ADVERSARIAL SYSTEMS

A Research Manifesto

On the Replication-Optimization Mechanism, metaethical foundations, and the necessity of a unified field

Murad Farzulla

ORCID: 0009-0002-7164-8704

murad@dissensus.ai

January 2026

"Equilibria are measure-zero. Dissensus is constitutive.
The universe persists by struggling against itself."

"Every sufficiently rigorous formalization of friction dynamics eventually derives cosmology from consent theory.
This is not scope creep. It is derivational inevitability."

Abstract

This manifesto establishes **Adversarial Systems** as a formal field of study, grounded in the Replication-Optimization Mechanism (ROM) and instantiated through the Axiom of Consent.

The argument proceeds in three stages. First, we present ROM as a substrate-neutral formalism capturing how complex systems persist through friction rather than equilibrium. Second, we derive the Axiom of Consent as ROM's metaethical instantiation: delegation necessarily produces friction; friction accumulation is measurable; legitimacy is stakes-weighted voice alignment. Third, we argue that these results necessitate a unified field—Adversarial Systems—because the mechanism is isomorphic across political economy, financial markets, cognitive architecture, and cosmological dynamics.

The wager is empirical. If ROM friction decomposition predicts observable outcomes across domains, adversarial systems is vindicated as unified science. If not, it is an interesting failure. Either way, the current disciplinary fragmentation is not merely inefficient. It is epistemologically incoherent.

1 The Mechanism: Replication-Optimization

1.1 Why a New Formalism

COMPLEX systems persist. This is the explanandum. Institutions survive regime changes. Markets absorb crashes and continue trading. Minds endure trauma and maintain coherence. Ecosystems withstand perturbation and regenerate. The question is: what mechanism governs persistence?

Traditional answers invoke equilibrium. Economic theory posits price convergence. Political theory assumes consensus formation. Thermodynamics predicts entropy maximization. These frameworks share an assumption: systems tend toward stable states; departures are pathological.

This assumption is false.

> Equilibria are measure-zero in generic dynamical systems. Complex systems do not persist *because* they reach stability. They persist *despite* never reaching it. Friction—the continuous pressure against stasis—is not noise to be eliminated. It is the mechanism by which identity is maintained.

The Replication-Optimization Mechanism (ROM) formalizes this insight.

1.2 ROM Defined

Definition: Replication-Optimization Mechanism

A system $\mathcal{S}$ instantiates ROM if it satisfies three conditions:

1. **Replication pressure** (ρ): Components face selection for persistence—continued existence requires active maintenance against dissolution.
2. **Optimization pressure** (σ): Components face selection for fitness—competitive advantage requires adaptation to environmental demands.
3. **Friction dynamics** (F): Replication and optimization are generically misaligned. The friction function $F = f(\alpha, \sigma, \varepsilon)$ decomposes into alignment (α), stakes (σ), and epistemic uncertainty (ε).

The dynamics are captured by:

$$\frac{dp_t(\tau)}{dt} = \sum_{\tau'} p_t(\tau') \cdot \phi(\tau') \cdot M(\tau' \to \tau) - p_t(\tau) \cdot \bar{\phi}_t \tag{1}$$

where $p_t(\tau)$ is the probability mass on configuration τ, $\phi(\tau)$ measures fitness (replication success under optimization pressure), and $M(\tau' \to \tau)$ captures transitions (learning, mutation, reform).

1.3 Substrate Neutrality

Equation (1) does not specify what τ represents. This is not ambiguity—it is generality. The formalism applies wherever the three conditions obtain:

- **Political systems**: τ = governance configurations; ρ = institutional persistence; σ = legitimacy optimization; F = consent-holding friction.
- **Financial markets**: τ = trading strategies; ρ = capital preservation; σ = return optimization; F = information asymmetry + infrastructure fragility.
- **Cognitive systems**: τ = belief configurations; ρ = self-model coherence; σ = environmental prediction; F = competing subsystem objectives.
- **Cosmological dynamics**: τ = spacetime density distributions; ρ = localization maintenance; σ = entropic dissipation; F = expansion-contraction tension.

Theorem: Equilibria are Measure-Zero

For generic ROM systems with continuous state space and non-degenerate fitness landscape, the set of fixed points has Lebesgue measure zero. Equilibrium configurations exist mathematically but are dynamically unattainable.

Corollary: Dissensus is structurally inevitable. Systems that appear stable are cycling through quasi-stable configurations faster than observation resolution. What we call "equilibrium" is coarse-grained persistence, not fine-grained stasis.

2 The Foundation: Axiom of Consent

2.1 From Mechanism to Metaethics

ROM is descriptive. It characterizes *how* systems persist. But complex systems involving agents raise normative questions: Who bears the costs of friction? Whose optimization targets count? How should persistence be distributed?

These questions require metaethical grounding. The Axiom of Consent provides it.

Axiom of Consent

In any system where agents delegate decision-making authority:

1. Delegation produces friction. The delegator's optimization target (σ_i) diverges from the delegatee's (σ_j) with probability 1.
2. Friction accumulates measurably. The friction function $F = f(\alpha, \sigma, \varepsilon)$ is computable from observable parameters.
3. Legitimacy is stakes-weighted voice. The legitimacy of an outcome is:

$$L(d) = \frac{\sum_i \sigma_i \cdot v_i(d)}{\sum_i \sigma_i} \tag{2}$$

where σ_i is stakeholder i's stakes and $v_i(d)$ is their voice (influence) over decision d.

2.2 The Belief-Transfer Mechanism

Delegation does more than create friction. It generates a psychological phenomenon with ethical weight: **belief-transfer**.

Theorem: Belief-Transfer under Delegation

Let $P_j(t)$ denote a delegatee's subjective probability that consent "belongs" to them rather than the delegator. Under Bayesian updating with asymmetric evidence availability:

$$\frac{dP_j}{dt} > 0 \quad \text{for all } t > 0$$

As delegation duration increases, delegatees come to perceive delegated consent as intrinsic property.

This is not pathology. It is optimal inference given the evidence structure. The delegatee experiences daily reinforcement (they exercise authority) while counterfactual evidence (delegator's continued consent) is invisible. Belief-transfer is Bayes-optimal on the available data.

Metaethical implication: Consent reclamation is structurally traumatic. When delegators attempt to reclaim authority (decolonization, employee exit, divorce), delegatees experience genuine loss—not theater, not bad faith, but the dissolution of what felt like owned property. This does not make reclamation wrong. It makes friction inevitable.

2.3 Stakes-Weighted Legitimacy

Equation (2) provides a quantitative legitimacy criterion. Its structure has normative consequences:

1. **Uniform voice fails**. Democratic theory's "one person, one vote" ignores stakes heterogeneity. A policy affecting someone's livelihood deserves more weight than one affecting their mild preference. Stakes-weighting captures this.

2. **Technocracy fails**. Pure expertise delegation ignores voice entirely. Competent decisions imposed without consent accumulate friction until legitimacy collapses. Colonial administration exhibits this pattern: peak bureaucratic sophistication immediately precedes decolonization.

3. **Optimal governance is productive dissensus**. Neither consensus (suppressed friction) nor gridlock (exhausted legitimacy) but managed contestation: enough friction to force accountability, not so much that coordination fails.

The Axiom of Consent transforms ROM from descriptive mechanism to normative framework. Friction is not merely observable—it is evaluable. Systems can be judged by whether their friction distribution respects stakes-weighted voice.

3 The Field: Adversarial Systems

3.1 Definition and Scope

We now have the components: ROM (mechanism) + Axiom of Consent (metaethics). The synthesis is **Adversarial Systems**.

Definition: Adversarial Systems

Adversarial Systems is the formal study of how complex systems persist through friction dynamics, with particular attention to:

1. The ROM formalism and its instantiations across substrates
2. Consent structures and their friction implications
3. Stakes-weighted legitimacy and its measurement
4. Coarse-graining and scale transitions
5. Empirical validation across domains

3.2 Why a Unified Field is Necessary

The case for disciplinary unification is not aesthetic preference. It is logical necessity.

Argument from derivational closure:

1. Political friction requires formalizing stakes, alignment, uncertainty.
2. These parameters are not primitive—they emerge from optimization dynamics.
3. Optimization dynamics require ROM: persistence under selection pressure.
4. ROM requires entropy as axiom: patterns dissolve unless maintained.
5. Entropy is dedensification: $S \propto \int \rho^{-1} dV$.
6. Dedensification dynamics imply cosmological structure.

Therefore: Rigorous formalization of political friction eventually derives cosmology. This is not scope creep. It is what happens when you refuse to hand-wave.

Argument from isomorphism:

The same mathematics governs:

- How institutions lose legitimacy (consent-friction accumulation)
- How markets crash (infrastructure-friction propagation)

- How minds develop trauma (adversarial-training optimization)
- How entropy increases (dedensification under expansion)

Isomorphic problems demand unified treatment. Seven disciplines studying the same phenomenon with incompatible notation is not intellectual diversity. It is coordination failure.

Argument from predictive power:
Cross-domain validation is the strongest test. If ROM friction decomposition predicts:

- Financial volatility asymmetry (infrastructure > regulatory shocks)
- Institutional failure timing (legitimacy erosion threshold)
- Therapeutic intervention efficacy (adversarial-training reversal)
- Cosmological observables (Hubble tension directional variance)

then the unified framework captures something real. Domain-specific theories cannot make cross-domain predictions. Adversarial Systems can.

3.3 Research Program

Immediate objectives (2–5 years):

1. **Empirical replication**: TARCH-X cryptocurrency study across asset classes; belief-transfer mechanism in organizational capture; trauma-as-optimization in longitudinal developmental studies.
2. **Formalization**: Complete lumpability proofs for arbitrary coarse-graining; extend ROM to quantum field theory; derive Born rule from traversal statistics.
3. **Policy application**: Climate governance as consent-holding failure (COP negotiations); AI regulation as friction-aware design; decolonization reparations via belief-transfer duration quantification.

Long-term vision (10–20 years):

1. Adversarial Systems as recognized field with dedicated journals and cross-departmental programs.
2. Friction reframed as constitutive across disciplines; equilibrium models recognized as measure-zero special cases.
3. Testable cosmological predictions: Hubble tension directional variance, dark energy equation of state evolution $w(z)$, CMB non-Gaussian signatures.

4 The Wager

The Wager

This research program makes falsifiable empirical claims:

1. ROM friction decomposition predicts volatility asymmetry in financial markets. (TARCH-X: infrastructure shocks produce 5.7× larger response than regulatory shocks, $p = 0.0008$, Cohen's $d = 2.75$.)
2. Stakes-weighted legitimacy forecasts institutional crises. (Testable: historical regime transitions should exhibit legitimacy erosion preceding collapse.)
3. Belief-transfer duration predicts reclamation trauma intensity. (Testable: longer delegation correlates with more conflictual separation.)
4. Entropy-as-dedensification implies cosmological observables. (Testable: Hubble tension should show directional variance if expansion rate varies with local density.)

If these predictions hold across domains, Adversarial Systems is vindicated as unified science.
If they fail, the framework is falsified.
Either outcome advances knowledge. Validation yields paradigm shift. Falsification yields precise boundaries. The current alternative—disciplinary fragmentation with no cross-domain predictions—yields neither.

5 Conclusion

THE argument is complete:

1. **ROM** provides substrate-neutral formalism for persistence under friction.
2. **Axiom of Consent** instantiates ROM in metaethics: delegation produces friction; friction is measurable; legitimacy is stakes-weighted voice.
3. **Adversarial Systems** unifies these results into a coherent field, with cross-domain predictions as validation criterion.

The normative implication is clear. If friction is constitutive—if systems persist *through* dissensus rather than *despite* it—then evaluation criteria must change:

- Good governance is not consensus but managed contestation.

- Market health is not volatility minimization but infrastructure robustness.
- Mental health is not affect stability but adaptive friction between competing goals.
- The universe does not decay toward equilibrium. It breathes.

Disciplinary silos can persist. Political theorists can ignore econometrics. Economists can dismiss postcolonial critique. Physicists can treat entropy as settled. This is permitted.

But do not then express surprise when institutions collapse, markets crash, minds break, and cosmological tensions accumulate—all exhibiting identical friction signatures that no single discipline can explain.

The universe is adversarial against itself. So are we. Understanding this is not optional. It is the condition of persistence.

Murad Farzulla

murad@dissensus.ai

ORCID: 0009-0002-7164-8704