ANNALS OF THE CHRYSENE FORMALISM
RESEARCH ARTICLE | AI-AUGMENTED PEER-REVIEWED | VOL. 1, NO. 4 | NOVEMBER 2026

Endogenous and Exogenous Sources of Encoded Chrysene Systems

Functional Selection, Adaptive Realization, and Engineered Emulation

Charles D. Schaper, Ph.D.
Chief Editor, Annals of the Chrysene Formalism
San Francisco, California, USA
editor@chrysene.com

Abstract

The genetic-code, protocell-ledger, and finite interior-lattice constructions of the Chrysene Formalism require a source of substituted molecular units and their organized assemblies. This fourth pilot study places endogenous assembly from a mixture and exogenous engineered preparation in a common finite mathematical framework. A source is represented by its admissible configurations, observable functions, preparation law, causal controls, and operating dynamics. Under full-support and placement hypotheses the two origins have the same reachable configurations, while their attainable distributions and trajectories can differ.

The principal comparison is functional rather than microscopic. For a positive endogenous reference law and a prescribed distribution of functional outputs, we derive the unique least-relative-entropy preparation: it changes the output distribution and preserves the reference conditional law within each functional fiber. Exact microscopic placement incurs an additional conditional information cost whenever the fiber is nontrivial. On a six-site registry column, the thirty-two agreement words give twenty reversal classes; selecting a two-member class retains one bit of variability that exact placement removes. We construct rejection, reversible local-rate, and finite-horizon conditioned realizations and distinguish their information objectives from their material and temporal requirements.

Robust-demand and responsive-kinetic examples give strict advantages over specified fixed-placement classes. Exact and approximate emulation theorems identify the sensing, randomization, local dynamics, and resource conditions by which an engineered source matches or approaches those capabilities. A coefficient-consistent lattice extension separates structural invariants from source-dependent laws. The resulting theory compares endogenous and exogenous organization without making molecular provenance itself an obstruction to engineering.

Keywords: Chrysene Formalism; biogenesis; endogenous source; exogenous source; molecular self-assembly; conditional selection; relative entropy; Markov dynamics; registry; engineered emulation.

1 Introduction

1.1 From encoded components to their source

A mathematical account of a protocell must coordinate the production of its components with the organization that makes those components functional. In the Chrysene Formalism, the preceding pilot studies addressed three successive parts of this construction. The genetic-code study formulated a typed substituted-chrysene carrier and the assignments by which its registry supports a code-table realization [Sch26d]. The protocell-ledger study coordinated the carbon traceability of genetic, membrane, and interior components within a common stock-based construction [Sch26c]. The finite interior chrysene lattice study then specified the geometric, algebraic, and operational structures through which a finite internal lattice can be characterized and coordinated with the surrounding protocell [Sch26a]. The present article is the fourth pilot study in that research sequence. Its subject is the source that supplies and organizes the substituted units on which those downstream constructions depend.

For an endogenous biogenesis scenario, a source cannot be introduced solely as a list of correctly prepared molecules. It must include a mechanism by which an available mixture yields the relevant stock, brings compatible units into contact, and retains the organizations required for subsequent use. A gen-

Citation: Schaper, C. D. (2026). Endogenous and Exogenous Sources of Encoded Chrysene Systems: Functional Selection, Adaptive Realization, and Engineered Emulation. *Annals of the Chrysene Formalism*, 1(4):1905–1927.
https://chrysene.com/index.php/acf/article/view/99
© 2026 Charles D. Schaper. Distributed under the CC Attribution-NonCommercial-NoDerivatives 4.0. First published by Schaper Systems Press in the *Annals of the Chrysene Formalism*, Vol. 1, No. 4 (2026). ISSN 3144-2021.

eral bottom-up mixture description, often summarized by the primordial-soup picture, leaves both the identity of the useful components and their productive encounters to be resolved. The substituted-chrysene program narrows this inventory by working from one molecular core and a specified family of substitutions. Within that hypothesis, the functional components are related transformations and assemblies of a shared stock rather than unrelated molecular starting points.

This common-core restriction makes the mathematical inventory simpler and permits shared structural and material accounting. It does not, by itself, supply a synthesis route, remove competitors, set an orientation, or establish a reaction rate. Those are questions about the source mechanism. The appropriate next step is therefore to place the source inside the same explicit framework as the code, the ledger, and the lattice, so that chemical assumptions and mathematical consequences can be connected without being identified.

1.2 Endogenous and engineered preparation

Article 79, *The Source*, proposes a reaction–transport mechanism initiated by substitution. Modification of a polyaromatic bulk is coupled to selective export and organization of the substituted-chrysene stock [Sch26r]. Its adopted kinetic model gives exit probabilities, demand-matching relations, and supply constraints. Article 84, *The Placed Source*, keeps the admitted material architecture but replaces its preparation by an external agent placing specified units and registry data [Sch26o]. Under explicit positivity and placement hypotheses, it identifies the same reachable configurations for the two sources. Extension from columns to laterally bridged assemblies requires the additional availability of the relevant bridge-forming operations.

The contrast is consequently not exhausted by asking which configuration can exist. A low-probability endogenous outcome and a reliably placed outcome can belong to the same reachable set. Conversely, a precisely placed configuration may discard variability that is harmless to the desired function, or may be less useful when the required output is not known until the environment changes. A comparison of sources must distinguish possibility, distribution, function, persistence, and intervention.

Article 90 already develops a parameterized comparison of source laws and separates chemical selection, registry selection, arm-pattern visibility, and persistence [Sch26q]. The present article takes that law-level description as a starting point and asks a different question: *what must be controlled to obtain a function, what may remain organized by the source itself, and which engineering capabilities recover the resulting performance?*

An exogenous origin does not imply permanently open-loop operation. Advanced molecular manufacturing could install a reference, a local selector, an environmental sensor, a reversible assembly pathway, or an entire autonomous organization. A system engineered externally may subsequently operate through the same local transitions as an endogenous system. We therefore distinguish the origin of preparation from the class of operations available after preparation. Fixed placement, randomized placement, adaptive control, and engineered autonomous assembly are different comparison classes.

1.3 Functional fibers as the organizing object

Let $\mathcal{X}$ be a finite space of admissible molecular configurations and let $\pi : \mathcal{X} \to \mathcal{Y}$ report the function of interest. The fiber $\pi^{-1}(y)$ comprises configurations equivalent for that function. Equivalence may mean a common registry class, a common specified product, or the same chosen observation. These are not interchangeable interpretations; each requires its own map.

The principal result determines the preparation law that changes a positive reference law μ as little as possible in relative entropy while imposing a desired output law. The solution retains μ's conditional distribution inside each functional fiber. Thus the target is selected without introducing distinctions that the function does not require. Exact placement of one member of a nontrivial fiber imposes an additional conditional specification.

The underlying entropy decomposition is standard. Its role here is to provide a common comparison principle for the source, the placed source, the registry quotient, and their engineered realizations, rather than to assert a new abstract entropy identity. Its physical interpretation depends on which mechanism realizes the law and which resources are charged. The paper therefore constructs explicit realizations and records their distinct requirements instead of equating information change with total manufacturing expenditure.

1.4 Results and research direction

The development has four connected parts.

(i) A common operational model distinguishes structural reachability from attainable observation laws and causal response. Emulation is quantified uniformly over the allowed environments.

(ii) Finite disintegration gives least-information functional selection, its strict comparison with over-specified microscopic placement, and a complete six-site registry example. Robust demand exposes the difference between diversity over words and diversity over functions.

(iii) Local rate tilting and finite-horizon conditioning construct mechanisms for the selected law. A two-state responsive model exhibits a strict finite-time advantage over uninformed preparation, while an engineered copy of the mechanism removes that gap.

(iv) Exact and approximate emulation bounds turn engineering approximation into a quantitative research problem.

Lattice cochains, finite codes, and common observable structures enter through explicit maps rather than through a presumed equivalence of different mathematical domains.

These results show how an endogenous mechanism can attain an optimality property unavailable to a specified restricted placed-source protocol. They also show how engineering can approach that capability: preserve the relevant functional conditional law, implement the required local transitions, provide the observations that adaptation uses, and account for their resources on the same basis. This is a comparison of realizations, not a universal ranking of natural and engineered origins.

1.5 Scope and conventions

All central state spaces and horizons are finite. Logarithms in relative entropy are natural, with bits obtained by division by $\log 2$. The source papers provide the molecular vocabulary and motivate the mechanisms; the mathematical constructions and their hypotheses are stated explicitly below. The fixed occupied set and the allowed transition graph are part of a realization. Where formation of the occupied set itself is the research question, vacant-site and attachment states must be added rather than assuming the completed geometry as an outcome.

The finite-ICL pilot is now Article 88 [Sch26a]. Earlier working manuscripts retained in the bibliography are identified by title and historical numbering where their current publication mapping is unconfirmed. The present manuscript is Article 99, formerly Article 102; it is distinct from the earlier arithmetic-geometry working manuscript cited below.

2 Finite molecular objects and their observations

The source comparison begins with a common object, rather than with two different descriptions of the same material. The stock alphabet of Article 79 and the placement configurations of Article 84 supply the finite-column model. The directional connector analysis of Article 94 specifies the additional compatibility data required for a laterally connected realization [Sch26r, Sch26o, Sch26s].

Definition 2.1 (Molecular configuration datum)**.** Fix a finite occupied set $\Omega \subset \mathbb{Z}^3$, finite sets of candidate lateral edges E_{lat} and stacking contacts E_z, a finite arm alphabet $\mathcal{H}$, and a finite reference set $\mathcal{R}$. A molecular configuration datum is a tuple

$$x = (\kappa, h, \epsilon, \eta, B, a).$$

Here κ assigns an admitted substitution family to each vertex; $h_v : \kappa(v) \to \mathcal{H}$ is its arm word; ϵ belongs to $\mathbb{F}_2^{\Omega}$ when the cipher coordinate is independently realized, and to a singleton otherwise; $\eta \in \mathbb{F}_2^{E_z}$ records the stacking-contact registry; $B \subseteq E_{\mathrm{lat}}$ is the installed connector set; and $a \in \mathcal{R}$ records the reference configuration together with the vertex placement frames needed to evaluate locant directions. Where these frames are decoded from other coordinates, $\mathcal{R}$ records that fixed decoding rather than an extra independent choice. The predicate below enforces agreement between the frames and any vertex-lifted registry. A coordinate which is fixed by the realization is retained as a constant, not counted as a freely selectable degree of freedom.

The primary locant families are

$$D = \{6, 12\}, \qquad I = \{3, 6, 12\}, \qquad T = \{3, 6, 9, 12\}.$$

Whether the inversion image of I is also admitted is part of the chosen alphabet. For every installed edge $b = \{v, w\}$, compatibility requires available endpoint locants at both v and w, correctly oriented for the edge direction, with compatible arm types. Distinct installed connectors may not use the same consumable terminal. The endpoint assignment is evaluated in the placement frame of the unit. Thus a scalar inequality comparing vertex degree with arm count is necessary in some realizations but is not a substitute for directional compatibility.

Definition 2.2 (Admissible configuration space)**.** Let $\mathsf{A}(x)$ be a specified predicate expressing local constitutional admission, endpoint compatibility, terminal allocation, registry constraints, and reference requirements. The admissible configuration space is the nonempty finite set

$$\mathcal{X} = \{x : \mathsf{A}(x)\}.$$

The predicate and the occupied set are held fixed in a comparison of source mechanisms.

Finiteness follows directly from the finite domains in Definition 2.1. No assertion that every assignment satisfies A is needed. The distinction between the stock and the interior lattice is included in this definition. For example, a stock model may admit D, I, T families, whereas a particular saturated interior-lattice model fixes its nodes to the specified tetrasubstituted constitutional type. The former alphabet must not be substituted into the latter model without changing A.

Definition 2.3 (Registry conventions)**.** On a column with ℓ vertices, the binary agreement field has $\ell - 1$ contact coordinates. More generally, on a chosen graph $G = (V, E)$, a

vertex-lifted binary registry has the form

$$\eta_{vw} = s_v + s_w, \qquad s : V \longrightarrow \mathbb{F}_2.$$

If contact values are instead independent primitive data, this is declared explicitly and the vertex-lift condition is not imposed.

Lemma 2.4 (Integrable binary registry)**.** *Let G be a finite graph with c connected components. A binary edge assignment η has a vertex lift if and only if its sum around every cycle is zero. The space of lifted edge assignments has dimension $|V| - c$ over $\mathbb{F}_2$.*

Proof. If $\eta_{vw} = s_v + s_w$, summing around a cycle counts every vertex label twice and gives zero. Conversely, choose a root in each component, set its label to zero, and define s_v as the sum of η along a path from the root to v. The cycle condition makes this value independent of the path. It then satisfies the required edge identities. The kernel of the map $s \mapsto \eta$ consists exactly of the assignments constant on each component. Rank–nullity gives the dimension. □

The lemma prevents two incompatible counts from being combined: arbitrary edge labels and integrable edge labels have the same cardinality on a forest, but generally not on a graph with cycles. All coefficient-dependent statements in this paper retain their specified coefficient field.

Definition 2.5 (Functional observation)**.** A functional observation is a map $\pi : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{Y}$ is replaced by $\pi(\mathcal{X})$ if necessary. A frame quotient $q : \mathcal{X} \to \mathcal{X}/G_{\mathrm{fr}}$ is a separate observation. Factorization $\pi = \bar{\pi} \circ q$ is asserted only when π is constant on every relevant frame orbit.

The product-factorization statement of Article 84 is used on its fixed-composition cluster locus. It does not identify configurations with different atom inventories merely because their registry fields agree. Likewise, an asymmetric cap supplies a reference datum only under the declared physical interpretation of that reference. These typing requirements make the comparison independent of accidental choices of coordinates.

3 Preparation laws and common reachability

For a finite set F, write $\mathcal{P}(F)$ for its probability distributions. A preparation law $\mu \in \mathcal{P}(\mathcal{X})$ specifies how often the source produces each admitted configuration. A placement map $P : \mathcal{U}_{\mathrm{place}} \to \mathcal{X}$ specifies the configurations that a placement protocol can prepare.

Definition 3.1 (Reachable sets)**.** The endogenous and placed reachable sets are, respectively,

$$R_{\mathrm{endo}} = \operatorname{supp}\mu, \qquad R_{\mathrm{place}} = P(\mathcal{U}_{\mathrm{place}}).$$

Their functional images are $\pi(R_{\mathrm{endo}})$ and $\pi(R_{\mathrm{place}})$.

Theorem 3.2 (Common reachability under full access)**.** *Suppose $\mu(x) > 0$ for every $x \in \mathcal{X}$, and suppose P is surjective. Then*

$$R_{\mathrm{endo}} = R_{\mathrm{place}} = \mathcal{X},$$

and the two functional images equal $\pi(\mathcal{X})$ for every common functional observation π.

Proof. Strict positivity gives $\operatorname{supp}\mu = \mathcal{X}$. Surjectivity gives $P(\mathcal{U}_{\mathrm{place}}) = \mathcal{X}$. Applying π to the equal sets proves the second assertion. □

This is the finite form of the reachability comparison of Article 84 [Sch26o]. In its column realization, the positive substitution and export rates from Article 79 provide positive weights for the admitted substitution degrees; the assembly law must also allow every remaining admitted arm and structural datum. For a laterally connected occupied set, equal access to its connector construction is an additional premise. Fixing Ω does not prove that either mechanism forms that geometry from an unrestricted mixture.

Lemma 3.3 (Sequential full support)**.** *Encode a configuration as a finite sequence of choices. If every prefix that can be extended to an admitted configuration has positive probability, and each next choice belonging to such an extension has positive conditional probability, the terminal law has full support on $\mathcal{X}$.*

Proof. Fix $x \in \mathcal{X}$ and follow its sequence of choices. Its probability is the product of the initial factor and finitely many strictly positive conditional factors. Hence it is positive. □

Positive one-coordinate marginals do not suffice for this conclusion. For instance, the law giving probability $1/2$ to each of 00 and 11 has positive marginals on both binary coordinates but excludes 01 and 10. Full support concerns compatible joint configurations.

Proposition 3.4 (Terminal emulation by randomized placement)**.** *If placement is surjective onto finite $\mathcal{X}$ and the placement controller can sample an arbitrary prescribed*

distribution on X, it can reproduce every preparation law $\mu \in \mathcal{P}(X)$.

Proof. For each $x \in X$, choose one placement instruction u_x with $P(u_x) = x$. Sample x with probability $\mu(x)$ and execute u_x. The resulting law is $\sum_x \mu(x)\delta_x = \mu$. □

This proposition concerns terminal distributions. It does not assert that the instructions have the same cost, nor that terminal placement reproduces an endogenous trajectory. It identifies the capability that must be excluded if a terminal distribution is to be strictly unavailable to the placed comparison class.

4 Causal protocols and operational equivalence

The parameterized laws of Article 90 distinguish chemical selection, registry selection, and persistence [Sch26q]. To compare operating capabilities as well as initial preparation, these laws are extended to finite histories.

Definition 4.1 (Admissible source protocol)**.** Fix a horizon $T \in \mathbb{N}$, a nonempty environment set $\mathcal{E}$, and finite state, observation, and action alphabets. A protocol consists of an initial preparation rule, transition rules, and a causal policy. At time t the policy may depend on its past actions and the observations available through time t, but not on an unobserved environment variable or a future observation.
For each environment $e \in \mathcal{E}$, the protocol induces a law $P_{u,e}$ on a common finite comparison record $Z_0, \dots, Z_T \in \mathcal{Z}$. The record includes the functional observations and any resource data used in the comparison. Different internal implementations may be mapped to this same record. The record has a declared finite observation resolution. Exact continuous waiting times or full continuous-time jump trajectories require a separate measurable-space model; any sampling or discretization used here is part of the observation map.

An environment can encode a fixed condition or an externally specified input sequence. Its influence through reaction rates or sensors is part of the transition and observation rules. A policy may use that influence only through its allowed observations. In particular, the index e in $P_{u,e}$ is not an instruction allowing the controller to know e.

Every finite path law admits a sequential factorization

$$P_{u,e}(z_{0:T}) = \nu_{u,e}(z_0) \prod_{t=0}^{T-1} K_{u,e,t}(z_{t+1} \mid z_{0:t}).$$

Conditional kernels on zero-probability histories may be extended arbitrarily. Approximation hypotheses below concern specified extensions on all histories; this is a sufficient condition that is straightforward to test.

Definition 4.2 (Protocol classes)**.** Let C_{endo} be a declared class of endogenous reaction–assembly protocols. Distinguish the following placed classes: fixed deterministic preparation, preparation with an admissible randomizer, observation-dependent control, and engineered initialization followed by autonomous molecular operation. Their available observations, transitions, and resource bounds form part of their definitions.

These classes classify capabilities rather than technological labels. An engineered system that contains the same reaction, retention, sensing, and repair mechanisms as an endogenous system belongs to an appropriate autonomous placed class.

Definition 4.3 (Operational emulation)**.** For two protocol classes C and $\mathcal{D}$, write $C \preceq \mathcal{D}$ when for every $u \in C$ there exists a single $v \in \mathcal{D}$ such that

$$P_{u,e} = P_{v,e} \qquad \text{for every } e \in \mathcal{E},$$

and v satisfies the required resource constraints whenever u does. Thus the order of quantifiers is $\forall u\, \exists v\, \forall e$.

The definition excludes a different emulator selected after learning an environment that the original protocol was not permitted to observe.

Theorem 4.4 (Emulation transfers performance)**.** *If $C \preceq \mathcal{D}$, every attainable expectation of an environment-dependent functional of the comparison record under a protocol in C is matched under a protocol in $\mathcal{D}$, simultaneously for all environments. If emulation holds in both directions, the two classes have the same attainable environment-indexed path-law families.*

Proof. Choose the emulator in Definition 4.3. For every e and every function f_e on the finite record space,

$$\sum_{z_{0:T}} f_e(z_{0:T}) P_{u,e}(z_{0:T}) = \sum_{z_{0:T}} f_e(z_{0:T}) P_{v,e}(z_{0:T}).$$

The second assertion follows by applying this matching of laws in both directions. □

Corollary 4.5 (No provenance-only separation)**.** *If the admissible placed class operationally emulates the endogenous class, no objective determined by the comparison*

record can have a strict optimum attainable only by the endogenous class.

Proof. Any proposed attaining endogenous protocol has an admissible placed emulator with the same objective value. □

The corollary does not prevent strict comparison results. It locates their content: a strict result must use a specified restriction in the placed class, such as unavailable feedback, a limited randomizer, a different transition set, or a binding resource constraint. Equal material reachability alone neither supplies nor excludes these operating capabilities.

4.1 Approximate emulation

For distributions p, q on a finite set F, use

$$d_{\mathrm{TV}}(p,q) = \frac{1}{2}\sum_{x\in F} |p(x) - q(x)|.$$

This convention gives $|\mathbb{E}_p f - \mathbb{E}_q f| \le M d_{\mathrm{TV}}(p,q)$ for $0 \le f \le M$. Indeed, writing $r = p - q$ and using $\sum r = 0$, the largest value of $\sum_x f(x)r(x)$ is obtained by putting $f = M$ on the positive part of r and $f = 0$ on the negative part.

Theorem 4.6 (Finite-horizon engineering error)**.** *Let two admissible protocols have initial laws $\nu_e, \widehat{\nu}_e$ and history kernels $K_{e,t}, \widehat{K}_{e,t}$ on the same comparison alphabet. Suppose, uniformly over environments and histories,*

$$d_{\mathrm{TV}}(\nu_e, \widehat{\nu}_e) \le \delta_0,$$
$$d_{\mathrm{TV}}(K_{e,t}(\cdot \mid h), \widehat{K}_{e,t}(\cdot \mid h)) \le \delta_{t+1}.$$

Then the laws of their complete records satisfy

$$d_{\mathrm{TV}}(P_e, \widehat{P}_e) \le \Delta_T := \min\left\{1, \sum_{j=0}^{T} \delta_j\right\}.$$

The same bound holds after applying any common deterministic observation map or stochastic observation channel to the record.

Proof. For distributions p, q on histories and extension kernels $K, \widehat{K}$,

$$\begin{aligned} d_{\mathrm{TV}}(pK, q\widehat{K}) &\le d_{\mathrm{TV}}(pK, qK) + d_{\mathrm{TV}}(qK, q\widehat{K}) \\ &\le d_{\mathrm{TV}}(p,q) \\ &\quad + \sup_h d_{\mathrm{TV}}(K(\cdot \mid h), \widehat{K}(\cdot \mid h)). \end{aligned}$$

The first term contracts because a stochastic kernel preserves total mass and the triangle inequality can be applied after summing over its output. The second is bounded by the average conditional distance, and hence by its supremum. Apply this inequality successively to the path prefixes. A probability distance is at most one, giving the truncation. The final assertion is the same contraction inequality for the common observation channel. □

In particular, an initial error δ_0 and a per-step kernel error δ give $\Delta_T \le \min\{1, \delta_0 + T\delta\}$. The estimate is for the entire observed trajectory, not merely its terminal state. It does not imply convergence uniformly over unbounded time horizons.

Definition 4.7 (Resource-bounded robust value)**.** Let $0 \le f_e \le M$ be a common reward on the record and let $0 \le C \le M_C$ be a common resource cost on that record. Define

$$c(u) = \sup_{e\in\mathcal{E}} \mathbb{E}_{P_{u,e}} C.$$

For a protocol class C with a nonempty feasible set at budget b, define

$$V_C(b,T) = \sup_{\substack{u\in C \\ c(u)\le b}} \inf_{e\in\mathcal{E}} \mathbb{E}_{P_{u,e}} f_e.$$

The cost convention is a worst-environment expected budget. A pathwise budget is a different constraint and is not substituted for it.

Theorem 4.8 (Performance and budget transfer)**.** *If each $u \in C$ feasible at b has a single emulator $v \in \mathcal{D}$ whose record-law error is at most Δ uniformly over environments, then*

$$V_{\mathcal{D}}(b + M_C\Delta, T) \ge V_C(b,T) - M\Delta.$$

If emulation is exact and resource preserving, the same-budget inequality $V_{\mathcal{D}}(b,T) \ge V_C(b,T)$ holds.

Proof. The expectation bound for total variation gives, for every environment,

$$\mathbb{E}_{P_{v,e}} f_e \ge \mathbb{E}_{P_{u,e}} f_e - M\Delta,$$
$$\mathbb{E}_{P_{v,e}} C \le \mathbb{E}_{P_{u,e}} C + M_C\Delta.$$

Taking the supremum over environments in the cost inequality makes v feasible at $b + M_C\Delta$. Taking the infimum in the reward inequality and then the supremum over feasible u proves the assertion. No maximizing protocol is presumed to exist. The exact case has zero error and no budget increase. □

Corollary 4.9 (An engineering approximation sequence)**.** *Fix a horizon. Suppose admissible engineered protocols approximate an endogenous protocol with uniform initial*

errors tending to zero and uniform per-step kernel errors tending to zero. Then their comparison-record laws converge in total variation, their bounded reward values converge, and their bounded expected resource costs converge, uniformly over environments.

Proof. There are finitely many steps. The bound in Theorem 4.6 therefore tends to zero. Apply the expectation bounds to the reward and cost. □

This is a precise sense in which engineering can approach the capability of an endogenous source. The engineering task is to approximate a preparation-and-response mechanism at a declared resolution, rather than only to manufacture one endpoint configuration. The estimate charges imperfect initialization, imperfect responses, and finite observation horizon separately. Where energy, feedstock, delay, and external control cannot be combined into a justified scalar cost, Definition 4.7 is applied componentwise and the resulting feasible performance sets are compared by their Pareto order.

5 A strict adaptive separation and its recovery by engineering

A small kinetic realization shows how equal molecular reachability can coexist with strict operational differences. The two states in this section represent two admitted functional classes of a fixed architecture. Their rates are explicitly assumed molecular response parameters; the example does not infer such rates from the chrysene core alone.

Definition 5.1 (Binary demand experiment)**.** The environment is $e \in \{0,1\}$ and the terminal output is $Y \in \{0,1\}$. Success is the event $Y = e$. The environment is unknown to an uninformed preparation policy and is fixed during the preparation or response interval.

Proposition 5.2 (Uninformed preparation bound)**.** *An environment-independent randomized preparation law has worst-environment success probability at most* $1/2$*. The unique maximizing law is uniform. Every fixed deterministic preparation has worst-environment success zero.*

Proof. Let $a = \Pr(Y = 1)$. The two success probabilities are $1 - a$ and a, so their minimum is at most $1/2$, with equality exactly when $a = 1/2$. A deterministic preparation has $a = 0$ or $a = 1$. □

Definition 5.3 (Environment-responsive kinetic witness)**.** Fix rates $\alpha > \beta > 0$. In environment e, the transition rate from $1 - e$ to e is α and the reverse rate is β. In the ordered states $0, 1$, the generators are

$$Q_0 = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix}, \qquad Q_1 = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}.$$

The coupling of environment to the generator is part of the realization. A molecular mechanism with these rates is responsive even if it has no separately identified digital sensor.

Theorem 5.4 (Finite-time endogenous response advantage)**.** *For either environment, if* p_0 *is the initial probability of the correct state, the correct-state probability at time* t *is*

$$p_e(t) = \frac{\alpha}{\alpha+\beta} + \left(p_0 - \frac{\alpha}{\alpha+\beta}\right) e^{-(\alpha+\beta)t}.$$

Consequently, for every initial law, the worst-environment success is strictly greater than $1/2$ *whenever*

$$t > \frac{1}{\alpha+\beta} \log\left(\frac{2\alpha}{\alpha-\beta}\right).$$

Its stationary success probability is $\alpha/(\alpha+\beta) > 1/2$*.*

Proof. The forward equation for the probability of the correct state is $p_e' = \alpha(1 - p_e) - \beta p_e$. Solving this scalar linear equation gives the formula. It is increasing in p_0, so a uniform lower bound over all initial laws is

$$p_e(t) \geq \frac{\alpha}{\alpha+\beta}\left(1 - e^{-(\alpha+\beta)t}\right).$$

This exceeds $1/2$ exactly when $e^{-(\alpha+\beta)t} < (\alpha-\beta)/(2\alpha)$, giving the stated sufficient dwell time. Letting t tend to infinity gives the stationary value. □

The rates are positive, so both states are accessible in both environments. The strict difference is therefore not a difference of reachable sets. It is a difference between an environment-blind preparation law and a mechanism whose transition law is coupled to the environment. A cost-qualified comparison additionally requires that the responsive realization satisfy the same declared budget; the theorem by itself is a success-probability statement.

Proposition 5.5 (Recovery by an engineered response mechanism)**.** *An engineered realization having the same initial law, environment coupling, and generators* Q_0, Q_1 *has exactly the same response law. At sampling interval* Δt*,*

the common discrete kernels are $\exp(\Delta t\, Q_e)$. *Engineered approximations to these kernels satisfy the finite-horizon estimate of Theorem 4.6.*

Proof. For a finite generator Q_e, the transition matrix is the unique solution of the linear initial-value problem $K'(t) = K(t)Q_e$, $K(0) = I$, namely $\exp(tQ_e)$. Equal initial laws and equal transition matrices give equal sampled path laws by their sequential factorization. The approximation assertion is Theorem 4.6. □

The example makes the progression from placement to source engineering explicit. Exact initial placement controls an endpoint. Randomized placement additionally controls an ensemble. An engineered response mechanism controls how that ensemble changes with the environment. Once the latter mechanism reproduces the endogenous kernels and respects the same resource constraints, the operational distinction vanishes even though the historical origins of the systems remain different.

6 Functional selection and retained source variability

The configuration selected by a source and the function obtained from that configuration need not have the same resolution. Article 84 compares reachable configurations and their product images, while Article 90 separates chemical and registry contributions to configuration surprisal [Sch26o, Sch26q]. We now fix the functional requirement first and determine the least change to the source law needed to satisfy it. The resulting construction retains the source's conditional variability wherever the required function permits it. Its mathematical basis is the finite relative-entropy chain rule; the source comparison and the registry instantiation are the applications developed here. Relative-entropy projection provides the standard information-geometric setting [Csi75, CT06].

6.1 Disintegration over a functional map

Definition 6.1 (Reference source and functional fibers)**.** Let $\mathcal{X}$ and $\mathcal{Y}$ be finite nonempty sets, let $\pi : \mathcal{X} \to \mathcal{Y}$ be surjective, and let $\mu \in \mathcal{P}(\mathcal{X})$ satisfy $\mu(x) > 0$ for every $x \in \mathcal{X}$. Write

$$\mathcal{X}_y = \pi^{-1}(\{y\}), \qquad \bar{\mu}(y) = \sum_{x \in \mathcal{X}_y} \mu(x).$$

The conditional reference law on $\mathcal{X}_y$ is $\mu_y(x) = \mu(x)/\bar{\mu}(y)$. All logarithms in this section are natural. For probability laws p, q on a finite set, with $q > 0$, set

$$D_{\mathrm{KL}}(p\|q) = \sum_{x:p(x)>0} p(x) \log \frac{p(x)}{q(x)}.$$

Thus zero summands are assigned their continuous value 0.

The law μ may be the assembly law induced by a specified endogenous reaction–transport and retention model. Its positivity is an explicit assumption about that model. On a constrained lattice, $\mathcal{X}$ is the admissible set already incorporating connector and registry compatibility, not an unconstrained product of site and edge alphabets. The map π may record a registry class or a chemical function; the two choices coincide only when the corresponding identification has been established.

Lemma 6.2 (Finite disintegration)**.** *Let $p \in \mathcal{P}(\mathcal{X})$ and $r = \pi_* p$. For every y with $r(y) > 0$ there is a unique probability law p_y on $\mathcal{X}_y$ such that*

$$p(x) = r(y)p_y(x), \qquad x \in \mathcal{X}_y.$$

If $r(y) = 0$, then p vanishes throughout $\mathcal{X}_y$, and the choice of p_y has no effect on p. In particular, for every prescribed $r \in \mathcal{P}(\mathcal{Y})$, the law

$$p_r^{\mu}(x) = r(\pi(x))\mu_{\pi(x)}(x) \tag{1}$$

is well defined and satisfies $\pi_* p_r^{\mu} = r$.

Proof. For $r(y) > 0$, divide $p(x)$ by its sum over $\mathcal{X}_y$. For $r(y) = 0$, nonnegativity forces every summand of that sum to vanish. Finally, summing (1) over $\mathcal{X}_y$ gives $r(y)$, since μ_y is a probability law. Summing once more over y gives total mass one. □

Lemma 6.3 (Nonnegativity and its equality condition)**.** *If $q > 0$, then $D_{\mathrm{KL}}(p\|q) \geq 0$, with equality if and only if $p = q$.*

Proof. Apply $\log t \leq t - 1$, with equality only at $t = 1$, to $t = q(x)/p(x)$ for $p(x) > 0$. It follows that

$$-D_{\mathrm{KL}}(p\|q) \leq \sum_{x:p(x)>0} q(x) - 1 \leq 0.$$

Equality requires $q(x) = p(x)$ on the support of p and zero q-mass outside it. These conditions are equivalent to $p = q$. □

Theorem 6.4 (Least-information functional selection)**.** *For any prescribed functional law $r \in \mathcal{P}(\mathcal{Y})$ and any $p \in \mathcal{P}(\mathcal{X})$ with $\pi_* p = r$,*

$$D_{\mathrm{KL}}(p\|\mu) = D_{\mathrm{KL}}(r\|\bar{\mu}) + \sum_{y:r(y)>0} r(y) D_{\mathrm{KL}}(p_y\|\mu_y). \tag{2}$$

Consequently,

$$\min_{\pi_* p = r} D_{\mathrm{KL}}(p\|\mu) = D_{\mathrm{KL}}(r\|\bar{\mu}), \tag{3}$$

and the unique minimizing law is p_r^μ of (1).

Proof. For $p(x) > 0$, put $y = \pi(x)$. Then $r(y) > 0$ and

$$\frac{p(x)}{\mu(x)} = \frac{r(y)}{\bar{\mu}(y)} \frac{p_y(x)}{\mu_y(x)}.$$

Take logarithms, multiply by $p(x)$, and sum first over each fiber and then over y. The fiber sum of p_y is one, giving (2). Lemma 6.3 proves the lower bound. For $p = p_r^\mu$, every positive-weight conditional law equals μ_y, so the bound is attained. Conversely, equality in the bound forces $p_y = \mu_y$ whenever $r(y) > 0$. On the remaining fibers $p = 0$ by Lemma 6.2. These conditions determine p_r^μ uniquely. □

Corollary 6.5 (Exact residual penalty)**.** *For every p with $\pi_* p = r$,*

$$D_{\mathrm{KL}}(p\|\mu) = D_{\mathrm{KL}}(r\|\bar{\mu}) + D_{\mathrm{KL}}(p\|p_r^\mu). \tag{4}$$

The last divergence is taken on the common allowed set $\bigcup_{r(y)>0} \mathcal{X}_y$, where p_r^μ is strictly positive.

Proof. On a positive-weight fiber, $p(x)/p_r^\mu(x) = p_y(x)/\mu_y(x)$. Summation gives the conditional term of (2). Both laws vanish outside the displayed allowed set. □

Equation (4) separates two tasks. The first term changes the distribution of required functions. The second term further reorganizes the configurations implementing those functions. When the functional specification leaves internal choices open, the latter change is avoidable for this objective. This is the precise sense in which source variability can be useful rather than a defect.

6.2 Exact configurations and functional classes

Corollary 6.6 (Class selection and microscopic specification)**.** *Fix $y \in \mathcal{Y}$. The unique least-divergence law supported on $\mathcal{X}_y$ is the conditional law μ_y, extended by zero outside $\mathcal{X}_y$, and*

$$D_{\mathrm{KL}}(\mu_y\|\mu) = -\log \bar{\mu}(y). \tag{5}$$

For every $x \in \mathcal{X}_y$,

$$D_{\mathrm{KL}}(\delta_x\|\mu) - D_{\mathrm{KL}}(\mu_y\|\mu) = -\log \mu_y(x). \tag{6}$$

The difference is positive if and only if $\mu_y(x) < 1$.

Proof. Apply Theorem 6.4 to $r = \delta_y$. The point law has divergence $-\log \mu(x)$, and $\mu(x) = \bar{\mu}(y)\mu_y(x)$ proves the difference formula. □

Proposition 6.7 (Strict source-class separation for the information objective)**.** *Suppose an admissible endogenous selection mechanism realizes μ_y, and suppose the comparison class consists of deterministic placements δ_x with $x \in \mathcal{X}_y$. If $|\mathcal{X}_y| \geq 2$, the least divergence in that deterministic class exceeds the endogenous value by*

$$-\log\left(\max_{x\in\mathcal{X}_y} \mu_y(x)\right) > 0. \tag{7}$$

If the exogenous class also realizes μ_y, the strict separation vanishes.

Proof. Minimize (6) over the finite nonempty fiber. Strict positivity of μ and $|\mathcal{X}_y| \geq 2$ imply that its largest conditional atom is less than one. If μ_y belongs to the exogenous class, it attains the common global lower bound (5). □

This proposition compares two preparations yielding the same functional output with probability one. It does not assign a microscopic-precision task to one source and a different functional task to the other: exact placement is here a restriction on the available protocol, not a requirement of the task. Expanding the placed-source protocol to include the minimizing mixture removes the gap.

Remark 6.8 (Information, material and work)**.** The divergence is measured in nats; division by $\log 2$ gives bits. It is a change-of-law objective, not automatically the number of laboratory control operations, the amount of feedstock consumed, or a work requirement. A thermodynamic identification requires a physical reference equilibrium state, a specified Hamiltonian, reservoirs, and a preparation proto-

col. The class-state entropy comparison in Article 96 motivates the present formulation [Sch26j]; the finite proof above uses no quantum-channel or work-cost assumption.

6.3 Uniform reference and entropy

Proposition 6.9 (Retained entropy at fixed function)**.** *Let μ be uniform on $\mathcal{X}$, put $n = |\mathcal{X}|$ and $n_y = |\mathcal{X}_y|$, and define $H(p) = -\sum_{p(x)>0} p(x) \log p(x)$. Among laws with $\pi_* p = r$,*

$$H(p) \le H(r) + \sum_y r(y) \log n_y. \tag{8}$$

Equality holds exactly for $p = p_r^\mu$, which is uniform within each positive-weight fiber.

Proof. For a uniform reference, direct expansion gives $D_{\mathrm{KL}}(p\|\mu) = \log n - H(p)$. Theorem 6.4 therefore identifies the unique entropy maximizer. Its value follows from $p_r^\mu(x) = r(y)/n_y$ on $\mathcal{X}_y$: summing $-p_r^\mu(x) \log p_r^\mu(x)$ over the fibers yields the right-hand side of (8). □

For a nonuniform reference, the minimizing conditional law is μ_y, which need not maximize Shannon entropy on its fiber. The general principle is preservation of the reference's conditional organization; maximal conditional entropy is its uniform-reference specialization.

Proposition 6.10 (Successive functional quotients)**.** *Let $\psi : \mathcal{Y} \to \mathcal{Z}$ be a surjection, $\widetilde{\mu} = \psi_* \bar{\mu}$, and $s = \psi_* r$. Then*

$$r_z = r(\cdot \mid z), \qquad \bar{\mu}_z = \bar{\mu}(\cdot \mid z)$$

are defined on $\psi^{-1}(\{z\})$ whenever $s(z) > 0$, and

$$D_{\mathrm{KL}}(r\|\bar{\mu}) = D_{\mathrm{KL}}(s\|\widetilde{\mu}) + \sum_{z:s(z)>0} s(z) D_{\mathrm{KL}}(r_z\|\bar{\mu}_z). \tag{9}$$

Thus prescribing a finer output distribution cannot have smaller minimum divergence than prescribing only its coarser image.

Proof. Apply Theorem 6.4 with configuration space $\mathcal{Y}$, reference $\bar{\mu}$, and functional map ψ. The reference is strictly positive because π is surjective and $\mu > 0$. Nonnegativity proves the final assertion. □

This is useful when registry classes are subsequently mapped to chemical products or when several encoded states implement the same response. A count of registry classes alone supplies neither the injectivity of that product map nor the cost of selecting a product; the relevant fibers are those of the map actually used.

7 Constructive selection and its supply cost

The minimizing law is an explicit mathematical object. To connect it to a source, the mechanism must be able to distinguish the functional classes on which it acts. The following finite construction isolates that requirement and makes the rejection cost visible.

Definition 7.1 (Functional acceptance mechanism)**.** An independent-trial functional acceptance mechanism draws successive configurations with law μ and accepts a draw x with probability $a(\pi(x))$, where $a : \mathcal{Y} \to [0,1]$. The trials and their acceptance randomness are independent. The first accepted configuration is the output.

Theorem 7.2 (Exact acceptance realization)**.** *For $r \in \mathcal{P}(\mathcal{Y})$, set*

$$c_{\max} = \min_{y:r(y)>0} \frac{\bar{\mu}(y)}{r(y)}.$$

Then $0 < c_{\max} \le 1$. For any $0 < c \le c_{\max}$, define

$$a(y) = \begin{cases} c\, r(y)/\bar{\mu}(y), & r(y) > 0, \\ 0, & r(y) = 0. \end{cases} \tag{10}$$

The mechanism of Definition 7.1 terminates almost surely and outputs p_r^μ. Its number N of draws obeys

$$\Pr(N > k) = (1-c)^k, \qquad \mathbb{E}[N] = c^{-1}. \tag{11}$$

Among class-only acceptance mechanisms producing p_r^μ from independent μ-draws, $c_{\max}$ is the largest acceptance probability per draw.

Proof. Positivity and finiteness make $c_{\max} > 0$. If every positive-target class had $\bar{\mu}(y) > r(y)$, summing over those classes would give a reference mass greater than one. Hence $c_{\max} \le 1$. The bound on c makes (10) an acceptance probability. Its unconditional acceptance probability is

$$\sum_y \bar{\mu}(y) a(y) = c \sum_y r(y) = c.$$

The joint probability of drawing x and accepting equals $c\, r(\pi(x))\mu(x)/\bar{\mu}(\pi(x)) = c\, p_r^\mu(x)$. Thus the conditional law of any accepted draw is p_r^μ. Independence gives the geometric tail and mean in (11), as well as almost-sure termination. Conversely, suppose a class-only mechanism has positive ac-

ceptance probability c' and accepted law p_r^{μ}. Comparing its accepted class probabilities gives $\bar{\mu}(y)a(y) = c'r(y)$. Since $a(y) \le 1$, one has $c' \le \bar{\mu}(y)/r(y)$ for every $r(y) > 0$. Therefore $c' \le c_{\max}$. □

Corollary 7.3 (A single functional target)**.** *For $r = \delta_y$, accept exactly the draws in $\mathcal{X}_y$. The output law is μ_y and the mean number of draws is $1/\bar{\mu}(y)$. The probability of obtaining an accepted configuration within k draws is $1 - (1 - \bar{\mu}(y))^k$.*

Proof. Here $c_{\max} = \bar{\mu}(y)$, so (10) is the indicator of the target class. Apply Theorem 7.2. □

The construction can represent a passive retention interface only if the specified acceptance law is physically implemented; it can equally represent an engineered sorter with the required readout. The source chemistry of Article 79 supplies a reference law and selective export mechanisms [Sch26r], but does not by itself supply every functional acceptance map in (10). The theorem therefore provides an explicit realization target for both endogenous and exogenous research. Its information optimum does not eliminate rejected material, waiting time, sensing, or the means of returning unaccepted material to the supply.

8 A hexamer comparison on the registry quotient

We now instantiate the comparison at the coding-stack scale used in Articles 84, 90 and 91 [Sch26o, Sch26q, Sch26e]. Fix the admissible composition and a reference making the registry field meaningful. The six-site orientation word is $s = (s_1, \dots, s_6) \in \mathbb{F}_2^6$. Its agreement word is

$$d(s) = (s_1 + s_2, \dots, s_5 + s_6) \in \mathbb{F}_2^5. \tag{12}$$

All additions in this display are in $\mathbb{F}_2$. This temporary six-site lift is distinct from the five-coordinate configuration space used below.

Lemma 8.1 (Agreement reduction)**.** *The map d is onto, and each fiber consists of exactly two words related by global complementation. Reversal of a six-site word induces reversal of its five-coordinate agreement word.*

Proof. For a prescribed agreement word η, choose s_1 freely and set $s_{i+1} = s_i + \eta_i$. The two choices of s_1 produce the only two lifts, differing at every coordinate. Reversing the order of the adjacent pairs reverses their list of sums. □

Put $\mathcal{X} = \mathbb{F}_2^5$ and let $R(\eta_1, \dots, \eta_5) = (\eta_5, \dots, \eta_1)$. Let $\mathcal{Y} = \mathcal{X}/\langle R \rangle$ and let π be the quotient map. These are registry classes. Their identification with distinct chemical products is not assumed.

Proposition 8.2 (Exact quotient and baseline law)**.** *The quotient $\mathcal{Y}$ has eight singleton classes and twelve two-element classes. If μ is uniform on $\mathcal{X}$, then*

$$\bar{\mu}(y) = \begin{cases} 1/32, & |\mathcal{X}_y| = 1, \\ 1/16, & |\mathcal{X}_y| = 2. \end{cases}$$

In particular, $|\mathcal{Y}| = 20$, but $\bar{\mu}$ is not uniform on $\mathcal{Y}$.

Proof. A word is fixed by reversal exactly when $\eta_1 = \eta_5$ and $\eta_2 = \eta_4$. The three free bits give $2^3 = 8$ fixed words. The remaining $32 - 8 = 24$ words form disjoint pairs. The push-forward mass is the size of each fiber divided by 32. □

Corollary 8.3 (Four-bit class selection and five-bit word selection)**.** *For a two-element registry class, conditional selection has divergence $4 \log 2$ and retains conditional entropy $\log 2$. Exact selection of either member has divergence $5 \log 2$. For a singleton class both preparations have divergence $5 \log 2$.*

Proof. Apply (5) to masses $1/16$ and $1/32$. In a two-element fiber, μ_y is uniform and has entropy $\log 2$. The excess in (6) is therefore $\log 2$. □

The calculation recovers the class-state information distinction appearing in Article 96 [Sch26j], now as a consequence of the functional-selection theorem. The missing bit specifies a member of the pair; it is necessary only if the task distinguishes those members. On a singleton fiber there is no internal choice to preserve. At the level of six-site lifts, the corresponding orbit sizes are two and four by Lemma 8.1; these do not create an additional independent contact bit.

If a physical product map further merges registry classes, the appropriate selection law follows by composing π with that product map. Proposition 6.10 then quantifies the information saved by specifying only the product.

9 Unknown demand and optimal functional coverage

Exact-target selection assumes that the desired function is known before preparation. A source supplying a variable environment may instead need to cover several possible functions before a particular one is requested. We formulate that distinct objective explicitly.

Definition 9.1 (Demand-independent preparation)**.** Let $m = |\mathcal{Y}| \geq 2$. A preparation law p is chosen without observing a demand $y \in \mathcal{Y}$. The demand may depend on the announced preparation law but not on the realized draw. Success means $\pi(X) = y$, and its guaranteed probability is

$$\Gamma(p) = \min_{y \in \mathcal{Y}} (\pi_* p)(y).$$

Theorem 9.2 (Minimax functional coverage)**.** *One has*

$$\max_{p \in \mathcal{P}(X)} \Gamma(p) = \frac{1}{m}.$$

A law is maximizing if and only if $\pi_* p$ is uniform on $\mathcal{Y}$. Every deterministic preparation has guaranteed success zero. Every full-support reference law μ has strictly positive guaranteed success.

Proof. The smallest of m nonnegative masses summing to one is at most their average $1/m$. Equality requires all masses to equal that average. Surjectivity supplies at least one lift of the uniform output law, for example (1). A deterministic preparation produces one output, leaving at least one other output with mass zero. Finally, positivity of μ and nonempty fibers give $\bar{\mu}(y) > 0$ for every y, so their finite minimum is positive. □

Theorem 9.3 (Least-information robust source law)**.** *Among all maximizers of Γ, the unique law minimizing $D_{\mathrm{KL}}(p\|\mu)$ is*

$$p^{\mu}_{\mathrm{rob}}(x) = \frac{\mu(x)}{m\,\bar{\mu}(\pi(x))}. \tag{13}$$

Its divergence is

$$D_{\mathrm{KL}}(p^{\mu}_{\mathrm{rob}}\|\mu) = \frac{1}{m} \sum_{y \in \mathcal{Y}} \log \frac{1/m}{\bar{\mu}(y)}. \tag{14}$$

Proof. Theorem 9.2 fixes the maximizing output law to $r(y) = 1/m$. Theorem 6.4 gives the unique minimizing lift and its divergence. □

The theorem is a lexicographic optimum: first maximize guaranteed functional coverage, then minimize the change to the reference source. It preserves all conditional organization within the functional fibers while correcting unequal output frequencies.

Example 9.4 (Balanced hexamer coverage)**.** In Proposition 8.2, the baseline uniform registry source has $\Gamma(\mu) = 1/32$, whereas the best guaranteed functional coverage is $1/20$. The least-information robust law assigns

$$p^{\mu}_{\mathrm{rob}}(\eta) = \begin{cases} 1/20, & R\eta = \eta, \\ 1/40, & R\eta \neq \eta. \end{cases}$$

Its divergence is exactly

$$\begin{aligned} D_{\mathrm{KL}}(p^{\mu}_{\mathrm{rob}}\|\mu) &= \frac{8}{20} \log \frac{32}{20} + \frac{12}{20} \log \frac{16}{20} \\ &= \log \frac{8}{5} - \frac{3}{5} \log 2. \end{aligned} \tag{15}$$

This is approximately 0.0541153 nats, or 0.0780719 bits.

Proof. Use (13) and the two fiber sizes. Equation (14) gives the first line of (15); collecting logarithms gives the second. The atom probabilities sum to one because the eight singleton words contribute $8/20$ and the twenty-four remaining words contribute $24/40$. □

The robust law also has a simple acceptance realization. For the hexamer baseline, $c_{\max} = \min(20/32, 20/16) = 5/8$. Equation (10) accepts palindromic words with probability one and nonpalindromic words with probability $1/2$. The mean number of draws is $8/5$. These statements describe an ideal class-sensitive selector; constructing a molecular selector that implements these probabilities is a separate realization problem.

Remark 9.5 (What the robust separation establishes)**.** The positive comparison is with a fixed demand-independent exact placement. An endogenous ensemble can cover several as-yet-unspecified functions, whereas one fixed placed output cannot cover all of them with positive guaranteed success. A randomized exogenous preparation having the same output law achieves the same coverage. If demand is known in advance, exact placement of that demanded output instead has success one; if demand is chosen after seeing the actual draw, the present minimax game no longer applies. These are different observation conditions, not contradictory rankings of the same experiment.

10 Constructive realization by admissible dynamics

The variational comparison becomes an engineering problem when the optimal law is required to arise from allowed transitions rather than from a formal assignment of probabilities. We give two realizations. The first modifies a reversible assembly network by a Metropolis acceptance factor while preserving its allowed moves [MRR+53]. The second conditions a finite as-

sembly trajectory on a terminal functional requirement. Both constructions retain the distinction between a selected function and a selected microscopic history.

Article 79 adopts reaction–transport kinetics with irreversible export and retention hypotheses [Sch26r]; it does not supply the reversible network assumed below. Our first construction is an additional, post-preparation realization that may be considered for either source. Its role is to identify the transitions, sensing, and initialization that would make the functional optimum attainable. Article 84's reachability comparison [Sch26o] and Article 90's parameterized laws [Sch26q] do not by themselves establish those operating capabilities.

10.1 Reversible assembly and functional rate selection

Definition 10.1 (An admissible reversible network)**.** Let $\mathcal{X}$ be a finite set with at least two elements and let μ be a probability law with $\mu(x) > 0$ for every $x \in \mathcal{X}$. An admissible reversible network is a continuous-time Markov generator $Q = (q(x,z))_{x,z\in\mathcal{X}}$ satisfying

$$q(x,z) \geq 0 \quad (x \neq z), \qquad q(x,x) = -\sum_{z\neq x} q(x,z),$$

and the detailed-balance identities

$$\mu(x)q(x,z) = \mu(z)q(z,x).$$

Its move graph has edge $\{x,z\}$ precisely when $q(x,z) > 0$. We initially require this graph to be connected.

The states may describe whole assemblies while each edge changes only a specified local part. Thus locality is a property of the move graph, not an automatic consequence of finiteness. Detailed balance makes its edges bidirectional. If a chemical mechanism supplies an irreversible edge only, it does not satisfy Definition 10.1 without an additional reversible realization.

Fix an onto functional map $\pi : \mathcal{X} \to \mathcal{Y}$. Write $\bar{\mu}(y) = \sum_{\pi(x)=y} \mu(x)$ and let r be a strictly positive probability law on $\mathcal{Y}$. The functional-selection optimizer is

$$p^*(x) = \mu(x)w(x), \qquad w(x) = \frac{r(\pi(x))}{\bar{\mu}(\pi(x))}. \tag{16}$$

The weight is constant within each functional fiber. It is positive and satisfies $\sum_x \mu(x)w(x) = 1$.

Theorem 10.2 (Rate realization of functional selection)**.**

Under Definition 10.1, define, for $x \neq z$,

$$q^*(x,z) = q(x,z)\min\left\{1, \frac{w(z)}{w(x)}\right\}, \tag{17}$$

and set $q^*(x,x) = -\sum_{z\neq x} q^*(x,z)$. Then:

(i) Q^* has exactly the same move graph as Q;
(ii) Q^* is reversible with respect to p^*;
(iii) p^* is its unique stationary law;
(iv) every initial law converges to p^* as time tends to infinity.

Transitions within a functional fiber retain their original rates.

Proof. Positivity of w makes the multiplicative factor in (17) strictly positive, so the support of the off-diagonal rates is unchanged. For distinct x, z,

$$\begin{aligned} p^*(x)q^*(x,z) &= \mu(x)q(x,z)\min\{w(x), w(z)\} \\ &= \mu(z)q(z,x)\min\{w(z), w(x)\} \\ &= p^*(z)q^*(z,x). \end{aligned}$$

Summing these identities shows that $p^*Q^* = 0$. Connectedness of the finite move graph gives irreducibility. The reversible generator is self-adjoint in the inner product weighted by p^*, and its quadratic form is nonpositive. Its nullspace consists exactly of constants: vanishing of the corresponding Dirichlet form forces equality of a function across every edge and hence across $\mathcal{X}$. All other eigenvalues are therefore strictly negative. Exponentiating the generator proves convergence to the unique stationary law. If $\pi(x) = \pi(z)$, then $w(x) = w(z)$ and the multiplier is one. □

Remark 10.3 (Local moves and available information)**.** The theorem preserves local *moves* but does not assert local *computability* of their acceptance factors. To implement the ratio $w(z)/w(x)$, a mechanism must distinguish the relevant functional classes or carry an equivalent local encoding of their weights. A globally defined output predicate may require distributed sensing, memory, or additional state variables. These are explicit implementation requirements. They are not supplied merely by the fact that the underlying moves alter one contact at a time.

An endogenous realization of these rates attains the same functional law as an engineered realization of them. The source comparison is therefore constructive: the engineered system need not prescribe every microscopic outcome. It can instead reproduce the allowed network and the functional acceptance ratios, leaving the conditional fluctuations inside each fiber intact.

10.2 Finite-time approach to the optimum

Definition 10.4 (Dirichlet form and relaxation gap)**.** For a reversible generator Q with invariant probability law μ, define

$$\mathcal{D}_Q(f,f) = \frac{1}{2}\sum_{x,z}\mu(x)q(x,z)(f(z)-f(x))^2.$$

Its spectral gap is

$$\lambda_Q = \inf_{\mathrm{Var}_\mu(f)>0}\frac{\mathcal{D}_Q(f,f)}{\mathrm{Var}_\mu(f)}.$$

For a finite connected move graph, $\lambda_Q > 0$.

Theorem 10.5 (A quantitative realization bound)**.** *Let $w_{\min} = \min_x w(x)$ and $w_{\max} = \max_x w(x)$. The generator in Theorem 10.2 satisfies*

$$\lambda_{Q^*} \geq \frac{w_{\min}}{w_{\max}}\lambda_Q. \tag{18}$$

For an initial law p_0, let $p_t = p_0 e^{tQ^*}$ and define

$$\chi^2(p_0\|p^*) = \sum_x p^*(x)\left(\frac{p_0(x)}{p^*(x)}-1\right)^2.$$

Then

$$d_{\mathrm{TV}}(p_t,p^*) \leq \frac{1}{2}e^{-\lambda_{Q^*}t}\sqrt{\chi^2(p_0\|p^*)}. \tag{19}$$

The same upper bound holds for $d_{\mathrm{TV}}(\pi_* p_t, r)$.

Proof. The balance identity in the preceding proof gives

$$\mathcal{D}_{Q^*}(f,f) \geq w_{\min}\mathcal{D}_Q(f,f).$$

Since variance is the minimum squared distance to a constant,

$$\begin{aligned}\mathrm{Var}_{p^*}(f) &\leq \sum_x p^*(x)(f(x)-\mathbb{E}_\mu f)^2\\ &\leq w_{\max}\,\mathrm{Var}_\mu(f).\end{aligned}$$

The gap inequality for Q now yields (18).
Let $g_t(x) = p_t(x)/p^*(x)$. Reversibility identifies the evolution of g_t with the self-adjoint semigroup e^{tQ^*} acting on functions. The function $g_0 - 1$ has p^*-mean zero. The finite-dimensional spectral theorem therefore gives

$$\|g_t - 1\|_{L^2(p^*)} \leq e^{-\lambda_{Q^*}t}\|g_0-1\|_{L^2(p^*)}.$$

Cauchy–Schwarz applied to $2d_{\mathrm{TV}}(p_t,p^*) = \sum_x p^*(x)|g_t(x)-1|$ proves (19). Finally, summing over the fibers of π cannot increase total variation, and $\pi_* p^* = r$. □

Corollary 10.6 (A sufficient operating time)**.** *If $\chi^2(p_0\|p^*) > 0$ and $0 < \delta < 1$, any nonnegative time satisfying*

$$t \geq \frac{w_{\max}}{w_{\min}\lambda_Q}\log\frac{\sqrt{\chi^2(p_0\|p^*)}}{2\delta}$$

ensures $d_{\mathrm{TV}}(\pi_* p_t, r) \leq \delta$. When the right-hand side is negative, time zero already satisfies this sufficient bound. If $p_0 = p^*$, the conclusion holds at every time.

The bound exhibits a cost that a terminal-law comparison alone hides. Strongly uneven functional weights can slow the guaranteed relaxation through $w_{\min}/w_{\max}$. An information-optimal target law is not thereby a time-optimal construction. Different admissible move graphs can have the same invariant law and different gaps.

10.3 Exact exclusions and disconnected supports

Proposition 10.7 (Zero target masses)**.** *Permit $r(y) = 0$ and put*

$$S = \{x \in \mathcal{X} : r(\pi(x)) > 0\}.$$

Use (17) only for distinct $x, z \in S$, with diagonal entries chosen from the rates within S. This defines a generator on S reversible with respect to $p^*|_S$, which is already normalized. If the induced move graph on S is connected, $p^*|_S$ is the unique stationary law and every law initially supported on S converges to it.
If the induced graph has components $S_1, \ldots, S_k$, each component has invariant law $p^*(\cdot \mid S_j)$. For an initial law p_0 on S, the limiting law equals p^* if and only if

$$p_0(S_j) = p^*(S_j) \qquad (1 \leq j \leq k).$$

Proof. All retained weights are positive, so the detailed-balance calculation applies on every induced edge. No retained edge connects different components, and each component is a finite irreducible chain (including the one-state case). Its initial total mass is conserved. Within that component the normalized law converges to $p^*(\cdot \mid S_j)$. Consequently the limiting mass at $x \in S_j$ is $p_0(S_j)p^*(x)/p^*(S_j)$, which equals $p^*(x)$ precisely under the displayed component-mass condition. □

Remark 10.8 (What an exact target requires)**.** Restricting the generator to S means forbidding exits from S. The proposition does not move an initial state outside S into it, nor does it redistribute mass between disconnected components. Those tasks require an initialization stage or ad-

ditional admissible transitions. Thus exact functional exclusions have a sharper implementation requirement than a strictly positive target law.

11 Functional conditioning of assembly histories

An assembly may be required to reach a function by a specified time, not merely to have that function at equilibrium. A conditioned path law retains all unselected features of the baseline history. This gives an exact dynamic analogue of retaining microscopic variability within a functional class. Conditioned-process constructions have a broader nonequilibrium setting [CT15]; the finite-horizon identities needed here are proved directly.

11.1 A terminal-event transform

Definition 11.1 (Baseline trajectory and backward success function)**.** Let $T \geq 1$, let ν be a probability law on finite $\mathcal{X}$, and let $P_t(x,z)$, $0 \leq t < T$, be stochastic matrices. The baseline path law is

$$\mathbb{P}(x_0,\dots,x_T) = \nu(x_0)\prod_{t<T} P_t(x_t,x_{t+1}).$$

Fix a nonempty functional target $B \subseteq \mathcal{Y}$ and set $A = \{\pi(X_T) \in B\}$. Assume $a = \mathbb{P}(A) > 0$. Define

$$h_T(x) = \mathbf{1}_{\{\pi(x)\in B\}},$$
$$h_t(x) = \sum_z P_t(x,z)h_{t+1}(z) \qquad (0 \leq t < T).$$

Thus $a = \sum_x \nu(x)h_0(x)$.

The recurrence defines $h_t(x)$ even for states of zero baseline occupancy at time t; it is the success probability for the specified future kernels when started at x at that time. This avoids invoking a conditional probability on a zero-probability event.

Theorem 11.2 (Functional Doob transform)**.** *Set*

$$\nu^A(x) = \frac{\nu(x)h_0(x)}{a}. \tag{20}$$

For $h_t(x) > 0$, define

$$P_t^A(x,z) = P_t(x,z)\frac{h_{t+1}(z)}{h_t(x)}. \tag{21}$$

On rows with $h_t(x) = 0$, choose any stochastic row. The resulting path law $\mathbb{P}^A$ satisfies

$$\mathbb{P}^A = \mathbb{P}(\cdot \mid A).$$

In particular it reaches the terminal target with probability one, and every transition it uses from a reachable state was already allowed by the baseline kernel.

Proof. Equation (20) defines a probability law. For $h_t(x) > 0$, the backward recurrence makes the row sum in (21) equal to one. The initial law assigns zero mass to $h_0 = 0$ states. A positive-mass transformed transition from a positive-h_t state must lead to a positive-h_{t+1} state. By induction, arbitrary zero-h_t rows are never used.
Write $\omega = (x_0,\dots,x_T)$ and $H_t = h_t(x_t)$. On a transformed path of positive mass the factors telescope:

$$\frac{\mathbb{P}^A(\omega)}{\mathbb{P}(\omega)} = \frac{H_0}{a}\prod_{t<T}\frac{H_{t+1}}{H_t} = \frac{H_T}{a} = \frac{\mathbf{1}_A(\omega)}{a}.$$

Every baseline path with positive mass ending in A has positive h_t along it, since its remaining transitions provide a positive-probability route to A. Hence the identity accounts for every such path; all other paths have zero conditional mass. This proves the assertion. The support statement follows from the factor $P_t(x,z)$ in (21). □

Remark 11.3 (Initialization is part of conditioning)**.** The initial reweighting (20) cannot generally be omitted. Retaining ν while changing only the kernels produces a different law, or may start in a state from which success is impossible. No initial change is needed when the initial state is fixed and has positive success probability. More generally, no change is needed exactly when $h_0(x) = a$ on the support of ν.

Remark 11.4 (Known future dynamics)**.** The construction uses a fixed, known sequence of future kernels. It does not grant either source advance knowledge of an unobserved environment. Under environmental uncertainty, the backward functions must be computed from the information actually available, or their implementation must be justified by an enlarged causal state model.

11.2 Optimality on path space

For a possibly vanishing finite reference law, extend relative entropy by setting $D_{\mathrm{KL}}(\mathbb{R}\|\mathbb{P}) = +\infty$ unless $\mathbb{R} \ll \mathbb{P}$. Otherwise use the finite sum over the positive mass of $\mathbb{R}$, where $\mathbb{P}$ is then

positive.

Theorem 11.5 (Least-information successful trajectory law)**.** *Among all probability laws $\mathbb{R}$ on $\mathcal{X}^{T+1}$ with $\mathbb{R}(A) = 1$, the unique minimizer of $D_{\mathrm{KL}}(\mathbb{R}\|\mathbb{P})$ is $\mathbb{P}^A$. Its value is*

$$D_{\mathrm{KL}}(\mathbb{P}^A\|\mathbb{P}) = -\log a.$$

More precisely, whenever $\mathbb{R} \ll \mathbb{P}$ and $\mathbb{R}(A) = 1$,

$$D_{\mathrm{KL}}(\mathbb{R}\|\mathbb{P}) = D_{\mathrm{KL}}(\mathbb{R}\|\mathbb{P}^A) - \log a. \tag{22}$$

Laws not absolutely continuous with respect to $\mathbb{P}$ have infinite divergence.

Proof. On every path in A with positive baseline mass, $\mathbb{P}^A = \mathbb{P}/a$. Substitution in the finite relative-entropy sum gives (22). Nonnegativity of relative entropy proves minimality, and its equality condition gives uniqueness. Setting $\mathbb{R} = \mathbb{P}^A$ yields the stated value. □

Corollary 11.6 (The informational price of a prescribed history)**.** *For a baseline-positive successful path $\omega \in A$,*

$$D_{\mathrm{KL}}(\delta_\omega\|\mathbb{P}) - D_{\mathrm{KL}}(\mathbb{P}^A\|\mathbb{P}) = -\log \mathbb{P}(\omega \mid A).$$

The excess is strictly positive unless that path is the only baseline-positive successful path.

The distinction is relevant to an engineered source. Specifying an entire assembly schedule fixes more information than guaranteeing its terminal function while retaining admissible alternative histories. The Doob construction gives the least-change history law, not a free implementation: its time-dependent success functions must be computed, encoded, sensed, or otherwise realized, and its initialization must be supplied.

Proposition 11.7 (A prescribed terminal functional law)**.** *Let q be the baseline law of $\pi(X_T)$, and let r be a probability law supported on* $\operatorname{supp} q$*. Define*

$$g(y) = \begin{cases} r(y)/q(y), & q(y) > 0, \\ 0, & q(y) = 0. \end{cases}$$

Replace $h_T(x)$ in Definition 11.1 by $g(\pi(x))$ and retain the backward recurrence. Then $\sum_x \nu(x)h_0(x) = 1$. The initial law $\nu^r(x) = \nu(x)h_0(x)$ and the positive-row kernels (21) produce the path law

$$\mathbb{P}^r(\omega) = \mathbb{P}(\omega)g(\pi(x_T)).$$

Its terminal functional law is r, and it uniquely minimizes path relative entropy among all laws with that terminal functional law. In particular,

$$D_{\mathrm{KL}}(\mathbb{P}^r\|\mathbb{P}) = D_{\mathrm{KL}}(r\|q).$$

Proof. The terminal weight has baseline expectation $\sum_{y:q(y)>0} q(y)r(y)/q(y) = 1$. The same backward summation and telescoping proof as in Theorem 11.2 therefore gives a normalized transformed path law of the stated form. Summing it over paths with terminal output y gives $r(y)$. Conditional on any terminal output with positive target mass, its path law is exactly the baseline conditional path law. Competitors not absolutely continuous with respect to $\mathbb{P}$ have infinite divergence and cannot minimize, since the target is supported on $\operatorname{supp} q$ and the displayed candidate has finite divergence. For the remaining competitors, restrict the path space to $\operatorname{supp}\mathbb{P}$ and the output space to $\operatorname{supp} q$. The reference is strictly positive there, so Theorem 6.4 applies. The finite relative-entropy chain rule, applied to the path space and its terminal-output map, decomposes every competing law's divergence into $D_{\mathrm{KL}}(r\|q)$ plus the r-weighted conditional divergences. Those terms are nonnegative and vanish together only for $\mathbb{P}^r$. □

11.3 A criterion for functional rather than microscopic control

Definition 11.8 (Strong functional lumpability)**.** The baseline kernels are strongly lumpable under π if, for each t, y, y', the sum

$$\sum_{\pi(z)=y'} P_t(x, z)$$

has the same value for every x with $\pi(x) = y$. Denote that value by $\bar{P}_t(y, y')$.

Proposition 11.9 (Conditioning can be controlled on the quotient)**.** *Under Definition 11.8, $h_t(x) = \bar{h}_t(\pi(x))$, where*

$$\bar{h}_T(y) = \mathbf{1}_B(y),$$
$$\bar{h}_t(y) = \sum_{y'} \bar{P}_t(y, y')\bar{h}_{t+1}(y').$$

For every positive-$\bar{h}_t(y)$ row, the conditioned quotient kernel is

$$\bar{P}^A_t(y, y') = \bar{P}_t(y, y')\frac{\bar{h}_{t+1}(y')}{\bar{h}_t(y)}.$$

Furthermore, if x and a destination fiber y' have positive conditioned transition probability, conditioning does not change the relative probabilities of the destination states

inside that fiber:

$$\frac{P_t^A(x,z)}{\sum_{\pi(v)=y'} P_t^A(x,v)} = \frac{P_t(x,z)}{\sum_{\pi(v)=y'} P_t(x,v)} \quad (\pi(z)=y').$$

Proof. The assertion about h_T is immediate. Assuming the assertion at time $t+1$, group the defining sum for h_t by destination fibers and apply strong lumpability. This proves the backward assertion by induction. Summing (21) over a destination fiber gives the quotient kernel. Within that fiber the multiplier $\bar{h}_{t+1}(y')/\bar{h}_t(y)$ is constant, so it cancels from the ratio in the last display. □

This is a precise route by which engineering can approach an endogenous capability without duplicating every microscopic specification. When lumpability holds, the future success weighting depends only on the current functional class; the microscopic conditional transitions can remain those of the underlying assembly network. Strong lumpability is sufficient, not necessary, for fiber-constant backward weights for a particular target and horizon. Without it, configurations with the same current output may have different future success probabilities; the actual transformed transition factors must be checked against the available observations and admissible local couplings. Additional internal state or sensing may be needed when those factors cannot be implemented from the functional observations alone. A shared terminal product map alone does not establish this stronger dynamic equivalence.

12 Coordination with lattice and response structures

The preceding results compare sources on a common state space. We now specify how additional structures from the corpus enter that state space or its observations. The relevant distinction is between a structure attached to an admissible configuration and a law assigning weight to configurations. The former is shared when the realization is shared; the latter may be source-dependent.

12.1 Coefficient-consistent registry fields

Let $G=(V,E_G)$ be a finite graph with c connected components. Choose an orientation for each edge, and let $\mathbb{K}$ be a finite field with q elements. Define

$$C^0(G;\mathbb{K}) = \mathbb{K}^V, \qquad C^1(G;\mathbb{K}) = \mathbb{K}^{E_G}.$$

For an oriented edge $v \to w$, put

$$(\delta s)(v,w) = s(w) - s(v).$$

An edge field is *integrable* if it is δs for some vertex field s. The binary registry convention is the case $\mathbb{K}=\mathbb{F}_2$. A quaternary extension is a different coefficient model, not a relabeling that creates a second physical bit for free.

Theorem 12.1 (Registry integration and its dimension)**.** *An edge field a is integrable if and only if its signed sum around every graph cycle is zero. Moreover,*

$$\dim_{\mathbb{K}} \operatorname{im}\delta = |V| - c, \qquad |\operatorname{im}\delta| = q^{|V|-c}.$$

It suffices to impose the cycle conditions on a fundamental cycle basis associated with a spanning forest.

Proof. If $a=\delta s$, the signed sum along a closed walk telescopes to zero. Conversely, choose one root in each component and assign it the value zero. For each other vertex define $s(v)$ as the signed sum of a along the unique forest path from the root to v. Every forest edge then satisfies $a=\delta s$. An edge outside the forest closes a fundamental cycle; the zero-sum condition on that cycle supplies the same equality on the remaining edge.
The kernel of δ consists exactly of functions constant on each component, and hence has dimension c. Rank–nullity gives the dimension of its image. A vector space of dimension $|V|-c$ over a q-element field has the stated cardinality. □

This theorem specifies when edge assignments may be sampled independently. A forest has no cycle conditions, so every edge field is integrable. On a graph with cycles, independent full-support edge draws assign positive mass to nonintegrable fields unless the sampling is subsequently constrained. The admissible configuration set must decide which interpretation is intended.

Example 12.2 (Columns and a connected finite lattice)**.** The $3\times3\times6$ vertex box has 54 vertices. Its vertical stacking graph consists of nine disjoint six-site columns, with 45 edges and nine components. It therefore has 45 independent binary integrable registry coordinates.
If all nearest-neighbor lateral edges are included as additional measured relations, the graph has

$$|E_G| = 36+36+45 = 117, \qquad c=1.$$

Its integrable binary edge fields have dimension 53, and its cycle space has dimension

$$117-54+1 = 64.$$

The extra eight independent coordinates compared with the vertical forest are relative references between columns; they are not additional native vertical contacts. The remaining edge data are constrained by the cycle conditions.

Whether the lateral molecular connectors physically carry these relations must be specified in the realization.

Article 94 and the earlier topos working manuscript motivate directional bridge and cochain descriptions [Sch26s, Sch26l]. Theorem 12.1 uses one coefficient field throughout and proves a graph-level statement directly. If a cellular model replaces all cycle conditions by plaquette conditions, it must additionally establish that the chosen boundaries span the relevant cycle space. No identification of real and quaternary cochains is needed here.

12.2 Reliability as a common stochastic module

The finite-code constructions in Article 98 and the earlier arithmetic-geometry working manuscript [Sch26k, Sch26b] can be connected to source comparison once the physical alphabet, encoder, noise, and decoder are fixed.

Definition 12.3 (An encoded functional channel)**.** Let $\mathcal{M}$ be a finite message set, let $\mathrm{Enc} : \mathcal{M} \to \mathcal{A}^n$ be an encoder, let $N(z \mid v)$ be a stochastic noise kernel on transmitted words, and let Dec be a decoder into $\mathcal{M} \cup \{\perp\}$. For a specified correct message m, define

$$s(m) = \sum_{z:\,\mathrm{Dec}(z)=m} N(z \mid \mathrm{Enc}(m)).$$

The source supplies a law ρ on $\mathcal{M}$; the average success is $\sum_m \rho(m)s(m)$.

Proposition 12.4 (Common reliability and source weighting)**.** *If two sources use the same encoder, noise kernel, and decoder, and induce the same message law, their decoded output laws and average success probabilities agree. If their message laws have total-variation distance at most δ, their average success probabilities differ by at most δ.*

Proof. The encoder, noise kernel, and decoder compose to a common stochastic kernel from messages to decoded outputs. Equal input laws have equal pushforwards. Since $0 \le s(m) \le 1$, the difference in expectations of s is bounded by total variation: subtracting any common constant and separating the positive and negative parts of $\rho - \widetilde{\rho}$ gives this bound. □

A minimum-distance theorem is a property of a code, not of the origin of its encoder. A source may nevertheless differ in the cost or feasibility of implementing the encoder and syndrome readout. Those constraints belong to the admissible policy class.

Five binary contact agreements have 2^5 possible words. Five symbols over $\mathbb{F}_4$ have 4^5 possible words. Importing a quaternary contact code into the binary registry model therefore requires an additional realized coordinate or a specified block encoding. The four-state algebra in Article 92 [Sch26m] provides an algebraic vocabulary; a physical encoding map supplies the separate realization.

12.3 Finite operator observations

Let $\mathcal{H} = \mathbb{C}^{\mathcal{X}}$ with distinguished basis $\{|x\rangle : x \in \mathcal{X}\}$. Embed a classical source law as

$$\rho_p = \sum_{x\in\mathcal{X}} p(x)|x\rangle\langle x|.$$

A classical observable $f : \mathcal{X} \to \mathbb{R}$ is represented by $A_f = \sum_x f(x)|x\rangle\langle x|$, so

$$\mathrm{Tr}(\rho_p A_f) = \sum_x p(x)f(x).$$

This is an exact finite-state representation. It asserts neither coherent preparation nor a quantum advantage.

Proposition 12.5 (A verified stochastic channel lift)**.** *For a stochastic kernel $K(z \mid x)$, with $x, z \in \mathcal{X}$, define*

$$L_{zx} = \sqrt{K(z \mid x)}\,|z\rangle\langle x|.$$

Then the map

$$\Phi_K(\rho) = \sum_{z,x} L_{zx}\rho L_{zx}^*$$

is completely positive and trace preserving, and $\Phi_K(\rho_p) = \rho_{pK}$. It removes off-diagonal input coherences.

Proof. The displayed Kraus representation gives complete positivity. Moreover,

$$\sum_{z,x} L_{zx}^* L_{zx} = \sum_x \left(\sum_z K(z \mid x)\right)|x\rangle\langle x| = I.$$

Thus the map preserves trace. Direct multiplication leaves the output coefficient at $|z\rangle\langle z|$ equal to $\sum_x K(z \mid x)\rho_{xx}$, proving both remaining statements. □

This construction supplies a valid interface to the channel language of Article 96 [Sch26j] without inferring a complete depolarizer from a registry symmetry alone. If coherent quantum behavior is the intended observation, the full channel and its accessible operations must instead be specified; agreement of diagonal probabilities is then insufficient.

The transfer and spectral structures developed in Articles 93 and 95 and the earlier arithmetic-geometry working manuscript

[Sch26i, Sch26h, Sch26b] may likewise be used as common response observables or distances for a declared finite realization. Equal configuration laws give equal distributions of any deterministic response attached to those configurations. A spectral lower bound or a change of coefficient algebra is not, by itself, a distinction between endogenous and exogenous origin.

12.4 Geometry and source probability

The single-core architecture gives a reusable basis for these modules: locant assignments, registry relations, finite graphs, code coordinates, and response operators can all be typed relative to the same constitutional framework. Coordination consists of specifying the maps between them. It does not require every structure to be equivalent to every other structure.

For example, a registry field may determine an encoded word, an encoded word may determine a decoder output law under noise, and a molecular configuration may determine an electrical response under a chosen Hamiltonian. Each arrow has a domain, a codomain, and assumptions. Composing verified arrows yields a functional map or a stochastic observation kernel to which the source comparison applies. A shared dimension, cardinality, or analogy does not supply a missing arrow.

13 Engineering toward endogenous capability

13.1 A constructive design sequence

The theory suggests a sequence of engineering questions rather than a demand to reproduce every microscopic feature of an endogenous source.

Specify the function. Choose the map π or the observable history whose performance matters. A target molecular word, a registry equivalence class, a decoded output, and a sustained adaptive response define different tasks. In particular, unnecessary refinement of the target can add conditional information requirements without improving the intended function.

Specify the accessible configuration graph. Determine which substitutions, arm states, attachments, references, and bridge operations are actually available. Article 79's source kinetics, the cleavage and supply analyses [Sch26r, Sch26n, Sch26p], and the directional-bridging model [Sch26s] constrain different parts of this graph. A positivity assumption in a finite mathematical model is not a replacement for identifying a physical transition.

Retain useful conditional organization. If the functional task admits multiple configurations, a placed system need not choose one permanently. It can prepare a randomized mixture or install local dynamics with the desired invariant law. The fiber-selection theorem identifies the particular conditional distribution that preserves the reference source organization while imposing the required output distribution.

Supply the information used by adaptation. The binary kinetic example separates a responsive mechanism from an uninformed preparation. To close this gap, engineering must supply the environmental coupling, not merely a more accurately prepared initial state. Local sensing may be chemical rather than electronic, but it remains an explicit part of the realization.

Control the approach and its expense. A stationary-law construction supplies a target distribution; the gap bound supplies a sufficient approach time. A conditioned trajectory supplies a successful path law; the backward success function identifies the information its implementation requires. Approximate emulation supplies error bounds, with the corresponding cost slack, over a specified horizon.

13.2 What can be specifically endogenous

An endogenous mechanism can exhibit a strict advantage relative to an explicitly restricted engineering protocol. A mechanism that retains a nontrivial functional fiber changes less information than one required to specify an exact microstate. A demand-independent diverse source can cover an unknown functional requirement that a single fixed output misses. An environment-responsive source can exceed the minimax success of an uninformed initial preparation. These are proved comparisons of distinct attainable laws or responses.

The same results also identify their boundaries. An engineered sampler can reproduce the selected terminal law. An engineered reversible network can reproduce its invariant law and relaxation behavior. An engineered sensor and generator can reproduce the responsive kinetic witness. An engineered realization of the conditioned kernels can reproduce the successful trajectory law. The stronger the allowed exogenous class, the more operations it can share with the endogenous mechanism.

Thus an endogenous-only claim is meaningful only with a stated comparison class and objective. For the broad engineering class envisaged here, an emulation theorem is as important as a separation theorem: it identifies how the supposed limitation can be overcome.

13.3 Information, supply, and physical work

The information objective compares laws relative to a declared reference source. It is not a universal currency for chemical manufacture. The acceptance construction quantifies a distinct expenditure in repeated draws. The local-dynamics construction quantifies a distinct time requirement. Environmental coupling, reference retention, error correction, and recovery from rejected assemblies may impose further costs.

Converting a relative-entropy difference into a lower bound on work requires a physical model of equilibrium, energies, reservoirs, allowed processes, and information reset. The present construction does not require such a conversion. This separation is useful: a rigorous information optimum remains meaningful when actual chemical energies are unknown, while an asserted energy optimum would not.

Likewise, a supply law with the correct mean composition does not ensure adequate supply in every finite cycle. Repeated trials, retention losses, waiting times, and incomplete bridging can be incorporated as state and cost variables. The finite-horizon formulation accommodates those additions, but does not silently identify mean throughput with reliability.

13.4 Research implications for the common-core hypothesis

The common molecular core reduces the number of independent constitutional starting points in the model and supports a shared material ledger. The remaining difficulty is concentrated in controlled transformations and organization: which arms survive, which contacts distinguish alignment, which references persist, which bridges form, and which outputs can be selected without over-specification.

This reframing is useful for both origins. Endogenous research asks whether local interactions in a mixture implement the requisite filters and transition laws. Exogenous research asks which of those operations can be installed, approximated, or replaced while preserving the same functional distribution and response. A hybrid realization may use manufacturing to establish the stock or reference and then allow local assembly to distribute the remaining degrees of freedom.

The source comparison therefore complements the genetic-code, ledger, and ICL studies. The earlier constructions specify downstream organizations and functions. The present framework specifies how a source law can be connected to those functions, and how the needed organization can be allocated between preparation, local dynamics, observation, and external intervention.

14 Conclusion

Endogenous and exogenous sources can occupy the same finite molecular configuration space while realizing different distributions and operating histories. Functional fibers provide the central comparison object. Their conditional laws identify what a source may retain while a desired function is imposed, and relative entropy separates necessary output selection from unnecessary microscopic specification.

Explicit acceptance, reversible-rate, and trajectory-conditioning constructions make this comparison operational. Robust demand and responsive kinetics exhibit strict advantages over designated fixed-placement classes. Exact and approximate emulation then show how engineering can recover the same capabilities under stated observation, transition, and resource conditions.

The resulting mathematical framework supports neither an automatic superiority of self-assembly nor an automatic superiority of placement. It identifies shared capabilities, conditional optima, and constructive routes between realizations. For the substituted-chrysene program, this completes the fourth pilot's source-level connection between a common molecular stock and its encoded, functional organization.

A Hypotheses and source precedence

A.1 Four levels of assertion

The manuscript distinguishes:

(i) definitions of finite mathematical objects;
(ii) theorems proved from their displayed hypotheses;
(iii) inherited molecular architectures used to motivate those objects;
(iv) physical realization requirements for the mechanisms.

The last level includes chemical accessibility, rate control, reference stability, sensing, and thermodynamic interpretations. A proof at the second level does not establish a requirement at the fourth.

A.2 Reachability

The common-reachability result requires positive probability for every admitted endogenous configuration and surjectivity of the placement operation. Conditional positivity along a valid sequential construction is sufficient. Merely positive one-site marginals are not sufficient: two bits constrained to agree can each have positive marginals while the configurations 01 and 10 remain impossible.

For a laterally bridged lattice, equal bridge access is an additional premise. Directional compatibility must be enforced, not replaced by a scalar arm count. The completed occupied geometry is held fixed unless geometry-changing states and moves are explicitly added.

A.3 Registry and the observed function

Registry is assigned to contacts. A six-site column has five native binary agreement coordinates. A symmetry quotient requires a declared physical equivalence, and a registry-only product map requires fixed composition and the applicable invariance. An asymmetric reference is retained when needed to distinguish the observed alternatives.

The source law is positive on the chosen admissible space. Zero target masses are allowed in finite conditioning but require restricted support or appropriate initialization in dynamic realizations. Uniformity on microstates and uniformity on functions are distinct.

A.4 Dynamics and engineering

Reversible rate tilting assumes a finite reversible baseline network with positive stationary law. This is an optional realization, not a reinterpretation of Article 79's irreversible export model. Exact conditioning of trajectories requires a positive baseline probability of the selected event and uses the backward success functions. Strong lumpability is an additional condition when control is claimed to require only the functional quotient.

Emulation must use one admissible policy uniformly across the allowed environments. Environmental information unavailable to the implementation cannot enter that policy. Approximation bounds use a common observation convention and either bounded common path costs or separately specified cost-preserving guarantees.

B Corpus coordination and scope of inheritance

The following source roles determine the dependency structure; they do not import every claim of each cited article.

Genetic code, protocell ledger, finite ICL
Articles 75, 77, and the finite-ICL Article 88 supply the three preceding pilot constructions [Sch26d, Sch26c, Sch26a]. They motivate the output functions, component organization, and finite interior setting.

The source and the placed source
Articles 79 and 84 anchor the reaction–transport and placement comparison [Sch26r, Sch26o]. For Article 84, the revised `acf84_placed (1).tex` controls: it places registry on contacts, identifies the reference issue, and states equal bridging access separately.

Cleavage, bounds, and source laws
The earlier cleavage and placement-budget working manuscripts supply conditional arm-state and source-bound considerations [Sch26n, Sch26p]. Article 90 supplies law factorization, visibility, and persistence questions [Sch26q]. Their role is to constrain possible mechanisms, not to identify every information statistic with an optimized capacity.

Symmetry, cipher, and transfer
Articles 91–93 supply registry, algebraic, and response structures [Sch26e, Sch26m, Sch26i]. These are shared modules when their realization is shared. The induced probability on their sectors or response values is source-dependent.

Bridging and spectral structure
Articles 94 and 95 motivate directional bridge constraints and finite metric or infinite-algebra questions [Sch26s, Sch26h]. Finite and infinite algebras are different scale models; their difference is not an endogenous/exogenous invariant.

Channels and control
Articles 96 and 97 motivate channel and information-control descriptions [Sch26j, Sch26g]. The present paper uses explicitly verified stochastic kernels and the Kraus construction in Proposition 12.5. It does not require a logistic error statistic to equal an induced H_∞ norm.

Codes, finite invariants, and cochains
Article 98 and the earlier arithmetic-geometry and topos working manuscripts motivate finite codes, graph and spectral invariants, and lattice cochains [Sch26k, Sch26b, Sch26l]. Encoders and coefficient maps are specified separately. Theorem 12.1 is proved over one field and does not rely on an equivalence between real, binary, and quaternary complexes.

B.1 Numbering and version identity

The bibliography distinguishes published identities from earlier working titles. In particular, Article 88 is the finite-ICL pilot. The earlier cleavage, placement-budget, arithmetic-geometry, and topos manuscripts retain descriptive historical labels pending confirmation of their current publication identifiers. The earlier arithmetic-geometry label "Article 99 does not identify the present source-comparison paper.

B.2 Preservation and relaxation

Independent vertex flips and independent contact flips define different registry processes on a chain: an interior vertex flip changes two adjacent contact agreements. A rate or mixing bound derived for one process cannot be transferred to the other without a generator calculation.

Similarly, convergence toward a stationary random ensemble is not preservation of a selected word. The present paper labels the target of each dynamic estimate explicitly: convergence to a selected distribution, success at a terminal event, or response to an environment.

C Verification and formal-proof coverage

The proofs in the main text are analytic finite-dimensional arguments. Two standalone companions provide reproducible

checks and a separately delimited Lean development [Sch26t, Sch26f]. Their coverage is stated independently of the fact that their LaTeX documents compile.

C.1 Executed finite verification

The accompanying `verify_a99.py` completed 213 checks in seven blocks:

(i) registry enumeration, reversal fibers, and balanced class weights;

(ii) finite information identities, acceptance laws, and minimax examples;

(iii) reversible generators, selected laws, spectral-gap comparisons, finite-time bounds, and disconnected supports;

(iv) fully enumerated event-conditioned and terminal-law trajectories;

(v) approximate emulation, resource slack, and responsive binary kinetics;

(vi) binary and quaternary finite-field graph ranks;

(vii) finite stochastic-channel Kraus identities.

Integer, rational, and finite-field checks use exact arithmetic. Logarithms, eigenvalues, matrix exponentials, linear optimization, and matrix-channel checks use floating-point arithmetic with the declared 2×10^{-11} absolute tolerance. The companion records the exact software versions, source hash, commands, individual checks, and results. These tests establish the reported finite instances, not universal theorems or chemical feasibility.

Negative controls test omitted initial reweighting, incorrect uniformity of quotient classes, inappropriate unchanged transition rates, forbidden support, missing resource slack, and nonintegrable edge fields. A nonlumpable three-state example with fiber-constant target-specific backward weights guards the sufficiency qualification following Proposition 11.9.

C.2 Machine-checked structural and rational core

The companion `FormalProof_A99.lean` was compiled using Lean 4.24.0. It contains 25 checked theorem declarations covering observation reachability, emulation transfer and transitivity, finite rational normalization identities, acceptance algebra, backward-weight row normalization, arbitrary-finite-horizon rational Doob telescoping, an integer-weight minimax bound, and exact hexamer identities.

The finite spaces in the rational normalization results are represented by explicit lists; duplicate-free exhaustive lists and the stated positivity premises supply their probability interpretation. The telescoping theorem assumes nonzero backward weights along the path. The code does not use `sorry`, added custom axioms, or `native_decide`. Its reported axiom dependencies are the standard Lean principles listed in the companion.

The full real-logarithmic KL optimality theorems, spectral analysis, continuous-time convergence, general total-variation/resource bounds, general finite-field integration theorem, and molecular realization claims are not machine-formalized by this module. Their mathematical support remains the displayed analytic proofs; selected finite instances are additionally checked by the Python suite. The formal companion maps each checked declaration to the corresponding manuscript component without representing an algebraic identity as an entire probability or physical theorem.

References

[Csi75] Imre Csiszár. *I*-Divergence Geometry of Probability Distributions and Minimization Problems. *The Annals of Probability*, 3(1):146–158, 1975. https://doi.org/10.1214/aop/1176996454.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, 2 edition, 2006. https://doi.org/10.1002/047174882X.

[CT15] Raphaël Chetrite and Hugo Touchette. Nonequilibrium Markov processes conditioned on large deviations. *Annales Henri Poincaré*, 16, 2015. https://doi.org/10.1007/s00023-014-0375-8.

[MRR+53] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. https://doi.org/10.1063/1.1699114.

[Sch26a] Charles D. Schaper. A Finite Axiomatic Framework for the Interior Chrysene Lattice and Protocell Organization. *Annals of the Chrysene Formalism*, 1(4):1778–1807, 2026. Article 88. https://doi.org/10.68366/2pyt7f91.

[Sch26b] Charles D. Schaper. Arithmetic Geometry, Spectral Zeta Theory, and Non-Commutative Geometry of the Chrysene Formalism: The (9, 8) Hermitian Curve over GF(4), Genus-1 Goppa Codes, the Ihara–Bass Zeta Function, and Connes' Spectral Triple, 2026. Earlier supplied working manuscript, formerly numbered 99; final publication number and DOI not assigned here.

[Sch26c] Charles D. Schaper. Carbon Traceable Protocell Construction From Substituted Chrysene Stocks. *Annals of the Chrysene Formalism*, 1(4):1567–1605, 2026. Article 77. https://doi.org/10.68366/wvna7r05.

[Sch26d] Charles D. Schaper. Constructive Genetic Code-Table Realization On A Typed Chrysene Carrier. *Annals of the Chrysene Formalism*, 1(4):1503–1544, 2026. Article 75. https://doi.org/10.68366/ees7wd38.

[Sch26e] Charles D. Schaper. Emergent Assembly Chirality and Flip Kinematics: Euclidean Stabilizers of the Frame Classes, Enantiotopic Faces of the C_{2h} Core, and Initiator Capping of the Interior Lattice. *Annals of the Chrysene Formalism*, 1(4):1836–1846, 2026. Article 91. https://chrysene.com/index.php/acf/article/view/91.

[Sch26f] Charles D. Schaper. Formal Proof Companion for Article 99, 2026. Standalone Lean 4.24.0 companion; machine-checked structural and exact rational core.

[Sch26g] Charles D. Schaper. Information-Theoretic Mechanics of Survival in Maximum-Entropy Environments: Static Minimax Disturbance Attenuation, Coordination-Deficit Costs, and the Placed Source as Categorical Right Adjoint. *Annals of the Chrysene Formalism*, 1(4):1890–1897, 2026. Article 97. https://chrysene.com/index.php/acf/article/view/97.

[Sch26h] Charles D. Schaper. K-Theory of the Sequence Algebra, Spectral Dimension Three, and Path-Calibrated Spectral Metrics: The Pimsner–Voiculescu Sequence on $Q \rtimes_\alpha \mathbb{Z}^3$, Unconditional Morita Non-Equivalence, and Exact Metric Compatibility. *Annals of the Chrysene Formalism*, 1(4):1876–1882, 2026. Article 95. https://chrysene.com/index.php/acf/article/view/95.

[Sch26i] Charles D. Schaper. Quantum Behavior of Substituted-Chrysene Molecular Assemblies: Bipartite Schur-Complement Transfer Anisotropy, Quadratic-to-Linear Stark Switching Across (D, I, T), and Registry-Gated Ballistic Transport. *Annals of the Chrysene Formalism*, 1(4):1856–1865, 2026. Article 93. https://chrysene.com/index.php/acf/article/view/93.

[Sch26j] Charles D. Schaper. Quantum Information Dynamics of the Source and the Placed Source: Lindblad Dynamical Scission, Von Neumann Entropy Production, and Landauer-Bounded Quantum-Classical Transduction. *Annals of the Chrysene Formalism*, 1(4):1883–1889, 2026. Article 96. https://chrysene.com/index.php/acf/article/view/96.

[Sch26k] Charles D. Schaper. Quaternary Coding Theory of the Biological Ledger over GF(4): Evaluation Codes, Plaquette Checks, and Conditional Syndrome Decoding. *Annals of the Chrysene Formalism*, 1(4):1898–1904, 2026. Article 98. https://chrysene.com/index.php/acf/article/view/98.

[Sch26l] Charles D. Schaper. The Chrysene Topos and Non-Cocommutative Hopf Algebra: Drinfeld Twists of the C_{2h} Step-Feature, Cellular Sheaf Cohomology of Topological Free Energy, and Centennial Synthesis of the Formalism, 2026. Earlier supplied working manuscript, formerly numbered 100; final publication number and DOI not assigned here.

[Sch26m] Charles D. Schaper. The Dual Cipher, Structurally Induced GF(4) Multiplication, and Real Spectral Structure: Companion-Matrix Realization of the Register Automorphism, Frobenius Reversal, and the Order-One Commutant Condition. *Annals of the Chrysene Formalism*, 1(4):1847–1855, 2026. Article 92. https://chrysene.com/index.php/acf/article/view/92.

[Sch26n] Charles D. Schaper. The Partially Sighted Source: Arm Cleavage as an Orientation Channel Without a Direction, 2026. Earlier supplied working manuscript, formerly numbered 85; final publication number and DOI not assigned here.

[Sch26o] Charles D. Schaper. The Placed Source: Exogenous Preparation of the Substituted-Chrysene Assembly, and the Equality of Its Reachable Set with the Endogenous One. *Annals of the Chrysene Formalism*, 1(4):1703–1731, 2026. Article 84. https://doi.org/10.68366/a4tjv735.

[Sch26p] Charles D. Schaper. The Placement Budget: Bounds Any Exogenous Assembler Must Satisfy, and the One Equilibrium Cannot Meet, 2026. Earlier supplied working manuscript, formerly numbered 86; final publication number and DOI not assigned here.

[Sch26q] Charles D. Schaper. The Source Duality: Endogenous and Exogenous Preparation of the Substituted-Chrysene Stock, Mixed Arm Words, and the Persistence Semigroup. *Annals of the Chrysene Formalism*, 1(4):1823–1835, 2026. Article 90. https://chrysene.com/index.php/acf/article/view/90.

[Sch26r] Charles D. Schaper. The Source: Substitution-Initiated Self-Separation of the Substituted-Chrysene Stock from a Polyaromatic Bulk, and the Supply It Delivers to the Protocell. *Annals of the Chrysene Formalism*, 1(4):1616–1638, 2026. Article 79. https://doi.org/10.68366/s8sfwq50.

[Sch26s] Charles D. Schaper. Topological Fault Tolerance, Fiedler Routability, and Partially Bridged Percolation of the Interior Chrysene Lattice: Homological Bypasses, Directional Attachment Constraints, and the Exclusion of Degenerate Star Graphs. *Annals of the Chrysene Formalism*, 1(4):1866–1875, 2026. Article 94. https://chrysene.com/index.php/acf/article/view/94.

[Sch26t] Charles D. Schaper. Verification Code for Article 99, 2026. Standalone computational companion; verify_a99.py and recorded finite checks.