w_i P_i.$$

Then for every probability law  $Q$  on  $\Omega$ ,

$$\sum_{i \in I} w_i D_{\text{KL}}(P_i \parallel Q) = \sum_{i \in I} w_i D_{\text{KL}}(P_i \parallel R) + W D_{\text{KL}}(R \parallel Q).$$

In particular,  $Q = R$  is a minimizer, and the minimum value is

$$\sum_{i \in I} w_i D_{\text{KL}}(P_i \parallel R).$$

*Proof.* See Appendix C. □

**Lemma 6.2** (Best macro kernel). *For each  $y \in Y$ , define*

$$K_\tau^*(y, \cdot) := \bar{p}_y^{(\tau)}(\cdot) = \text{Law}(Y_{t+\tau} \mid Y_t = y).$$

Then

$$\min_K \mathcal{L}_\tau(K) = \sum_{x \in X} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel \bar{p}_{\Pi(x)}^{(\tau)}\right),$$

and the minimum is attained by  $K_\tau^*$ .

*Proof.* See Appendix C. □

**Lemma 6.3** (Conditional mutual information identity). *In this finite deterministic setting,*

$$I(X_t; Y_{t+\tau} \mid Y_t) = \sum_{x \in X} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel \bar{p}_{\Pi(x)}^{(\tau)}\right).$$

*Proof.* See Appendix C. □

*Proof of Theorem 3.9.* By definition, the closure deficit in the theorem is

$$\text{CD}_\tau(\Pi) = I(X_t; \Pi(X_{t+\tau}) \mid \Pi(X_t)) = I(X_t; Y_{t+\tau} \mid Y_t).$$

Lemma 6.3 expresses this quantity as

$$\text{CD}_\tau(\Pi) = \sum_{x \in X} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel \bar{p}_{\Pi(x)}^{(\tau)}\right).$$

Lemma 6.2 identifies the same expression as the minimum of the variational objective:

$$\text{CD}_\tau(\Pi) = \min_K \mathcal{L}_\tau(K), \quad K_\tau^*(y, \cdot) = \bar{p}_y^{(\tau)}(\cdot).$$

This proves the finite variational identity and the formula for the minimizing macro kernel.

For the positivity claim, suppose there exist  $x, x' \in X$  with  $\Pi(x) = \Pi(x') =: y$ ,  $\mu(x) > 0$ ,  $\mu(x') > 0$ , and  $p_x^{(\tau)} \neq p_{x'}^{(\tau)}$ . Then not all positive-mass fiber laws in  $\Pi^{-1}(y)$  coincide with their weighted average  $\bar{p}_y^{(\tau)}$ . Hence at least one state  $z \in \Pi^{-1}(y)$  with  $\mu(z) > 0$  satisfies

$$D_{\text{KL}}\left(p_z^{(\tau)} \parallel \bar{p}_y^{(\tau)}\right) > 0.$$

The sum formula for  $\text{CD}_\tau(\Pi)$  therefore contains at least one strictly positive term and all remaining terms are nonnegative, so

$$\text{CD}_\tau(\Pi) > 0. \quad \square$$

**Remark 6.4** (Scope of exact closure). If every macro fiber is a singleton, then  $\Pi$  loses no information and the closure deficit vanishes trivially. More generally, if all packaged future laws  $p_x^{(\tau)}$  agree inside each fiber  $\Pi^{-1}(y)$ , then the minimizing macro kernel reproduces those common laws and  $\text{CD}_\tau(\Pi) = 0$ . The theorem becomes informative exactly when a macro fiber contains distinct packaged future laws.

**Remark 6.5** (Attribution is separate). The theorem isolates a finite obstruction: distinct packaged future laws inside a single macro fiber force a positive closure deficit. Which primitive or modeling decision produces such a pair is a separate question and is independent of the variational identity proved here.

## 7 Objecthood under strict contraction

This section proves Theorem 3.11. The theorem is formulated on a finite simplex equipped with total variation distance and a strict Dobrushin contraction bound. Within this admissible regime, the induced Markov operator has a unique fixed distribution, its iterates converge geometrically, and any two  $\varepsilon$ -approximate fixed points must lie within an explicit separation bound.

Let  $Y$  be a finite set. For  $\mu, \nu \in \Delta(Y)$ , define the total variation distance by

$$d_{\text{TV}}(\mu, \nu) := \frac{1}{2} \sum_{y \in Y} |\mu(y) - \nu(y)|.$$

We also use the equivalent variational form

$$d_{\text{TV}}(\mu, \nu) = \sup_{A \subseteq Y} (\mu(A) - \nu(A)).$$

Let  $P : Y \rightarrow \Delta(Y)$  be a Markov kernel and let  $T : \Delta(Y) \rightarrow \Delta(Y)$  be the induced affine map

$$T(\mu) := \mu P.$$

The Dobrushin coefficient of  $P$  is

$$\lambda(P) := \max_{y, y' \in Y} d_{\text{TV}}(P(y, \cdot), P(y', \cdot)).$$

This contraction coefficient and its standard consequences for finite chains are discussed in [6, 18].

**Lemma 7.1** (Dobrushin contraction in total variation). *For all  $\mu, \nu \in \Delta(Y)$ ,*

$$d_{\text{TV}}(\mu P, \nu P) \leq \lambda(P) d_{\text{TV}}(\mu, \nu).$$

*Proof.* See Appendix A. □

**Lemma 7.2** (Strict contraction gives a unique fixed distribution). *Assume  $\lambda(P) < 1$ . Then there exists a unique distribution  $\pi \in \Delta(Y)$  with  $\pi P = \pi$ . Moreover, for every  $\mu \in \Delta(Y)$  and every integer  $t \geq 0$ ,*

$$d_{\text{TV}}(\mu P^t, \pi) \leq \lambda(P)^t d_{\text{TV}}(\mu, \pi).$$

*Proof.* See Appendix A. □

**Lemma 7.3** (Approximate fixed-point separation bound). *Assume  $\lambda(P) \leq \lambda < 1$ . If  $\mu, \nu \in \Delta(Y)$  satisfy*

$$d_{\text{TV}}(\mu P, \mu) \leq \varepsilon, \quad d_{\text{TV}}(\nu P, \nu) \leq \varepsilon,$$

*then*

$$d_{\text{TV}}(\mu, \nu) \leq \frac{2\varepsilon}{1 - \lambda}.$$

*Proof.* See Appendix A. □

*Proof of Theorem 3.11.* The uniqueness and convergence claims follow from Lemma 7.2 once  $\lambda(P) < 1$ . For the approximate objecthood statement, suppose  $\mu$  and  $\nu$  are  $\varepsilon$ -stable packaged states in the admissible TV/Dobrushin regime. Lemma 7.3 then gives the bound

$$d_{\text{TV}}(\mu, \nu) \leq \frac{2\varepsilon}{1 - \lambda(P)}.$$

In particular, when  $\varepsilon = 0$  every fixed distribution coincides with the unique stationary distribution from Lemma 7.2. Thus the theorem rules out multiplicity of well-separated robust objects inside the stated contraction regime. □

**Remark 7.4** (Regime and metric scope). This front is formulated in total variation distance with the Dobrushin coefficient as the contraction parameter. The theorem asserts an obstruction only in that regime. Other metrics, other contraction notions, or noncontractive operators lie outside the scope of this statement.

**Remark 7.5** (Formalization status). For this front, the formalized component is the uniqueness consequence of strict contraction for fixed distributions. The total-variation/Dobrushin framework and the  $\varepsilon$ -stability separation bound are used here at the level of the paper proof and are not part of the formalized core.

## 8 Bounded interfaces and no-ladder obstructions

This section isolates two finite obstructions to indefinite refinement under audited closure. The first is a simple idempotence observation. The second is a counting argument: under a fixed deterministic interface, there are only finitely many definable predicates, hence no infinite strictly increasing ladder of distinct definable distinctions.

### 8.1 Idempotent saturation

Let  $X$  be a finite set and let  $e : X \rightarrow X$  be a deterministic packaging operator. For an integer  $n \geq 1$ , write  $e^{on}$  for the  $n$ -fold iterate  $e^{on} = \underbrace{e \circ \dots \circ e}_{n \text{ times}}$ .

**Lemma 8.1.** *If  $e \circ e = e$ , then  $e^{on} = e$  for all integers  $n \geq 1$ .*

*Proof.* See Appendix D. □

*Proof of Theorem 3.13.* Fix  $n \geq 1$ . By Lemma 8.1 we have  $e^{on} = e$ , hence for every  $x \in X$ ,

$$e^{on}(x) = e(x).$$

In particular, the packaged state stabilizes after one step, and any packaged signature depending only on  $e^{on}(x)$  stabilizes after one step as well. □

### 8.2 Definability count for a fixed interface

Let  $f : X \rightarrow Y$  be a deterministic interface observation into a finite set  $Y$ . Define the family of predicates definable from  $f$  by

$$\text{Def}(f) := \{g \circ f : X \rightarrow \{0, 1\} \mid g : Y \rightarrow \{0, 1\}\}.$$

Only the restriction of  $g$  to the image  $\text{im}(f) \subseteq Y$  matters.

**Lemma 8.2.** *The map*

$$A \subseteq \text{im}(f) \quad \longmapsto \quad \mathbf{1}_A \circ f$$

*is a bijection between  $\mathcal{P}(\text{im}(f))$  and  $\text{Def}(f)$ . Consequently,*

$$|\text{Def}(f)| = 2^{|\text{im}(f)|}.$$

*Proof.* See Appendix D. □

*Proof of Theorem 3.15.* Assume for contradiction that there exists an infinite strictly increasing sequence  $(h_n)_{n \geq 1} \subseteq \text{Def}(f)$ , meaning  $h_n(x) \leq h_{n+1}(x)$  for all  $x \in X$  and  $h_n \neq h_{n+1}$  for each  $n$ . By Lemma 8.2, the set  $\text{Def}(f)$  has finite cardinality  $2^{|\text{im}(f)|}$ , so it cannot contain infinitely many distinct elements. This contradiction shows that no such infinite strictly increasing sequence exists. □

**Remark 8.3** (Extension escape). Theorem 3.15 assumes a fixed interface map  $f$ . If one allows the interface to vary with the step index (for example by adding coordinates or refining the lens family so that  $|\text{im}(f_n)|$  grows), then the bound  $|\text{Def}(f_n)| = 2^{|\text{im}(f_n)|}$  can increase and longer refinement chains become possible. Such interface expansion is outside the scope of the bounded-interface front.

## 9 Discussion

The eight theorem fronts delimit a finite obstruction layer for audited emergence claims. In the finite deterministic setting treated here, directed arrow, force-like affinity, exact closure, objecthood under contraction, and bounded-interface laddering are all constrained by explicit theorem hypotheses and explicit escape routes. The results therefore organize the finite obstruction structure of the theory: each front identifies a precise hypothesis set and the corresponding obstruction it entails.

Supplementary material records provenance, computational audit artifacts, fragility diagnostics, and formalization status. The main paper reserves its appendices for finite probability calculations, graph exactness derivations, variational closure identities, and packaging/definability details.

## A Finite probability, KL, and contraction calculations

This appendix records the finite calculations used by the arrow, protocol, and contraction fronts.

*Proof of Lemma 4.1.* For  $y \in \Omega'$ , write  $A_y := \{x \in \Omega : g(x) = y\}$  and abbreviate  $p_x := P(x)$ ,  $q_x := Q(x)$ ,  $p_y := \sum_{x \in A_y} p_x$ ,  $q_y := \sum_{x \in A_y} q_x$ . By definition,

$$D_{\text{KL}}(g_{\#}P \parallel g_{\#}Q) = \sum_{y \in \Omega': p_y > 0} p_y \log\left(\frac{p_y}{q_y}\right).$$

For each  $y$  with  $p_y > 0$ , the log-sum inequality gives

$$p_y \log\left(\frac{p_y}{q_y}\right) \leq \sum_{x \in A_y: p_x > 0} p_x \log\left(\frac{p_x}{q_x}\right),$$

with the usual convention that the right-hand side is  $+\infty$  if some  $p_x > 0$  but  $q_x = 0$ . Summing over  $y$  and using that the sets  $A_y$  partition  $\Omega$  yields

$$D_{\text{KL}}(g_{\#}P \parallel g_{\#}Q) \leq \sum_{x \in \Omega: p_x > 0} p_x \log\left(\frac{p_x}{q_x}\right) = D_{\text{KL}}(P \parallel Q).$$

□

*Proof of Lemma 4.2.* The maps  $\text{rev}$  and  $f^{(T)}$  commute pointwise:  $f^{(T)} \circ \text{rev} = \text{rev} \circ f^{(T)}$ . Pushforward respects composition, hence

$$f_{\#}^{(T)}(\text{rev}_{\#}P) = (f^{(T)} \circ \text{rev})_{\#}P = (\text{rev} \circ f^{(T)})_{\#}P = \text{rev}_{\#}(f_{\#}^{(T)}P).$$

Substituting  $P = P_T^{\mu, K}$  and using the definitions from Section 2 gives

$$P_T^{\mu, K, f, \text{rev}} = \text{rev}_{\#}P_T^{\mu, K, f} = f_{\#}^{(T)}(P_T^{\mu, K, \text{rev}}).$$

□

*Proof of Lemma 4.3.* Fix  $T \geq 0$  and a path  $\gamma = (x_0, \dots, x_T) \in X^{T+1}$ . By definition,

$$P_T^{\pi, K}(\gamma) = \pi(x_0) \prod_{t=0}^{T-1} K(x_t, x_{t+1}), \quad P_T^{\pi, K, \text{rev}}(\gamma) = P_T^{\pi, K}(x_T, \dots, x_0) = \pi(x_T) \prod_{t=0}^{T-1} K(x_{t+1}, x_t).$$

Using detailed balance repeatedly,

$$\pi(x_0)K(x_0, x_1) = \pi(x_1)K(x_1, x_0), \quad \pi(x_1)K(x_1, x_2) = \pi(x_2)K(x_2, x_1), \quad \dots$$

Multiplying these identities for  $t = 0, \dots, T-1$  yields

$$\pi(x_0) \prod_{t=0}^{T-1} K(x_t, x_{t+1}) = \pi(x_T) \prod_{t=0}^{T-1} K(x_{t+1}, x_t),$$

which is exactly  $P_T^{\pi, K}(\gamma) = P_T^{\pi, K, \text{rev}}(\gamma)$ . Since this holds for every  $\gamma$ , the two path laws are equal.  $\square$

*Proof of Lemma 7.1.* Fix  $A \subseteq Y$  and write  $\eta := \mu - \nu$ . Since  $\eta(Y) = 0$ , if  $y_0 \in Y$  is arbitrary then

$$(\mu P)(A) - (\nu P)(A) = \sum_{y \in Y} \eta(y) P(y, A) = \sum_{y \in Y} \eta(y) (P(y, A) - P(y_0, A)).$$

Let  $Y_+ := \{y \in Y : \eta(y) \geq 0\}$  and  $Y_- := Y \setminus Y_+$ . Because  $\sum_y \eta(y) = 0$ ,

$$\sum_{y \in Y_+} \eta(y) = \sum_{y \in Y_-} (-\eta(y)) = d_{\text{TV}}(\mu, \nu).$$

Hence

$$\begin{aligned} (\mu P)(A) - (\nu P)(A) &\leq \sum_{y \in Y_+} \eta(y) \sup_{u, v \in Y} (P(u, A) - P(v, A)) \\ &\leq d_{\text{TV}}(\mu, \nu) \sup_{u, v \in Y} (P(u, A) - P(v, A)). \end{aligned}$$

For fixed  $u, v$ , the variational characterization of total variation gives

$$P(u, A) - P(v, A) \leq d_{\text{TV}}(P(u, \cdot), P(v, \cdot)) \leq \lambda(P).$$

Therefore

$$(\mu P)(A) - (\nu P)(A) \leq \lambda(P) d_{\text{TV}}(\mu, \nu) \quad \text{for every } A \subseteq Y.$$

Taking the supremum over  $A$  yields the stated inequality.  $\square$

*Proof of Lemma 7.2.* Fix  $\mu \in \Delta(Y)$  and set  $\mu_t := \mu P^t$ . By Lemma 7.1,

$$d_{\text{TV}}(\mu_{t+1}, \mu_t) = d_{\text{TV}}(\mu_t P, \mu_{t-1} P) \leq \lambda(P) d_{\text{TV}}(\mu_t, \mu_{t-1})$$

for every  $t \geq 1$ . Inducting on  $t$  gives

$$d_{\text{TV}}(\mu_{t+1}, \mu_t) \leq \lambda(P)^t d_{\text{TV}}(\mu P, \mu).$$

Hence for integers  $m > n$ ,

$$\begin{aligned} d_{\text{TV}}(\mu_m, \mu_n) &\leq \sum_{k=n}^{m-1} d_{\text{TV}}(\mu_{k+1}, \mu_k) \\ &\leq d_{\text{TV}}(\mu P, \mu) \sum_{k=n}^{m-1} \lambda(P)^k. \end{aligned}$$

Since  $0 \leq \lambda(P) < 1$ , the geometric tail tends to zero as  $n \rightarrow \infty$ . Thus  $(\mu_t)_{t \geq 0}$  is Cauchy in total variation. Because  $\Delta(Y)$  is a closed subset of the finite-dimensional space  $\mathbb{R}^Y$ , it is complete for this metric, so there exists  $\pi \in \Delta(Y)$  with  $\mu_t \rightarrow \pi$ . The map  $T(\rho) = \rho P$  is continuous, hence

$$\pi P = \lim_{t \rightarrow \infty} \mu_{t+1} = \lim_{t \rightarrow \infty} \mu_t = \pi.$$

So  $\pi$  is a fixed distribution.

If  $\pi'$  is another fixed distribution, then Lemma 7.1 gives

$$d_{\text{TV}}(\pi, \pi') = d_{\text{TV}}(\pi P, \pi' P) \leq \lambda(P) d_{\text{TV}}(\pi, \pi').$$

Since  $\lambda(P) < 1$ , this forces  $d_{\text{TV}}(\pi, \pi') = 0$ , hence  $\pi = \pi'$ . This proves uniqueness.

Finally, applying Lemma 7.1 with  $\nu = \pi$  and using  $\pi P = \pi$  yields

$$d_{\text{TV}}(\mu P^t, \pi) = d_{\text{TV}}(\mu P^t, \pi P^t) \leq \lambda(P)^t d_{\text{TV}}(\mu, \pi).$$

□

*Proof of Lemma 7.3.* By the triangle inequality,

$$d_{\text{TV}}(\mu, \nu) \leq d_{\text{TV}}(\mu, \mu P) + d_{\text{TV}}(\mu P, \nu P) + d_{\text{TV}}(\nu P, \nu).$$

Using the hypotheses and Lemma 7.1,

$$d_{\text{TV}}(\mu, \nu) \leq \varepsilon + \lambda(P) d_{\text{TV}}(\mu, \nu) + \varepsilon \leq 2\varepsilon + \lambda d_{\text{TV}}(\mu, \nu).$$

Rearranging gives

$$(1 - \lambda) d_{\text{TV}}(\mu, \nu) \leq 2\varepsilon,$$

and division by  $1 - \lambda > 0$  proves the claim. □

## B Graph exactness and closed-walk derivations

This appendix records the graph calculations used by the force and affinity fronts.

*Proof of Lemma 5.5.* Because  $G$  is a tree it is connected and acyclic, hence for any  $v$  there is a unique simple path from  $r$  to  $v$ . If  $v \neq r$ , the final edge of that path has the form  $\{\text{par}(v), v\}$  for a neighbor  $\text{par}(v)$  of  $v$ . Uniqueness of the path gives uniqueness of  $\text{par}(v)$ . □

**Proposition B.1** (Potential construction on a rooted tree). *Let  $G = (V, E)$  be a finite tree with root  $r$ , and let  $a : \vec{E} \rightarrow \mathbb{R}$  be antisymmetric. Define  $\phi(r) = 0$  and, for  $v \neq r$ ,*

$$\phi(v) = \phi(\text{par}(v)) + a(\text{par}(v), v).$$

*Then  $a(u, v) = \phi(v) - \phi(u)$  for every oriented edge  $(u, v) \in \vec{E}$ .*

*Proof.* Fix an undirected edge  $\{u, v\} \in E$ . Since  $G$  is a rooted tree, exactly one of  $u, v$  is the parent of the other. If  $\text{par}(v) = u$ , then by definition,  $\phi(v) - \phi(u) = a(u, v)$ . If instead  $\text{par}(u) = v$ , then  $\phi(u) - \phi(v) = a(v, u)$ , hence by antisymmetry  $a(u, v) = -a(v, u) = \phi(v) - \phi(u)$ . Thus the identity holds on every oriented edge. □

*Expanded proof of Theorem 3.5.* For a tree, Proposition B.1 gives the potential explicitly. For a forest, apply the same rooted construction on each connected component after choosing one root per component. The resulting componentwise potentials patch together to a global potential because no edge joins distinct components.  $\square$

*Expanded proof of Theorem 3.7.* Assume  $a$  is exact:  $a(u, v) = \phi(v) - \phi(u)$  for some  $\phi : V \rightarrow \mathbb{R}$ . Let  $w = (v_0, v_1, \dots, v_k)$  be a closed walk with  $v_k = v_0$ . Then

$$S_a(w) = \sum_{i=0}^{k-1} a(v_i, v_{i+1}) = \sum_{i=0}^{k-1} (\phi(v_{i+1}) - \phi(v_i)) = \phi(v_k) - \phi(v_0) = 0.$$

The sum telescopes because each interior term appears once with positive sign and once with negative sign.  $\square$

## C Variational closure identities

This appendix records the finite variational calculations used by the closure front.

*Proof of Lemma 6.1.* Expand both Kullback–Leibler sums on the finite support:

$$\begin{aligned} \sum_{i \in I} w_i D_{\text{KL}}(P_i \parallel Q) &= \sum_{i \in I} \sum_{\omega \in \Omega} w_i P_i(\omega) \log \frac{P_i(\omega)}{Q(\omega)}, \\ \sum_{i \in I} w_i D_{\text{KL}}(P_i \parallel R) &= \sum_{i \in I} \sum_{\omega \in \Omega} w_i P_i(\omega) \log \frac{P_i(\omega)}{R(\omega)}. \end{aligned}$$

Subtracting the second identity from the first gives

$$\sum_{i \in I} \sum_{\omega \in \Omega} w_i P_i(\omega) \log \frac{R(\omega)}{Q(\omega)} = \sum_{\omega \in \Omega} \left( \sum_{i \in I} w_i P_i(\omega) \right) \log \frac{R(\omega)}{Q(\omega)}.$$

Since  $\sum_i w_i P_i(\omega) = WR(\omega)$  by definition of  $R$ , this becomes

$$W \sum_{\omega \in \Omega} R(\omega) \log \frac{R(\omega)}{Q(\omega)} = W D_{\text{KL}}(R \parallel Q).$$

Rearranging yields the displayed identity. Because Kullback–Leibler divergence is nonnegative, the final term is minimized at  $Q = R$ , and then the minimum value is exactly  $\sum_i w_i D_{\text{KL}}(P_i \parallel R)$ .  $\square$

*Proof of Lemma 6.2.* Decompose the objective by macro fibers:

$$\mathcal{L}_\tau(K) = \sum_{y \in Y} \sum_{x: \Pi(x)=y} \mu(x) D_{\text{KL}}(p_x^{(\tau)} \parallel K(y, \cdot)).$$

Fix a macrostate  $y$  with  $\mu(\Pi^{-1}(y)) > 0$ . Apply Lemma 6.1 on the fiber  $\Pi^{-1}(y)$  with indices  $i = x$ , weights  $w_x = \mu(x)$ , laws  $P_x = p_x^{(\tau)}$ , and comparison law  $Q = K(y, \cdot)$ . The corresponding mixture law is

$$R_y = \frac{1}{\sum_{x: \Pi(x)=y} \mu(x)} \sum_{x: \Pi(x)=y} \mu(x) p_x^{(\tau)}.$$

By the finite law of total probability and the identity  $Y_t = \Pi(X_t)$ , this mixture is exactly  $\text{Law}(Y_{t+\tau} \mid Y_t = y) = \bar{p}_y^{(\tau)}$ . Hence the minimizing row is  $K(y, \cdot) = \bar{p}_y^{(\tau)}$ , and the minimum contribution of the fiber is

$$\sum_{x:\Pi(x)=y} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel \bar{p}_y^{(\tau)}\right).$$

Summing over all  $y$  gives the formula for  $\min_K \mathcal{L}_\tau(K)$ . If  $\mu(\Pi^{-1}(y)) = 0$ , the row  $K(y, \cdot)$  contributes zero and may be chosen arbitrarily; setting it to  $\bar{p}_y^{(\tau)}$  still gives a minimizer.  $\square$

*Proof of Lemma 6.3.* Expand conditional mutual information into a finite sum:

$$I(X_t; Y_{t+\tau} \mid Y_t) = \sum_{x \in X} \sum_{y \in Y} \sum_{y' \in Y} \Pr[X_t = x, Y_t = y, Y_{t+\tau} = y'] \log \frac{\Pr[Y_{t+\tau} = y' \mid X_t = x, Y_t = y]}{\Pr[Y_{t+\tau} = y' \mid Y_t = y]}.$$

Because  $Y_t = \Pi(X_t)$  is a deterministic function of  $X_t$ , the joint probability vanishes unless  $y = \Pi(x)$ , and conditioning on  $(X_t = x, Y_t = \Pi(x))$  is the same as conditioning on  $X_t = x$  alone. Therefore the sum reduces to

$$\sum_{x \in X} \sum_{y' \in Y} \Pr[X_t = x, Y_{t+\tau} = y'] \log \frac{\Pr[Y_{t+\tau} = y' \mid X_t = x]}{\Pr[Y_{t+\tau} = y' \mid Y_t = \Pi(x)]}.$$

Using stationarity,  $\Pr[X_t = x] = \mu(x)$ . Using the definitions of  $p_x^{(\tau)}$  and  $\bar{p}_{\Pi(x)}^{(\tau)}$ , this becomes

$$\sum_{x \in X} \mu(x) \sum_{y' \in Y} p_x^{(\tau)}(y') \log \frac{p_x^{(\tau)}(y')}{\bar{p}_{\Pi(x)}^{(\tau)}(y')}.$$

This is exactly

$$\sum_{x \in X} \mu(x) D_{\text{KL}}\left(p_x^{(\tau)} \parallel \bar{p}_{\Pi(x)}^{(\tau)}\right).$$

$\square$

## D Packaging iterations and definability counts

This appendix records the finite calculations used by the ladder fronts.

*Proof of Lemma 8.1.* Proceed by induction on  $n$ . For  $n = 1$  the claim is immediate. If  $e^{\circ n} = e$ , then

$$e^{\circ(n+1)} = e^{\circ n} \circ e = e \circ e = e,$$

using idempotence. This proves  $e^{\circ n} = e$  for all integers  $n \geq 1$ .  $\square$

*Proof of Lemma 8.2.* Surjectivity: Let  $h \in \text{Def}(f)$ , so  $h = g \circ f$  for some  $g : Y \rightarrow \{0, 1\}$ . Define

$$A := \{y \in \text{im}(f) : g(y) = 1\} \subseteq \text{im}(f).$$

For any  $x \in X$  we have  $h(x) = g(f(x)) = 1$  iff  $f(x) \in A$ , hence  $h = \mathbf{1}_A \circ f$ .

Injectivity: If  $A, B \subseteq \text{im}(f)$  and  $\mathbf{1}_A \circ f = \mathbf{1}_B \circ f$ , then for any  $y \in \text{im}(f)$  pick  $x \in X$  with  $f(x) = y$ . Evaluating at  $x$  gives  $\mathbf{1}_A(y) = \mathbf{1}_B(y)$ , so  $A = B$ . The cardinality statement follows because  $|\mathcal{P}(\text{im}(f))| = 2^{|\text{im}(f)|}$ .  $\square$

**Proposition D.1** (Finite strictly increasing chains are bounded). *If  $S$  is a finite set, then every strictly increasing chain of distinct elements of  $S$  has length at most  $|S|$ .*

*Proof.* A strictly increasing chain of distinct elements is in particular an injective sequence in  $S$ . A finite set of size  $|S|$  admits no injective sequence longer than  $|S|$ .  $\square$

*Expanded proof of Theorem 3.15.* Lemma 8.2 identifies  $\text{Def}(f)$  with the finite power set  $\mathcal{P}(\text{im}(f))$ . Therefore Proposition D.1 applies with  $S = \text{Def}(f)$  and shows that no strictly increasing chain of distinct definable predicates can be infinite.  $\square$

## References

- [1] Norman Biggs. *Algebraic Graph Theory*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2 edition, 1993.
- [2] Peter Buchholz. Exact and ordinary lumpability in finite markov chains. *Journal of Applied Probability*, 31(1):59–75, 1994.
- [3] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, Hoboken, NJ, 2 edition, 2006.
- [4] Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146–158, 1975.
- [5] Reinhard Diestel. *Graph Theory*, volume 173 of *Graduate Texts in Mathematics*. Springer, 5 edition, 2017. doi:10.1007/978-3-662-53622-3.
- [6] R. L. Dobrushin. Central limit theorem for nonstationary markov chains. i. *Theory of Probability and Its Applications*, 1(1):65–80, 1956.
- [7] Massimiliano Esposito. Stochastic thermodynamics under coarse graining. *Physical Review E*, 85(4):041125, 2012.
- [8] J. Nathan Foster, Michael B. Greenwald, Jonathan T. Moore, Benjamin C. Pierce, and Alan Schmitt. Combinators for bidirectional tree transformations: A linguistic approach to the view-update problem. *ACM Transactions on Programming Languages and Systems*, 29(3):17, 2007.
- [9] Frank P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, 1979.
- [10] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. D. Van Nostrand, Princeton, NJ, 1960.
- [11] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951. doi:10.1214/aoms/1177729694.
- [12] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2 edition, 2017.
- [13] Lek-Heng Lim. Hodge laplacians on graphs. *SIAM Review*, 62(3):685–715, 2020.
- [14] Christian Maes and Karel Netočný. Time-reversal and entropy. *Journal of Statistical Physics*, 110(1–2):269–310, 2003.

- [15] J. R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, 1997.
- [16] Jürgen Schnakenberg. Network theory of microscopic and macroscopic behavior of master equation systems. *Reviews of Modern Physics*, 48(4):571–585, 1976.
- [17] Udo Seifert. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Reports on Progress in Physics*, 75(12):126001, 2012.
- [18] Eugene Seneta. *Non-Negative Matrices and Markov Chains*. Springer, New York, revised edition, 2006.
- [19] Ioannis Tsiokos. To cast a stone with six birds: A closure deficit account of randomness under packaging and budget, 2026.
- [20] Ioannis Tsiokos. To classify a stone with six birds: Primitive activation and observable separation of layer birth classes, 2026.
- [21] Ioannis Tsiokos. Six birds: Foundations of emergence calculus, 2026.
- [22] Ioannis Tsiokos. To lay a stone with six birds: Finite-state semantics for packaging, directionality, and coarse graining, 2026. Preprint.
- [23] Ioannis Tsiokos. Six birds protocol trap: Holonomy without entropy production, 2026.
- [24] Ioannis Tsiokos. To throw a stone with six birds: On agents and agenthood, 2026.