Real Patterns Need Closure: Transition Autonomy as a Dynamical Criterion for Macro-Objecthood

Anonymous Author[1*]

[1*]Anonymous Institution, Anonymous, Anonymous.

Corresponding author(s). E-mail(s): anonymous@example.com;

Abstract

Dennett's real-pattern realism links ontology to compression and prediction, but compression alone is too permissive: contrived codings and dynamically idle aggregates can satisfy it. This paper argues that the missing ingredient is an explicit closure condition. A candidate macro-object qualifies when, for a fixed regime, horizon, and admissible intervention class, macrostate information is sufficient for macro-transitions, so within-class micro-differences do not change macro-level what-follows. In exact Markov settings, strong lumpability provides a benchmark realization of this condition. In non-ideal settings, closure is graded through leakiness and convergent diagnostics under fixed constraints. The result is realist but disciplined: it excludes high-leak composites, supports graded verdicts in borderline cases, and remains compatible with microphysical completeness. By tying commitment strength to robustness under anti-gerrymandering tests, the framework explains why selective realism is principled rather than discretionary. The paper contributes a philosophical criterion and an audit protocol for applying it, while leaving domain-specific estimator engineering to downstream methodological work.

Keywords: real patterns, closure, macro-objecthood, strong lumpability, coarse-graining, pattern realism

1 Introduction: Why Pattern Realism Needs a Stricter Criterion

Dennett's core thought is still compelling. Scientific and everyday inquiry often succeeds by tracking patterns rather than exhaustively tracking microstates, and this success is not always reducible to convenience language. His criterion is explicit: "A

pattern exists in some data—is real—if there is a description of the data that is more efficient than the bit map, whether or not anyone can concoct it" (Dennett 1991, p. 34). In that sense, real-pattern realism captures something correct about the structure of explanation.

The persistent difficulty is permissiveness. Dennett himself anticipates the crudest version of this worry: a coding scheme that simply "baptizes" individual frames with proper names achieves no genuine compression, because it collapses to the full bit map whenever the target extends beyond the memorized cases (Dennett 1991, pp. 32–33). His response is that the coding must be part of an entirely general system, not an ad hoc renaming trick. That constraint excludes the most obvious cheats. But it does not close the gap. A coding strategy can be genuinely general and non-trivially compressive while still grouping together microstates whose onward transition profiles diverge. Such constructions compress, yet they require persistent within-class bookkeeping to maintain forecast quality, precisely because the macro-label hides transition-relevant heterogeneity. This is the residual permissiveness that survives Dennett's own anti-cheating move, and it is the target of the present paper.

The worry has a well-established lineage beyond Dennett's initial discussion. Haugeland pressed the question of when pattern detection crosses from descriptive convenience into genuine ontological contact (Haugeland 1998), and later work on idealization continues to press the gap between descriptive utility and epistemic warrant (Elgin 2017, chap. 2). The present paper takes that critique seriously: residual permissiveness, not the trivial kind Dennett already blocks, is the central obstacle standing between real-pattern realism and a defensible macro-ontology.

A contrast makes the gap concrete. Lewis observed that we have no principled reason to deny existence to the mereological sum of "the right half of my left shoe plus the moon plus the sum of all her Majesty's ear-rings," however sensible it is to ignore such things in ordinary thought (Lewis 1986, p. 213). That is correct as far as it goes. But consider the difference between that composite and a hurricane. Knowing a hurricane's pressure organization and rotational structure tells you what it will do next without tracking every molecule. The shoe-moon composite carries no comparable transition structure; predicting its future requires independently tracking each component. What separates them is not existence but autonomous macro-dynamics. This paper argues that closure of transition structure is the criterion that marks that difference.

The thesis, then, is that real-pattern realism needs an explicit closure condition to avoid permissive drift.

For a fixed regime, horizon, and admissible intervention class, a candidate macro-object qualifies when macrostate information is sufficient for macro-transitions, so within-class micro-differences do not change macro-level what-follows.

This thesis is a tightening move, not a replacement project. Existing pattern realism captures compression and projectibility insights. The present contribution adds a discriminating condition that excludes gerrymandered candidates by transition structure rather than by intuitive naturalness. The formal benchmark for this condition is strong lumpability in exact Markov settings. The non-ideal extension is graded closure via leakiness and convergent diagnostics under fixed constraints.

The argument is philosophical, not a methods paper in disguise. The paper does not claim to deliver a universal estimation recipe. It provides a criterion and a disciplined protocol for applying it. Estimator selection, finite-sample behavior, and domain-specific implementation remain downstream methodological tasks.

1.1 Novelty and Positioning

The novelty claim has three parts.

1. Beyond Dennett alone: compression realism is retained, but permissiveness is reduced by an explicit anti-gerrymandering condition.
2. Beyond formal closure results alone: strong lumpability is used as a benchmark for an ontological criterion, not only as a mathematical property.
3. Beyond pure interventionist pragmatism: admissibility is fixed upstream and physically constrained, so verdicts are not back-fit to analyst preference.

This is a conditional metaphysical proposal. Under structural realist and interventionist commitments, closure under admissible conditions is sufficient for macro-objecthood in regime. Readers with different priors can still accept a narrower conclusion: closure is at least a necessary anti-gerrymandering constraint on serious macro-ontology claims.

This positioning is deliberately dialogical. With Dennett, the paper keeps compression realism while rejecting permissiveness (Dennett 1991). With Ladyman and Ross, it keeps structural and projectibility ambitions while insisting on an explicit transition criterion for objecthood verdicts (Ladyman and Ross 2007). With Rosas, it treats formal closure diagnostics as benchmark machinery, then uses them to support a philosophical criterion that remains stable in non-ideal cases (Rosas et al. 2024). With causal-emergence work, it treats macro-level gains as corroborating diagnostics rather than as a standalone ontology test (Hoel et al. 2013). With Kim-style exclusion pressure, it defends non-redundancy at explanatory and control levels without denying microphysical implementation (Kim 1998).

A related line of work traces the cognitive origins of real-pattern commitments, linking predictive coding and free-energy minimization to how bounded agents build and revise their representational schemes (Gładziejewski 2025). That work helps explain why agents converge on some macro-descriptions and abandon others, which strengthens the anti-arbitrariness side of pattern realism. But representational success can remain task-relative unless a further test determines when a candidate grouping tracks dynamical structure in the world rather than merely serving local inferential needs. The closure criterion supplies that additional test.

A stronger reading of Dennett is worth confronting directly. One can interpret real-pattern realism as already requiring projectibility and cross-context robustness, not merely compression in a single dataset. On that steelman, Dennett's criterion already excludes fragile, narrowly-tuned codings. Even so, projectibility constrains which predictions are worth tracking; it does not test whether the macro-description carries autonomous transition structure. A partition can be projectible across contexts and still require persistent within-class micro-bookkeeping to sustain its forecasts, because the macro-label groups together microstates whose onward profiles diverge

once perturbation or horizon extension is applied. Closure catches exactly that residual failure mode. It asks not only whether a pattern projects, but whether macro-level what-follows is self-contained at the declared grain. That is the gap even the most charitable reading of Dennett leaves open.

In short, projectibility can still tolerate successful forecasting that depends on recurring micro-level repair. Closure does not.

1.2 Scope and Non-Claims

The paper is restricted to induced closure in spatiotemporal systems. It does not offer a complete metaphysics of levels, and it does not settle all questions about abstract objects or full mereology. Where those debates arise, they are handled only to protect the central claim from predictable misreadings. Still, the criterion is stated over state-transition structure: spatiotemporal dynamics are the paradigm case, not the only intelligible format in which closure questions can be posed.

Methodologically, the paper does not depend on the strongest anti-analytic rhetoric sometimes associated with naturalized metaphysics. The claim is narrower: closure is a usable philosophical constraint that can be motivated by dynamical and coarse-graining practice, whether or not one accepts every meta-metaphysical commitment in the broader ETMG program (Ladyman and Ross 2007).

The roadmap is direct. Section 2 states the criterion and admissibility framework. Section 3 gives the exact benchmark. Section 4 extends the criterion to non-ideal settings. Section 5 addresses the strongest objections. Section 6 summarizes gains and limits.

1.3 Three Nearby Positions

It helps to separate three nearby stances that are often run together.

1. Compression-prediction realism: successful compression and forecasting are treated as sufficient for realist commitment.
2. Pure pragmatism: model choice is guided by utility alone, with no ontological consequence.
3. Closure realism: compression and prediction matter, but objecthood requires transition autonomy under fixed regime, horizon, and admissible intervention class.

The third stance preserves the explanatory strengths of the first while avoiding its permissiveness. It also avoids reducing ontology to convenience, which is the pressure on the second.

1.4 Contribution Map

The paper makes three linked contributions.

1. It gives a criterion-level tightening of real-pattern realism by adding closure as a discriminating objecthood condition.
2. It provides a bridge from exact benchmark cases to non-ideal cases without changing criterion content.

3. It offers a disciplined verdict structure with explicit downgrade conditions, so realism claims remain answerable to failure.

The central payoff is selective commitment. The framework can support strong commitments where transition autonomy is robust, while avoiding forced commitments where evidence is fragile or regime-sensitive.

2 Closure and Admissibility

2.1 Criterion in Plain Language

Closure can be stated without heavy formalism. A candidate macro-object passes when knowing its current macrostate is enough to determine macro-level what-follows over a specified horizon and intervention class. If two microstates inside one macrostate produce different macro-transitions, closure fails at that grain.

The key point is dynamic. The criterion concerns transition sufficiency, not mere descriptive fit. A representation can summarize trajectories elegantly and still fail objecthood if it requires persistent within-class micro-bookkeeping to preserve forecast quality.

2.2 Formal Statement

Let micro-process be X_t and candidate macro-process be $Z_t = g(X_t)$. For horizon L, closure asks whether Z_t is sufficient for forecasting Z_{t+1}^{L} under a fixed regime and admissible intervention class.

The horizon object here is the L-step path distribution, not only one-step prediction. One-step tests can be used as diagnostics, but objecthood claims are indexed to the declared horizon target.

Informationally, the predictive side asks whether adding X_t beyond Z_t yields material gain for macro-future prediction. Interventionally, the causal side asks whether macro-transition laws remain stable across admissible perturbations without needing within-class micro-identification.

These are not competing tests. They are two readings of the same target: transition autonomy at the macro level.

2.3 Why This is More Than Compression

Compression is necessary but not sufficient for objecthood. A compressed code can hide transition-relevant heterogeneity and still score well on narrow tasks. Closure blocks that loophole by asking whether hidden heterogeneity matters for macro-level what-follows.

This is why closure functions as an anti-gerrymandering condition. It does not reward convenient coding alone. It rewards coarse-grainings that carry stable transition structure under declared constraints.

Put differently, this paper accepts Dennett's claim that compression success is evidence of objective patterning, but denies that compression success is the full ontological test (Dennett 1991). It also accepts Ladyman and Ross's concern that real

patterns should be projectible and structurally disciplined, while adding an explicit transition criterion for adjudicating borderline cases (Ladyman and Ross 2007). The aim is not to replace those insights. The aim is to make their realist force less permissive.

This also answers a familiar Haugeland-style pressure point. The paper does not infer objecthood directly from "there is a useful pattern in the data." It adds an explicit transition-autonomy condition that determines when pattern talk has ontological force rather than merely descriptive convenience (Haugeland 1998).

2.4 Admissibility and Hierarchical Evaluation

Closure is always relative to an intervention class. The standard objection is that this invites circularity. The response is procedural and hierarchical. Interventions are the empirical case of a broader admissible-transformation idea: what matters is fixing a defensible class of allowed manipulations before closure is scored.

Admissibility is fixed upstream of objecthood verdicts by three constraints:

1. Epistemic admissibility: interventions are measurable and implementable by bounded agents.
2. Dynamical admissibility: interventions preserve the target regime class.
3. Explanatory admissibility: interventions target variables with cross-context generalizability and non-trivial counterfactual reach, rather than one-off manipulations tuned to a single episode.

These constraints are physically anchored in available control channels and implementation structure. They are not analyst preferences about favored ontologies. The explanatory constraint is set before closure verdicts and does not assume in advance which candidate partition will prove stable. This makes the interventionist notion of a "possible experiment" concrete and bounded, rather than leaving it at conceptual possibility where nearly any hypothetical manipulation might count (Woodward 2015). This aligns with Woodward (2021)'s argument that correct variable choice is a matter of satisfying "proportionality"—minimizing the omission of dependency relations. The procedure constrains judgment without eliminating it: domain expertise still shapes regime identification and control-channel selection, but that shaping is declared, auditable, and separated from partition scoring.

Operationally, explanatory admissibility can be audited without looking at which partition wins. A candidate intervention class is stronger when it is implemented through control channels that already exist in the domain and are replicable across sites, operators, or runs. It is weaker when success depends on narrow, one-off tuning that fails under small, feasible perturbations of the same intervention family.

A related distinction matters here. Pearl defends the $do(x)$ operator as an ideal "surgery" that can be defined mathematically regardless of physical feasibility (Pearl 2000). For causal inference, that idealization is productive: it lets one reason about effects even when direct manipulation is unavailable. But macro-objecthood requires tighter constraints than logical derivation. A partition that achieves closure only under physically unrealizable interventions, ones that no bounded agent could implement

without destroying the target regime, does not support the kind of stable, intervention-guiding structure that warrants realist commitment. The admissibility constraints above ensure that closure is assessed relative to control channels that are actually available in the regime, not relative to arbitrary mathematical surgeries. This is not a rejection of Pearl's framework. It is a restriction on which of its outputs carry ontological weight.

Evaluation order matters.

1. Fix regime, horizon, admissible intervention class, diagnostics, and model class.
2. Compare candidate partitions under those fixed constraints.
3. If constraints are revised after inspecting results, restart and disclose the revision.

This order blocks back-fitting and makes disagreement tractable. If verdicts remain highly sensitive across nearby defensible admissibility specifications, commitment should be downgraded. Stated as a discipline, admissibility has two stages: first, intervention classes are fixed from independent physical constraints and established control practice in the target domain; second, closure is evaluated only relative to that fixed class.

A structural worry deserves explicit treatment here. If regime individuation already commits to which macro-variables matter, the protocol risks a version of circularity: closure testing presupposes regime identification, but regime identification might presuppose partition outcomes. The response is that defensible regime specifications draw on quantities characterizable without presupposing which candidate partition will succeed. Temporal parameters, energy scales, flow rates, and enforcement intensities are typically identifiable through physical indicators that competing candidate partitions can agree on before either is evaluated. In the traffic illustration, "weekday commuter traffic in one metropolitan corridor" is characterized by clock schedules, total vehicle counts, and road network topology, all quantities that investigators favoring either the lane-flow or vehicle-ID-cluster partition can agree on before either closure test is run. Regime specifications of this kind are anchored in lower-level or cross-cutting physical indicators rather than in the outcome of the very test they enable. The key principle is that regime specifications should be grain-neutral across competing candidate partitions: a regime can itself be a macro-level description while still not presupposing which finer-grained partition will close within it.

The independence is imperfect in some domains, particularly biology and social science, where macro-variable choice and regime characterization can be genuinely entangled. The framework treats that entanglement as a source of evidential weakness rather than a concealed defect. When alternative, independently motivated regime specifications produce divergent closure verdicts, the verdict ceiling should be downgraded to at most qualified. More broadly, regime individuation and partition evaluation should be understood as iterative rather than viciously circular: initial regime characterizations are provisional, closure testing can refine them, and the process stabilizes into reflective equilibrium, in the sense Goodman articulated: rules of inference and particular inferential judgments are justified by their mutual coherence rather than by either serving as the unilateral foundation for the other (Goodman 1955). If iteration does not stabilize across nearby defensible starting specifications,

the result should be treated as indeterminate for that regime rather than forced into a robust verdict. Here stabilization means that comparative ordering across independently specified candidate partitions stops shifting under small, defensible protocol variations; it does not presuppose one privileged grain in advance. The discipline comes from requiring that all revisions be disclosed and that each revision restart the evaluation rather than silently inheriting prior results.

2.5 Pattern Reality Versus Macro-Objecthood

One distinction is essential for avoiding confusion. A pattern can be real in a weaker sense without satisfying this paper's objecthood criterion. A representation can capture regularities that are descriptively useful while still failing closure under admissible interventions.

The criterion in this paper is stricter. It targets macro-objecthood, not any and every projectible summary. This is why the paper can preserve a generous attitude toward pattern detection while denying ontological standing to high-leak candidates.

The same distinction also clarifies what this framework contributes to composition debates. In van Inwagen's terms, the standing question is when some things compose a further thing (van Inwagen 1990). The present proposal does not offer a full mereology. It gives a domain-indexed discipline condition: composition claims about macro-objects are warranted when transition autonomy is robust under declared constraints, and should be withheld when candidate sums remain transition-fragile.

That distinction clarifies dialectical burden. The paper does not need to show that non-closed representations are worthless. It only needs to show that they do not meet the objecthood standard defended here.

2.6 Admissibility Disputes

Admissibility disputes are expected, especially across domains. The adjudication rule is to compare candidate intervention classes by what they physically permit and whether they preserve the same target regime. Boundary-pressure perturbations can be admissible for storm dynamics, while molecule-by-molecule remote rewriting is not. In institutional settings, changing enforcement intensity can be admissible, while instant arbitrary rewriting of all agent commitments is not. The ontological consequence is direct. A storm satisfies closure under boundary perturbations because its pressure organization and rotational structure determine macro-transitions without tracking individual molecules. Under molecule-by-molecule rewriting, the same candidate fails, because admissibility now permits interventions that reach inside the macroclass and exploit within-class differences. The candidate has not changed; the admissible control channel has, and the verdict changes with it. This is why admissibility must be declared and justified before scoring, not adjusted afterward to protect a preferred outcome.

For transparency, three template classes are often useful as a starting grid: boundary nudges, coarse actuator controls, and policy levers. These templates are not universal, but they make admissibility comparison public and repeatable.

One guardrail is non-negotiable. Interventions that directly target preservation of partition labels, rather than system variables, are inadmissible because they trivialize closure by design. The test concerns whether transition structure is autonomous under physically meaningful controls, not whether labels can be protected by stipulation.

When two intervention classes are both admissible and target the same regime, the framework permits plural testing rather than forced monism. Verdicts should report which admissibility class was used and how sensitive results are across nearby defensible classes. A candidate partition that appears closed only under one narrowly tuned class should be downgraded when sensitivity appears across nearby defensible alternatives.

2.7 Canonical Criterion Statement

For ease of reference, the core criterion can be stated compactly:

A candidate coarse-graining $Z = g(X)$ qualifies for macro-objecthood in regime when, under fixed horizon L and admissible intervention class $\mathcal{I}$, macro-transition structure is autonomous up to declared tolerance, so within-class micro-differences do not materially improve macro-future prediction or intervention-guided control.

Candidate partitions must be specified independently of the particular sample trajectory used to score closure, for instance by a pre-declared construction rule or learning objective fixed before evaluation. Otherwise closure collapses into bespoke encoding, where any dataset can be made to look autonomous by post hoc recoding.

In exact Markov settings, strong lumpability is a sufficient benchmark realization of this criterion. In non-ideal settings, leakiness-centered diagnostics estimate distance from that ideal under fixed constraints.

2.8 Predictive and Interventional Closure

Predictive and interventional closure should be distinguished but not separated. Predictive closure concerns whether macrostate information screens off transition-relevant micro-detail for macro-future forecasting. Interventional closure concerns whether macro-transition structure remains stable under admissible perturbation.

The evidential relation is asymmetric. Strong interventional closure typically implies predictive adequacy for the same target and horizon, while predictive adequacy alone can persist in cases where interventional stability fails under distribution shift. This is why the criterion treats predictive evidence as important but not final.

This asymmetry also clarifies burden of proof. Claims to robust macro-objecthood need either direct interventional support or compelling indirect evidence that tracks intervention-relevant invariance. Otherwise, the responsible verdict is qualified or indeterminate.

2.9 Why Closure Carries Ontological Weight

A reviewer can still ask why transition autonomy should count as objecthood rather than as a mere success condition for modeling. The answer depends on the paper's explicit commitments. Under structural realism, what ontology should track is stable relational and causal organization rather than intrinsic micro-identity as such. Under

interventionism, what counts as causally relevant structure is structure that supports stable manipulation and control.

The meta-philosophical stance here follows what Woodward calls the methodological path to ontology: causal-interventionist tools are used to determine what earns realist commitment, rather than seeking independent metaphysical "truth-makers" or "grounds" that stand apart from the practices of manipulation and prediction (Woodward 2015). This is not a retreat to instrumentalism. It is the claim that scientific ontology, disciplined by closure, is the only ontology required for macro-objecthood verdicts. The demand for a further metaphysical foundation beyond stable dynamical autonomy under admissible intervention mistakes a philosophical habit for a substantive requirement.

Given those commitments, closure does not function as a convenient heuristic layered on top of ontology. It is the criterion that identifies when a candidate macro-description tracks stable structure at the level where explanation and intervention are being assessed. This is why the framework can remain compatible with microphysical completeness while still defending non-redundant macro commitment.

The connection to Pearl's notion of modularity is worth noting. Pearl argues that causal models work because nature consists of autonomous mechanisms: changing one structural equation need not disrupt the others (Pearl 2000). Closure is related to this property but not identical to it. Pearl treats modularity as a structural assumption about a given causal model; the present criterion asks a prior question, namely whether a candidate coarse-graining has earned the right to appear as a variable in such a model at all. A partition that requires persistent micro-bookkeeping to preserve its transition structure has not isolated an autonomous mechanism at the macro grain in the relevant sense. So closure can be understood as an ontological precondition for the kind of modularity Pearl's framework presupposes, rather than as a restatement of it.

This is also why the proposal is not vulnerable to standard "bare structure" worries. The criterion is modal and dynamical, not merely extensional. It is about counterfactual transition organization under interventions, not about abstract isomorphism alone.

This also addresses Newman-style triviality pressure in structural realism, where purely extensional structure can be cheap under redescription (Ainsworth 2009, pp. 141–145). Closure is not extensional fit alone. It requires transition and intervention stability under declared constraints, which arbitrary recodings typically fail. When apparent success depends on privileged representational accidents, that encoding dependence is itself evidence of leakiness.

Why not stop at "good variable" language? Because a good variable can be merely convenient, and the worry is that closure only tracks inferential convenience rather than what is objectively there. The framework rejects that deflation because closure claims are indexed to intervention classes, not only to passive prediction. A representation can score well on retrospective fit by exploiting accidental correlations and still fail quickly when admissible interventions shift transition pathways. A closure-supporting partition survives those shifts because the screened-off structure is not accidental. If a candidate supports reliable counterfactual control and regime-stable transition laws, it is tracking objective modal structure: stable counterfactual

dependencies and lawlike transition regularities under admissible intervention. On the paper's commitments, treating that structure as merely representational would undercut the very realism the framework is designed to preserve. The modal profile, not the in-sample fit, carries the ontological burden. Predictive sufficiency can reveal compression, but interventional stability identifies difference-makers: it tests whether the same macro-transition structure survives admissible manipulation rather than merely fitting observed trajectories.

The framework therefore distinguishes epistemic humility from ontological deflation. Finite agents will often have partial, noisy, and regime-limited evidence. But that uncertainty does not collapse the difference between dynamically autonomous and dynamically incoherent partitions. That difference can be difficult to estimate, but it is still a difference in the world rather than in preference.

2.10 Levels of Claim and Representation

Some recurring misunderstandings come from sliding between different levels of claim. The paper uses the following distinction throughout.

1. World dynamics: the implemented process itself.
2. Pattern type: stable transition structure supported by that process.
3. Pattern token: a concrete instance under specific boundary conditions.
4. Representation: a model, equation, classifier, or narrative that tracks the pattern.

A representation can fail while the pattern type remains real. A token can fail while the type remains robust. A token can also succeed while the type is weak, for example in narrow calibration windows. Closure claims in this paper are primarily type-level claims under specified regimes, then secondarily claims about token reliability under perturbation.

This is why the criterion does not equate model performance with ontology. It asks whether the represented transition structure is genuinely autonomous, not whether one representation currently performs well.

2.11 Closure, Underdetermination, and Objecthood Discipline

One remaining concern is underdetermination. Even after the criterion is stated, a reviewer may argue that many distinct coarse-grainings can be tuned to look acceptable on available data. If so, closure might seem to collapse back into model choice rather than objecthood discipline.

The framework's answer is to separate three questions that are often conflated.

1. Which candidate partitions are descriptively serviceable in a dataset?
2. Which candidates remain transition-autonomous under admissible perturbation?
3. Which of those candidates remain stable under modest horizon and admissibility variation?

Underdetermination is strongest at the first question and weaker at the second. It is often weakest at the third. Many candidates can fit observed trajectories. Far fewer preserve counterfactual structure when the regime is probed. Fewer still remain

stable across nearby defensible setups. This staged filtering is exactly where closure earns its metaphysical role.

The framework does not claim that every domain yields one uniquely privileged partition. It claims that objecthood commitments should be constrained by transition autonomy and robustness, rather than by fit alone. When underdetermination persists after those filters, the responsible result is qualified or plural commitment under transparent constraints, not forced monism and not unconstrained relativism.

Before moving to the formal benchmark, the framework can be summarized in three lines.

1. **Criterion:** macro-objecthood requires transition autonomy under fixed regime, horizon, and admissible intervention class.
2. **Evaluation:** when exact closure fails, leakiness-centered diagnostics estimate comparative distance from closure.
3. **Commitment rule:** commitment strength tracks evidential stability, with explicit downgrade to qualified or indeterminate status when evidence is unstable.

3 Exact Benchmark: Strong Lumpability

Strong lumpability is introduced as a benchmark, not as destiny. It gives a clean exact case in Markov settings where closure can be stated and checked without ambiguity.

Let partition cells under g be macroclasses. Strong lumpability holds when microstates within the same macroclass induce identical transition probabilities to all macroclasses (Kemeny and Snell 1960). When this condition holds, macro-transitions are autonomous by construction.

More precisely: for any two microstates x, x' in macroclass C_i, and any macroclass C_j, strong lumpability requires $\sum_{y \in C_j} P(y \mid x) = \sum_{y \in C_j} P(y \mid x')$.

Lemma. If the micro-process is first-order Markov and the partition induced by g is strongly lumpable, then the induced macro-process is itself Markov and transition-autonomous at the macro level.

Proof sketch. Define a macro-kernel by summing micro-transition probabilities from any representative $x \in C_i$ into each macroclass C_j. Strong lumpability guarantees this sum is representative-independent. The macro-kernel is therefore well-defined, and the induced process over macroclasses is Markov with transitions given by that kernel.

The philosophical significance is straightforward. A partition that satisfies this condition is not merely useful. It preserves transition structure at the macro grain. A partition that fails it mixes transition-heterogeneous microstates and therefore lacks macro autonomy.

Equivalent language from computational mechanics clarifies why this is not superficial bookkeeping. For any macro-process Z, two prediction machines can be defined. The ε-machine is the minimal model that predicts Z's future from Z's own past, using only macro-level information. The υ-machine is the minimal model that predicts Z's future from the full micro-past X, using everything available (Shalizi and Crutchfield 2001; Rosas et al. 2024). Closure holds when these two machines are equivalent: the ε-machine already captures everything the υ-machine knows about macro-futures, and

extra micro-information is redundant for the macro target. This gives an exact ideal where closure is not approximate.

The important limitation is equally clear. Exact strong lumpability is uncommon in open, noisy, and path-dependent systems. That limitation motivates the graded extension. It does not undermine the criterion.

3.1 Minimal Formal Illustration

The anti-gerrymandering role can be shown with a minimal symbolic example. Let microstates be $\{x_1, x_2, x_3, x_4\}$ with transition rows:

$$P(x_1, \cdot) = (\alpha, \beta, 0, 0),$$
$$P(x_2, \cdot) = (\alpha, \beta, 0, 0),$$
$$P(x_3, \cdot) = (0, 0, \gamma, \delta),$$
$$P(x_4, \cdot) = (0, 0, \gamma, \delta),$$

with $\alpha + \beta = 1$ and $\gamma + \delta = 1$.

Partition A groups $\{x_1, x_2\}$ and $\{x_3, x_4\}$. Partition B groups $\{x_1, x_3\}$ and $\{x_2, x_4\}$. In A, members of each macroclass have matching onward profiles to macroclasses, so macro-transitions are autonomous. In B, members of one macroclass differ in onward profiles whenever $\alpha \neq \gamma$, so the macro-label hides a transition-relevant difference.

Partition B can still be made predictive by adding repeated within-class bookkeeping. That is exactly the high-maintenance case this paper excludes from robust macro-objecthood. On a fixed sample trajectory, this bookkeeping can make Partition B look nearly as predictive as Partition A. The difference appears under admissible perturbation: A preserves macro-transition structure without renewed microstate tracking, whereas B does not.

The philosophical lesson is simple. Both partitions are definable. Only one preserves transition autonomy. Closure therefore discriminates between legitimate macro-candidates and merely codable aggregates.

3.2 If an Unusual Partition Closes

A common reaction is that a strange disjunctive partition might satisfy the formal condition in a symmetric system. That is correct. On this framework, if such a partition genuinely supports autonomous macro-transitions under fixed constraints, it counts as real at that grain.

This is not a concession to arbitrariness. The criterion tracks transition coherence, not intuitive naturalness. If one wants a separate naturalness filter, that is an additional criterion and should be declared as such.

The exclusion claim should therefore be read carefully. The framework excludes dynamically incoherent, high-leak aggregates. It does not exclude every unusual grouping.

3.3 What the Formal Machinery Alone Does Not Provide

The mathematical property and the philosophical criterion are different things. Four elements of the present framework have no counterpart in the formal literature alone.

First, intervention indexing. Causal-state constructions classify histories by equivalence of future distributions under the observed process, which is primarily an observational-predictive equivalence relation (Shalizi and Crutchfield 2001). The present criterion adds an admissible intervention class as a parameter of the closure test. Objecthood claims are indexed to what the system does under perturbation, not only under passive observation. This is what separates the criterion from a purely statistical diagnostic. Rosas et al.'s framework distinguishes informational, causal, and computational closure as complementary formal diagnostics (Rosas et al. 2024). The present paper treats predictive sufficiency as evidential, but ties objecthood commitment to admissible-intervention stability and to the absence of hidden micro-identity bookkeeping.

Second, an explicit ontological interpretation. Strong lumpability tells you when a coarse-graining preserves transition structure. It does not tell you whether that preservation warrants realist commitment. The philosophical work is to argue that, under stated commitments, transition autonomy under admissible interventions is sufficient for macro-objecthood, and to give the conditions under which that claim should be downgraded or withdrawn.

Third, a graded commitment protocol. The formal literature offers exact conditions and, more recently, approximate-closure measures. It does not supply a framework for translating those measures into warranted ontological verdicts with explicit robustness requirements and downgrade rules. The selective commitment structure (robust, qualified, indeterminate) is a philosophical addition.

Fourth, an anti-gerrymandering argument. The argument that closure is the right anti-gerrymandering condition, and that compression without closure is insufficient for objecthood, because compression can be won by codings that effectively preserve hidden micro-identities, is not a theorem. It is a philosophical claim defended by the structure of Sections 2 through 4.

A similar point applies to causal-emergence work. Effective-information analyses identify when macro-descriptions gain determinism relative to micro-descriptions (Hoel et al. 2013). That is a valuable diagnostic, but a partition can increase effective information while remaining fragile under intervention or horizon variation. The present framework treats such gains as corroborating evidence within a closure-governed criterion, not as a standalone ontology test.

So the paper is not computational mechanics or causal-emergence analysis with new labels. It is a philosophical criterion that borrows formal tools while adding intervention-indexed objecthood conditions, an explicit commitment protocol, and a sustained argument about why closure is the right fix for pattern-realism's permissiveness problem.

4 Approximate Closure Without Instrumentalist Drift

4.1 Why Approximation is Expected

In complex systems, boundaries leak and couplings shift across regimes. Exact closure is therefore unusual. A credible realism criterion cannot require perfect closure everywhere. The mathematical study of approximate aggregation has a long history, including analyses of nearly decomposable dynamic systems that separate short-run within-subsystem dynamics from long-run aggregate behavior (Ando and Simon 1961, pp. 111–117). The present extension is philosophical rather than technical. It uses graded closure to support graded ontological verdicts, not to improve estimation.

Approximation here is not concession to arbitrariness. It is the expected non-ideal form of the same criterion, provided constraints are fixed and diagnostics are comparative.

The right ontology test is type-level before token-level. A pattern type can be robustly closed for a regime even when a specific token fails because of boundary violation, atypical perturbation, or timescale mismatch. One anomalous token is therefore not decisive. The relevant question is whether failures are exceptional or systematic for the type under the stated constraints. Systematic failure, where within-class micro-differences repeatedly matter for macro-transitions across tokens under fixed constraints, is what triggers type-level downgrade rather than token-level exception.

4.2 Leakiness as Canonical Target

Leakiness measures how much within-class micro-detail still improves macro-future prediction once current macrostate is fixed. A canonical quantity is the conditional mutual information $I(X_t; Z_{t+1} \mid Z_t)$ (Cover and Thomas 2006, eqs. (2.60)–(2.61)). Low values support closure. High values indicate hidden transition-relevant heterogeneity.

In this paper, leakiness is the default target quantity. Other diagnostics, such as within-class transition divergence and predictive gain from added micro-features, function as estimators or proxies when direct estimation is limited.

Leakiness should therefore be read comparatively and through robustness checks, not as a context-free absolute cutoff. Absolute tolerance is domain-relative, but high sensitivity of leakiness rankings to modest defensible modeling choices is itself a downgrade trigger.

Comparative ranking is necessary but not sufficient. A candidate can beat nearby alternatives and still fail objecthood if absolute leakiness remains high and instability persists under modest robustness checks. The framework therefore requires both relative superiority and minimum viability.

The viability floor can be stated without a numeric threshold. A partition fails minimum viability if, across the declared robustness checks, within-class micro-features yield consistent, non-negligible gains in macro-future prediction or intervention response relative to the best macro-only model. When that pattern persists across diagnostics and perturbation tests, the partition has not achieved the transition autonomy the criterion requires, regardless of how it ranks against competitors.

The term "non-negligible" can be given a comparative anchor without fixing a universal numeric threshold. Leakiness is most clearly non-negligible when adding within-class micro-features yields prediction gains comparable in magnitude to the gains the macro-partition itself provides over a naive baseline model. If the macro-partition delivers substantial predictive improvement over no partition at all, and within-class micro-features add only marginal further gain, that residual leakiness is negligible for most objecthood purposes. If micro-features instead add gains approaching or matching the macro-partition's own contribution, the partition is not doing most of the relevant work, and minimum viability is not met. This framing makes the judgment about a single tractable quantity, namely relative predictive contribution, rather than an abstract absolute cutoff, and it makes cross-domain comparisons more tractable. The two requirements work in sequence: the viability floor is applied first as a minimum gate, excluding candidates regardless of how they compare to rivals, while comparative ranking then assigns verdict categories among candidates that pass it. To keep this anchor non-discretionary, the baseline model class and gain metric must be fixed upstream with regime and admissibility, and post hoc revisions trigger restart and disclosure.

Verdict-category boundaries carry genuine vagueness because the underlying property of transition autonomy is itself graded. This is a feature rather than a defect: biological fitness is real and explanatorily significant without admitting a sharp boundary between fit and unfit. What matters primarily is ordering stability. When partition A consistently shows lower leakiness and greater perturbation stability than partition B across diagnostics, that ordering should be robust even if the exact placement of the boundary between "qualified" and "robust" carries some domain-conventional element. A partition that dominates alternatives across diagnostics and remains stable under modest changes to horizon, intervention distribution, and robustness checks earns the stronger verdict regardless of where one draws the categorical line.

4.3 Procedural Safeguards in Non-Ideal Cases

Approximate closure claims require explicit safeguards.

1. Fix regime, horizon, admissibility class, and diagnostics before comparative scoring.
2. Compare candidate partitions under the same fixed setup.
3. Test robustness under modest changes in horizon and intervention distribution.
4. Apply a minimum-viability floor: if all candidates remain high-leak and fragile, return no robust objecthood verdict for that regime.
5. Report verdicts as robust, qualified, or indeterminate.

This keeps threshold choice procedural rather than discretionary. It also addresses a predictable confusion. Estimation difficulty is an epistemic limit on access, not a defect in the metaphysical criterion itself.

Selection discipline. One concern from both reviewers and readers is that flexibility can silently re-enter through upstream choices. The framework answers this with a single discipline rule: admissibility, state construction, horizon, and diagnostics are fixed before scoring, and post hoc revisions trigger restart and disclosure. This does

not eliminate judgment. It makes judgment auditable and prevents favored partitions from being protected by moving criteria.

4.4 Regime Dependence and Projectibility

Regime dependence does not imply observer-relativity. It states that closure depends on actual transition structure under specified constraints. A pattern can be closed in one regime and leaky in another because the structure has changed.

The metaphysical stance is explicit: regime and horizon index the modal profile being claimed, not a concession that anything goes.

Projectibility provides a further check. Defensible regime specifications should support stable induction under modest counterfactual variation. Narrowly engineered regimes that protect a favored partition usually fail under small context shifts. When that occurs, the candidate was never robustly closed in the relevant sense.

This is where the paper is closest to Ladyman and Ross. Projectibility is not treated as an optional methodological virtue. It functions as a realism-relevant stress test on whether a closure claim survives beyond calibration conditions (Ladyman and Ross 2007). Closure without projectible robustness is too cheap to support robust objecthood claims.

Regime dependence itself should be split into two forms. Ontic regime dependence occurs when system structure changes, such as phase transitions or boundary-condition shifts. Epistemic regime dependence occurs when measurement or control access changes while underlying structure remains fixed. The first changes what is there to be tracked. The second changes what can be warranted from available evidence.

4.5 One Distributed Illustration

Macro does not mean large or spatially contiguous. Coarse-graining can be logical and distributed. Monetary systems illustrate this point without requiring new machinery.

The macro-transition structure of transactions can remain stable across heterogeneous micro-realizations, such as cash tokens, ledger records, and digital balances, under admissible legal and financial interventions. If that stability holds, macro-objecthood is warranted in regime. If implementation channels degrade, closure can fail even while symbolic representations persist.

A minimal failure case clarifies the point. A state can retain formal legal tender rules while losing reliable payment enforcement and settlement implementation. In that case, the representation persists but transition autonomy degrades, and closure-based commitment should be downgraded.

Organisms give a complementary illustration. In many physiological regimes, membrane and regulatory organization screen off large amounts of molecular variation for the target transitions, so intervention on organ-level variables remains predictively and manipulatively effective. When those screening structures fail, for instance under severe systemic breakdown, the same macro-description can become leaky and require finer-grained tracking. The point is not that organism talk is always closed. The point is that closure can be regime-robust without being regime-universal.

Across these illustrations, the verdict structure does real discriminative work. The monetary example in a well-functioning system approaches robust objecthood in the declared regime: macro-transition structure remains stable across heterogeneous realizations and admissible interventions. The organism example falls into the qualified range: within established physiological regimes, membrane organization and regulatory control screen off large amounts of molecular variation, making organ-level interventions reliably effective, but leakage appears at ecological or pathological boundaries where those screening structures degrade. The failed-implementation case represents a downgrade: formal representations persist while transition autonomy erodes, warranting withdrawal of robust commitment. The three cases illustrate the verdict categories in operation, not as abstract labels but as outcomes determined by the same diagnostic criteria applied to different evidential situations.

The illustration does one job only. It shows that closure tracks transition structure, not spatial shape. Nothing in the argument depends on taking a stance on broader social ontology.

4.6 Failure Conditions and Downgrade Rules

The framework should also say clearly when commitment should be withdrawn or downgraded. Four failure patterns are especially relevant.

1. Persistent high leakiness across defensible diagnostics under fixed constraints.
2. Strong disagreement among diagnostics that does not resolve under modest robustness checks.
3. High sensitivity of verdicts to small, defensible changes in admissibility or horizon.
4. Candidate sets where every partition remains above the minimum-viability floor.

When these patterns appear, the right response is not forced binary judgment. The right response is explicit downgrade to qualified or indeterminate status. This keeps the criterion resilient without pretending that every case must produce a sharp ontology verdict.

Conversely, when closure persists across distinct, independently motivated admissibility classes within the same regime, that cross-admissibility stability raises evidential confidence for robust commitment.

Type-token reminder: a downgraded token does not automatically refute type-level closure. The key question is whether instability is exceptional at token level or systematic for the type under the stated constraints.

4.7 Stable Versus Merely Entailed Patterns

The graded framework also supports an important distinction. A pattern can be entailed by microhistory and laws without being stable as a macro-handle. Entailment alone is cheap. Stability under admissible perturbation is demanding.

This distinction matters for permissiveness debates. A contrived disjunctive construction can be true of a realized trajectory while still failing closure. It can require continual within-class micro repair, fail under modest regime shifts, and lose interventional reliability.

A historical case makes the pattern concrete. Phlogiston theory compressed combustion phenomena under a single macro-variable, but the partition required persistent bookkeeping to survive: when metals gained weight upon burning, theorists introduced "negative phlogiston"; when different substances showed different weight changes, further ad hoc parameters appeared. Each patch was a new within-class distinction imported to preserve macro-prediction. By contrast, oxygen theory required no such ongoing repair: the macro-variable (oxidation state) screened off the relevant chemistry without persistent bookkeeping. The critique here is structural, not retrospective ridicule. A candidate that continually imports corrections to preserve macro-prediction behaves like a non-autonomous partition, and the closure framework detects this failure directly.

So the claim is not that gerrymandered constructions are false. The claim is that they usually fail to qualify as macro-objects. They may remain descriptions, but not object-level descriptions in the relevant regime. The same applies to computationally adequate models more broadly. A higher-level model can be useful for prediction or control without meeting closure standards. Computational adequacy means the model serves some purpose; closure requires that transition-relevant micro-differences are screened off in the specified regime. A computationally adequate but high-leak representation remains a valuable tool, but it does not automatically earn macro-object status.

4.8 Practical Qualification Under Sparse Intervention Access

In many domains, direct interventions are sparse, ethically constrained, or expensive. This does not invalidate the criterion. It changes evidential strategy. The evidential order is asymmetric: predictive success can raise confidence, but only interventional stability can underwrite objecthood commitment.

Where intervention data are limited, predictive closure functions as an operational indicator while interventional closure remains the ontological target. A model can fit historical trajectories and still fail under novel perturbation. For that reason, claims should be qualified by available intervention access and downgraded when out-of-regime behavior is unstable.

The paper therefore does not treat in-sample prediction as decisive. The standard remains robustness under admissible perturbation, even when that robustness must be assessed indirectly.

This point also explains why diagnostics should be triangulated. Predictive closure can look strong in-sample while interventional closure fails under admissible perturbation. Partial observability can hide within-class heterogeneity that later appears as instability. Nonstationarity can make a previously low-leak partition unstable outside its calibration window. Triangulation does not remove these risks, but it makes them visible.

4.9 Conceptual Contrast: Near-Tie Prediction, Different Ontology Verdicts

Two partitions can show similar short-horizon observational performance and still receive different closure verdicts. This is where the criterion does real philosophical work.

Suppose partition P_1 and partition P_2 produce comparable one-step observational fit in calibration data. Under fixed admissible interventions, P_1 preserves stable macro-transition structure while P_2 requires repeated within-class micro refinements to maintain performance. Observationally, they may look close. Structurally, they are not.

On the present framework, P_1 receives the stronger objecthood verdict because it remains transition-autonomous under the fixed constraints. P_2 can remain useful as a representation, but its dependence on recurring hidden repair counts against macro-object standing.

A concrete sketch helps. In traffic modeling, one partition may track density and flow by lane segment. A rival partition may track arbitrary vehicle-ID clusters that happen to fit one week of observations. In-sample one-step fit can be similar. Under admissible perturbations, such as ramp metering changes and speed-limit adjustments, the lane-flow partition remains stable while the ID-cluster partition becomes brittle. Under modest horizon extension, leakiness for the ID-cluster partition rises sharply. The verdict is then robust or qualified objecthood for the lane-flow partition and downgrade for the rival.

Mini-walkthrough under the reporting schema:

1. Target partitions: P_{lane} (lane-segment density and flow) versus P_{id} (vehicle-ID clusters).
2. Regime and horizon: weekday commuter traffic in one metropolitan corridor; horizon set to short-run control windows and one modest extension beyond calibration.
3. Admissibility class: control channels that traffic operators actually use in that regime (ramp metering and speed-limit controls), excluding interventions that require implausible vehicle-level actuation.
4. Diagnostics: predictive gain from within-class microfeatures, within-class transition-profile divergence, and intervention response invariance under admissible perturbations.
5. Verdict: P_{lane} remains comparatively stable and earns robust or at least qualified objecthood in the declared regime; P_{id} remains instrumentally useful in calibration but is downgraded once perturbation and modest horizon extension are applied.

4.10 Reporting and Evidential Discipline

To keep conclusions comparable across cases, closure assessments should report five items explicitly.

1. Target partition and rationale.
2. Regime and horizon specification.

3. Admissible intervention class and justification.
4. Diagnostics used and their agreement or disagreement.
5. Final verdict category: robust, qualified, or indeterminate.

This reporting structure is simple, but philosophically important. It prevents silent shifts in target, timescale, or admissibility from being mistaken for genuine ontological progress.

For consistent application, the assessment sequence should be explicit.

1. Specify candidate partitions for one target process.
2. Fix regime, horizon L, admissible intervention class $\mathcal{I}$, diagnostics, and model classes in advance.
3. Evaluate all candidates comparatively under that same fixed setup.
4. Report verdict category and sensitivity under modest perturbation checks.

This sequence is not an optional implementation preference. It is part of what makes closure claims epistemically disciplined rather than post hoc.

Diagnostics do not replace criterion-level argument. They provide evidence about whether a candidate satisfies the criterion under declared constraints. This distinction matters because reviewers can otherwise misread the framework as reducing ontology to whichever metric happens to be available. The evidential logic is comparative and convergent. No single proxy is treated as infallible. Confidence increases when different diagnostics track the same rank-order among candidate partitions and when those rankings remain stable under modest perturbation tests. Confidence decreases when diagnostics diverge persistently or when rankings are brittle under small design changes.

This is why verdict categories are graded. Robust verdicts require convergence and stability. Qualified verdicts fit mixed but non-trivial evidence. Indeterminate verdicts fit persistent instability. The diagnostic set can vary by domain, but it should answer one fixed question: do within-class micro-differences still matter for macro-what-follows? In practice, three checks are usually enough.

1. Leakiness proxy based on predictive gain from added within-class micro-features.
2. Within-class transition-profile divergence.
3. Intervention-response invariance under admissible perturbation.

Agreement across these checks increases warrant. Persistent disagreement is evidence for revision or downgrade, not for forced commitment. When interventional evidence is sparse or unavailable, the remaining diagnostics can still support qualified status, but robust objecthood requires that the interventional leg not be entirely missing.

Minimal triangulation recipe: require stability under modest horizon variation and under at least two defensible admissibility perturbations, while at least two diagnostics agree on candidate rank-order. If this condition fails, return qualified or indeterminate status rather than robust objecthood.

5 Objections and Replies

The objections are organized as a progression from foundational worries to evidential-discipline worries. The order is deliberate, but the core answer remains stable: closure under fixed constraints supplies ontological discipline without demanding a single privileged descriptive level.

Three standing commitments apply throughout the replies below and will not be restated each time. First, the framework does not deny microphysical completeness; it claims non-redundancy at the explanatory and interventional level for declared macro-targets. Second, all verdicts are indexed to regime, horizon, and admissible intervention class, so objections that presuppose context-free ontological claims address a view the paper does not hold. Third, the strongest conclusion is conditional on structural realist and interventionist commitments; readers who withhold those commitments can still accept the narrower anti-gerrymandering result.

5.1 "Instrumentalism and Admissibility"

Objection: target selection is interest-relative, so the view still collapses into usefulness.

Reply: target selection can be interest-shaped without making closure verdicts preference-shaped. Once regime, horizon, and admissibility are fixed, whether macro-transitions are autonomous is a system fact under explicit constraints.

The stronger version invokes Dennett's Martian (Dennett 1991, p. 42). If an ideal micro-predictor can forecast everything, macro realism looks optional. The mistake is to infer ontological redundancy from representational power. Micro omniscience can coexist with objectively better macro bottlenecks for target dynamics. Refusing those bottlenecks increases representational burden. It does not erase macro transition structure.

The deeper point is that observer power changes convenience, not closure facts. Whether a partition is transition-autonomous under fixed constraints does not vary with who computes it.

Admissibility version of the same objection: admissibility criteria already encode what counts as the relevant level. The framework answers this by hierarchical order and disclosure rules. Admissibility is fixed by physically realizable, regime-preserving control channels before partition scoring. Post hoc admissibility revision triggers restart. This makes back-fitting explicit and sanctionable.

Residual disagreement can remain. The framework does not promise universal convergence from one pass. It promises disciplined comparison and explicit downgrade when verdicts are unstable across nearby defensible admissibility specifications. This is a feature for skeptical readers: hard cases are flagged as unresolved rather than settled by stipulation.

Instrumentalist pressure hides in two forms: ontology should track predictive utility only, or ontology is unnecessary once predictive utility is available. The framework rejects both. Utility without closure is too permissive, and closure without ontology understates what stable intervention-guiding structure provides. What closure adds is not metaphysical extravagance. It is commitment discipline. If a candidate repeatedly

supports stable counterfactual control under fixed constraints, refusing any ontological standing starts to look like an empty verbal policy rather than a substantive alternative. The paper does not force maximal realism. It forces explicit criteria for when anti-realism remains credible.

This is also where the cheap-coding objection is handled directly. A recoding can improve local fit and still fail closure under admissible perturbation. The criterion is therefore representation-hostile in exactly the way permissiveness critiques demand.

5.2 "Causal Exclusion Still Defeats Macro Claims"

Objection: if microphysics is causally complete, macro-level causal claims are redundant (Kim 1998).

Reply: closure does not posit a second fundamental cause layer. It identifies when macro-description is sufficient for prediction and intervention at the relevant target level. Microphysical implementation remains untouched.

The non-redundancy claim is explanatory and control-theoretic. A closed macro-partition carries autonomous transition structure for target dynamics. In that respect, macro-level variables are not placeholders for arbitrary omitted micro-details. They are level-appropriate carriers of intervention-relevant structure.

This reply aligns with interventionist causation. Macro-level interventions can be successful and projectible across perturbations without denying microphysical realization (Woodward 2003; Pearl 2000). Compatibility with microphysical completeness is a design feature, not a concession.

The overdetermination worry can be handled directly. The framework does not claim two independent sufficient causes at one event. It claims one implemented process with multiple adequate descriptions for different explanatory and control targets. Exclusion pressure weakens once adequacy is indexed to target and intervention class rather than to a single privileged descriptive level.

Relative to Kim's strongest formulations, the key move is target-indexing rather than layer multiplication (Kim 1998, 2005). A closed macro-description can be non-redundant without competing for micro-level fundamentality, because non-redundancy is inferred when micro-detail ceases to add control-relevant information for the target outcome under fixed constraints. This is not a linguistic escape. It is a claim about which counterfactual dependencies remain stable for a declared class of interventions. Compatibility with microphysical completeness by itself is too weak to do this work, because it leaves open whether macro-variables are dispensable shorthand. Closure narrows that space: when it holds, dispensability must be argued against a stated record of interventional and predictive sufficiency, not simply asserted.

5.3 "Autonomy, Pluralism, and Distinctiveness"

Objection: the view sounds like a formal restatement of familiar special-sciences autonomy claims, not a distinctive realism thesis.

Reply: the paper is continuous with autonomy insights, but it is not equivalent to them. Generic autonomy claims often remain permissive about what qualifies as a

level-worthy grouping. The present view adds a discriminating closure condition with explicit exclusion and downgrade rules. That addition changes verdict structure.

The difference is practical and philosophical. Practical, because the framework gives a protocol for ruling out high-maintenance, transition-incoherent candidates. Philosophical, because it ties realism commitment to transition autonomy under admissible interventions rather than to explanatory convenience alone. In this respect, causal-emergence style results are treated as evidence within a closure-governed framework, not as a replacement for the criterion itself (Hoel et al. 2013).

The paper should therefore be read as a constrained realism refinement of autonomy views, not as an unrelated alternative and not as a mere restatement. The conceptual difference can be stated directly. Generic autonomy claims assert that higher levels matter. Closure specifies *when* they matter, *how much* they matter (graded by leakiness), and *when the claim should be withdrawn* (explicit downgrade conditions). A realism criterion without withdrawal conditions is not a criterion at all. It is an assertion.

A more specific comparison is worth drawing. The classical arguments for the autonomy of the special sciences established that higher-level kinds can be explanatorily indispensable precisely because they abstract over heterogeneous micro-realizations (Putnam 1967; Fodor 1974). Multiple realizability is in fact a structural feature of any genuinely closed macro-kind: if a partition closes, different microstates within the same macroclass by construction produce the same macro-transitions, making the kind multiply realized in the relevant sense. Multiple realizability is therefore necessary for closure. But it is not sufficient. A purely extensional claim about kind membership says nothing about the dynamical coherence of the realizers. A disjunctive aggregate can have heterogeneous realizers while completely failing closure, because those realizers can have utterly divergent transition profiles. What closure adds is the dynamical requirement: the realizers must not only fall under the same label but must produce the same macro-transitions under the declared constraints. This gap, between what a kind groups together extensionally and how those instances behave dynamically, is precisely where gerrymandered kinds exploit multiple-realizability intuitions without earning genuine macro-object status. The present framework formalizes the transition condition that the autonomy tradition always needed to separate genuine higher-level kinds from merely disjunctive aggregates (Shapiro 2000). The monetary illustration in §4.5 provides a concrete instance: coins, ledger entries, and digital balances are heterogeneous realizers, but what earns money macro-object status is not the heterogeneity of realizers but the transition-equivalence of those realizers under admissible financial interventions.

Pluralism version of the objection: if multiple grains can satisfy closure, ontology becomes permissive again. The reply is that plurality of closed grains does not imply arbitrariness. Once regime, horizon, and admissibility are fixed, which partitions close is determined by transition structure, not by analyst choice. When multiple candidates remain viable, robustness and minimality rank them as an objective relation among candidates under fixed constraints. In many domains these relations are further structured by coarse-graining order: when one closed partition is a refinement of

another, their relation is nesting, not rivalry, which disciplines pluralism rather than dissolving it.

This is a virtue for complex systems, not an embarrassment. Forcing a single grain in every context would be a stronger and less defensible claim than the paper needs.

5.4 "The Markov Template is Too Restrictive"

Objection: strong lumpability presupposes first-order Markov microdynamics and excludes memory-dependent systems.

Reply: the criterion is transition sufficiency, not first-order Markovity. Strong lumpability is an exact benchmark. In memory-bearing systems, state can be enriched with relevant history and the same closure question can be asked over that enriched state. This changes state representation, not criterion content.

Concrete cases make this less abstract. Immune response modeling often depends on historical exposure profiles, and institutional dynamics often depend on path-dependent enforcement and trust trajectories. In both cases, a memoryless state is too thin, but a minimally history-enriched state can still support closure testing at the target grain.

Minimality still matters. Enrichment should be the least extension needed to preserve closure performance under fixed constraints. The framework also accepts an important edge case: if the minimally sufficient enriched state approaches micro-level complexity, then closure may be recovered without meaningful compression. In that regime, the result is not robust macro-objecthood. It is a warning that no informative macro-partition has been found at the target grain.

5.5 "Formal Closure Collapses the View into Formalism"

Objection: formal systems are closed by stipulation, so closure cannot ground empirical objecthood.

Reply: stipulated and induced closure must be separated. Formal systems can have objective internal closure under constitutive rules. The present criterion concerns induced closure in implemented spatiotemporal dynamics. Internal formal coherence does not by itself settle empirical macro-objecthood claims.

This does not reduce formal systems to arbitrary invention. Formal systems are invented, but not freely. What counts as intelligible formal practice is constrained by stable identity conditions, compositional inference, and coherent consequence. Those constraints explain why formal work can guide empirical reasoning without itself deciding empirical objecthood.

No collapse follows. The paper's argument is narrower and clearer: induced closure is the objecthood criterion for implemented macro-patterns in regime.

5.6 "Closure Is Too Strong and Eliminates Most Special Sciences"

Objection: if closure requires that within-class micro-differences not matter for macro-transitions, then most special-science kinds will fail. Biological species, psychological

states, and economic categories are notoriously leaky. The criterion eliminates the ontology it was supposed to discipline.

Reply: this objection conflates robust objecthood with any objecthood verdict at all. The framework provides three verdict categories, not one. Robust objecthood requires stable closure under fixed constraints. Qualified objecthood fits cases where closure is partial, regime-limited, or supported by convergent but incomplete evidence. Indeterminate status fits cases where evidence is too unstable to warrant commitment.

Most special-science kinds fall into the qualified range, and that is the correct result. Biological species exhibit substantial transition autonomy within ecological and evolutionary regimes while leaking at boundary cases (hybrid zones, ring species, horizontal gene transfer). Psychological kinds often support predictive and interventional regularity within clinical or experimental regimes while failing under broader perturbation. The framework does not eliminate these kinds. It gives them the status their closure evidence supports: qualified objecthood with explicit regime limitations, rather than either full robust status or wholesale elimination.

The objection therefore proves too much. A criterion that granted robust objecthood to every special-science kind regardless of leakiness would be the permissive framework this paper is designed to replace. The gain from graded verdicts is that they match evidential reality rather than forcing a binary choice between full realism and eliminativism.

5.7 "Predictive Success and Failure Conditions"

Objection: a model can predict well without warranting ontological commitment.

Reply: that is correct, and the framework does not deny it. Prediction alone is not the criterion. The criterion is transition autonomy under admissible interventions and fixed constraints. Predictive success functions as evidence only when it aligns with closure diagnostics and interventional stability.

This point carries two immediate consequences.

First, protocol language has a bounded philosophical role. The paper includes diagnostics to avoid a familiar failure mode, where a criterion is asserted but left too unconstrained to guide verdicts. The claim remains conceptual. Diagnostics do not replace argument. They operationalize what the argument says should matter.

Second, the framework is falsifiable in its own terms. If candidate partitions repeatedly fail closure diagnostics across defensible regimes and admissibility classes, commitment should be withdrawn or downgraded. The same applies when verdicts are unstable under small, defensible changes in horizon, admissibility, or diagnostic proxy. A proposal that cannot risk downgrade is not a serious realism criterion.

The framework does not promise certainty from one metric or one pass. It promises explicit criteria for when confidence increases, when confidence should pause, and when commitment should retreat.

5.8 "Horizon-Relativity Makes the Criterion Too Weak"

Objection: if closure is indexed to horizon, any candidate can be made to pass at a short enough horizon.

Reply: horizon indexing is a constraint, not an escape clause. Two requirements prevent it from becoming one.

First, horizons must be anchored in independently identified timescale structure in the target domain: known relaxation windows, intervention latency, control-response periods, or similar physically motivated temporal features. A horizon chosen only because it makes a favored partition look closed has no independent standing.

Second, verdicts must survive a modest-extension stability check. A candidate that passes at horizon L but fails immediately at $L + \Delta$ for small, defensible Δ receives, at best, qualified status. Robust objecthood requires that closure not be razor-thin in temporal scope.

Together, these requirements force explicit timescale discipline and prevent silent switching between incompatible temporal targets. Horizon flexibility is acceptable when it tracks independently motivated structure. It is not acceptable when it is introduced only after score inspection to salvage a preferred partition.

This distinction is central for pre-emption. Critics are right that any framework can be trivialized if timescales are unconstrained. The present framework avoids that outcome by tying horizons to ex ante justification and by requiring modest-extension robustness. If robustness fails, the verdict downgrades. The criterion does not break. It reports that the claim was too strong for the evidence.

5.9 "The Criterion Is Too Conservative"

Objection: by demanding closure and downgrade discipline, the framework may be too conservative and may miss emerging macro-objects.

Reply: conservatism is partly intentional. The paper aims to avoid premature ontological inflation. Still, the framework is not rigid. It allows qualified verdicts for emerging patterns when closure evidence is partial but improving, and it allows verdict revision as regimes stabilize and intervention evidence accumulates.

So the criterion does not require all-or-nothing maturity before any commitment. It requires that commitment strength track available closure evidence. Woodward (2021) similarly argues that while minimal interventionist criteria establish causality, invariance is a graded virtue. Our criterion of closure can be seen as the dynamical realization of this: a macro-object is "proportional" precisely when it minimizes the omission of relevant micro-detail. This is a virtue for dynamic systems where objecthood can be developmental rather than instantaneous.

This is also where compression and closure reconnect in a non-vacuous way. Early in a domain's development, compression gains may appear before robust closure is demonstrated. The framework treats this as evidential lead, not ontological conclusion. As closure evidence accumulates, commitment can strengthen. If closure evidence fails to accumulate, commitment should remain qualified or be withdrawn even when short-run compression remains attractive. The criterion is therefore neither eliminativist nor inflationary: it does not deny standing to emerging patterns, but it does not award standing before the dynamical evidence warrants it.

6 Conclusion: A Disciplined Realism Claim

Real-pattern realism is plausible but underconstrained. This paper adds the missing discriminating condition: closure of macro-transition structure under fixed regime, horizon, and admissible intervention class. Strong lumpability provides the exact benchmark. Leakiness and convergent diagnostics extend the same criterion to non-ideal settings without changing its content.

What the framework buys is selective commitment. It supports robust commitment where closure is stable, qualified commitment in borderline cases, and explicit downgrade where verdicts depend on fragile admissibility or horizon choices. The proposal is intentionally middle-range: stronger than compression-only pattern realism because it adds exclusion and downgrade structure, weaker than a total-level metaphysics because it is restricted to induced closure in spatiotemporal systems under declared constraints.

This is also where the view separates itself from eliminativist pressure. The framework does not treat higher-level descriptions as fictions by default, and it does not grant them standing by convenience alone. It asks whether they track stable transition structure under admissible intervention.

The paper also marks clear limits. It is not a complete metaphysics of levels and not a universal methods manual. Its contribution is a philosophical criterion with application discipline. Methodological implementation details remain downstream of this argument, and refining them does not alter the criterion's philosophical content. Future work should test the criterion across domains by comparing candidate partitions under fixed constraints, rather than by multiplying new concepts. If rejected, it should be rejected because one rejects its stated commitments, not because circularity, instrumentalist drift, or scope ambiguity were left unresolved.

Acknowledgements

Acknowledgements are omitted for blind review.

Declarations

Competing interests The author declares no competing interests.
Data availability Data sharing is not applicable to this article as no datasets were generated or analyzed.

References

Ainsworth PM (2009) Newman's objection. The British Journal for the Philosophy of Science 60(1):135–171. https://doi.org/10.1093/bjps/axn051

Ando A, Simon HA (1961) Aggregation of variables in dynamic systems. Econometrica 29(2):111–138. https://doi.org/10.2307/1909285

Cover TM, Thomas JA (2006) Elements of Information Theory, 2nd edn. Wiley, Hoboken, NJ. https://doi.org/10.1002/047174882X

Dennett DC (1991) Real patterns. The Journal of Philosophy 88(1):27–51. https://doi.org/10.2307/2027085

Elgin CZ (2017) True Enough. MIT Press, Cambridge, MA

Fodor JA (1974) Special sciences (or: The disunity of science as a working hypothesis). Synthese 28(2):97–115. https://doi.org/10.1007/BF00485230

Gładziejewski P (2025) Real patterns, the predictive mind, and the cognitive construction of the manifest image. Synthese 206:225. https://doi.org/10.1007/s11229-025-05311-0

Goodman N (1955) Fact, Fiction, and Forecast. Harvard University Press, Cambridge, MA

Haugeland J (1998) Having Thought: Essays in the Metaphysics of Mind. Harvard University Press

Hoel EP, Albantakis L, Tononi G (2013) Quantifying causal emergence shows that macro can beat micro. Proceedings of the National Academy of Sciences 110(49):19790–19795. https://doi.org/10.1073/pnas.1314922110

van Inwagen P (1990) Material Beings. Cornell University Press

Kemeny JG, Snell JL (1960) Finite Markov Chains. Van Nostrand. Reprinted 1976, Springer

Kim J (1998) Mind in a Physical World. MIT Press

Kim J (2005) Physicalism, or Something Near Enough. Princeton University Press

Ladyman J, Ross D (2007) Every Thing Must Go: Metaphysics Naturalized. Oxford University Press

Lewis D (1986) On the Plurality of Worlds. Blackwell

Pearl J (2000) Causality: Models, Reasoning, and Inference. Cambridge University Press

Putnam H (1967) Psychological predicates. In: Capitan WH, Merrill DD (eds) Art, Mind, and Religion. University of Pittsburgh Press, Pittsburgh, p 37–48

Rosas FE, Geiger BC, Luppi AI, Seth AK, Polani D, Gastpar M, Mediano PAM (2024) Software in the natural world: A computational approach to hierarchical emergence. arXiv preprint arXiv:2402.09090. URL https://arxiv.org/abs/2402.09090

Shalizi CR, Crutchfield JP (2001) Computational mechanics: Pattern and prediction, structure and simplicity. Journal of Statistical Physics 104(3-4):817–879. https://doi.org/10.1023/A:1010388907793

Shapiro LA (2000) Multiple realizations. The Journal of Philosophy 97(12):635–654. https://doi.org/10.2307/2678460

Woodward J (2003) Making Things Happen: A Theory of Causal Explanation. Oxford University Press

Woodward J (2015) Methodology, ontology, and interventionism. Synthese 192(11):3577–3599. https://doi.org/10.1007/s11229-014-0479-1

Woodward J (2021) Causation with a Human Face: Normative Theory and Descriptive Psychology. Oxford University Press, New York. https://doi.org/10.1093/oso/9780197585412.001.0001