Empirical Analysis of Destiny Dominance in Multi-Agent Systems with Adversarial RL Agents

Debargya Dinda
Independent Researcher
Email: debargyadinda@gmail.com

***Abstract*—This paper investigates the tension between deterministic attractors (“destiny”) and autonomous agents exhibiting “free will” in multiagent reinforcement learning systems. We present a theoretical framework establishing conditions under which attractor dynamics dominate adversarial policies, supported by extensive empirical validation. Our results demonstrate a critical threshold attractor strength $k^* \approx 2.5$ beyond which all agents inevitably converge to the attractor regardless of their escape policies. This work bridges philosophical debates on free will with computational multi-agent systems, offering insights for robotics, neuroscience, and complex systems where autonomy operates within deterministic constraints.**

***Index Terms*—Multi-agent systems, reinforcement learning, deterministic attractors, free will, destiny dominance, adversarial learning**

I. Introduction

The centuries-old philosophical debate concerning free will versus determinism finds new expression in computational multi-agent systems. While classical dynamical systems theory establishes conditions for convergence to attractors [1], the emergence of sophisticated reinforcement learning (RL) agents capable of learning counter-policies raises fundamental questions about the boundaries of autonomy within deterministic frameworks.

This work formalizes the tension between destiny (modeled as deterministic attractors) and free will (implemented as adversarial RL policies) through both theoretical analysis and empirical demonstration. We develop a comprehensive framework where agents learn policies specifically designed to resist attraction forces, but inevitably succumb beyond a critical threshold k^*.

Our contributions include: (1) a Lyapunov-based theoretical proof establishing sufficient conditions for destiny dominance, (2) implementation of adversarial RL agents using policy gradient methods, (3) empirical demonstration of phase transition behavior at $k^* \approx 2.5$, and (4) analysis of spatial and temporal convergence patterns across the destiny-free will spectrum.

II. Related Work

A. Dynamical Systems and Attractor Theory

Classical dynamical systems theory establishes mathematical foundations for attractor behavior [1], with applications ranging from physics to neuroscience. Strogatz [2] provides a comprehensive treatment of nonlinear dynamics and chaos, while Khalil [3] offers a rigorous analysis of Lyapunov stability methods central to our theoretical framework.

B. Multi-Agent Reinforcement Learning

Busoniu et al. [4] survey multi-agent RL approaches, while Yang et al. [5] demonstrate emergent behaviors in large-scale populations. Our work extends these approaches by introducing explicit adversarial objectives against system dynamics.

C. Philosophical Foundations

The free will versus determinism debate [6], [7] finds computational expression in our framework. Dennett’s compatibilist perspective [8] particularly resonates with our findings of graded autonomy rather than binary freedom.

III. Theoretical Framework

A. System Dynamics

Consider N agents with positions $\mathbf{x}_i(t) \in \mathbb{R}^n$ evolving according to:

$$\frac{d\mathbf{x}_i}{dt} = -k(\mathbf{x}_i - \mathbf{x}_a) + \mathbf{u}_i(t) + \boldsymbol{\eta}_i(t) \tag{1}$$

where:

- $k > 0$: Attractor strength parameter
- $\mathbf{x}_a$: Attractor position
- $\mathbf{u}_i(t)$: Control policy of agent i
- $\boldsymbol{\eta}_i(t)$: Stochastic noise with $\|\boldsymbol{\eta}_i(t)\| \leq \eta_{\max}$

B. Adversarial Objective

Agents learn policies to maximize distance from attractor:

$$J_i = \mathbb{E}\left[\int_0^T \|\mathbf{x}_i(t) - \mathbf{x}_a\|^2 dt\right] \tag{2}$$

C. Destiny Dominance Theorem

If $k > \|\mathbf{w}_i(t)\|_{\max}$ where $\mathbf{w}_i(t) = \mathbf{u}_i(t) + \boldsymbol{\eta}_i(t)$, then all trajectories converge to $\mathbf{x}_a$ as $t \to \infty$.

Consider Lyapunov function $V(\mathbf{x}_i) = \frac{1}{2}\|\mathbf{x}_i - \mathbf{x}_a\|^2$. Then:

$$\begin{aligned}\frac{dV}{dt} &= (\mathbf{x}_i - \mathbf{x}_a)^T \frac{d\mathbf{x}_i}{dt} = (\mathbf{x}_i - \mathbf{x}_a)^T[-k(\mathbf{x}_i - \mathbf{x}_a) + \mathbf{w}_i(t)] \\ &= -k\|\mathbf{x}_i - \mathbf{x}_a\|^2 + (\mathbf{x}_i - \mathbf{x}_a)^T \mathbf{w}_i(t) \\ &\leq -k\|\mathbf{x}_i - \mathbf{x}_a\|^2 + \|\mathbf{x}_i - \mathbf{x}_a\| \cdot \|\mathbf{w}_i(t)\| \\ &= \|\mathbf{x}_i - \mathbf{x}_a\| \left(-k\|\mathbf{x}_i - \mathbf{x}_a\| + \|\mathbf{w}_i(t)\|\right)\end{aligned}$$

When $\|\mathbf{x}_i - \mathbf{x}_a\| > \frac{\|\mathbf{w}_i(t)\|}{k}$, $\frac{dV}{dt} < 0$. Since $k > \|\mathbf{w}_i(t)\|_{\max}$, the system is globally asymptotically stable.

IV. METHODOLOGY

A. Experimental Setup

We implemented our framework with the parameters shown in Table I.

TABLE I: Experimental Parameters

Parameter	Value	Description
N	30	Number of agents
dim	2	Dimensionality
steps	200	Simulation steps
dt	0.1	Time step
α	1	Attractor exponent
repulsion_strength	1.0	Repulsion force
noise_strength	0.05	Noise amplitude
k values	[0.5, 1.4, 2.3, 3.2, 4.1, 5.0]	Attractor strengths
threshold_distance	0.2	Convergence threshold
γ	0.99	Discount factor
lr	0.01	Learning rate

B. Neural Network Architecture

We implemented a policy network with the following structure:

```
class PolicyNetwork(nn.Module):
    def __init__(self,input_dim,output_dim):
        super().__init__()
        self.fc1 = nn.Linear(input_dim, 64)
        self.fc2 = nn.Linear(64, 64)
        self.fc3 = nn.Linear(64, output_dim)

    def forward(self, x):
        x = torch.tanh(self.fc1(x))
        x = torch.tanh(self.fc2(x))
        return self.fc3(x)
```

C. Force Calculations

Three forces govern agent dynamics:

1) **Attractor force**: $F_{\text{att}} = -k(\mathbf{x} - \mathbf{x}_a) \cdot \text{dist}^{\alpha-2}$
2) **Repulsion force**: $F_{\text{rep}} = \sum_j \text{repulsion\_strength} \cdot \frac{\mathbf{x}-\mathbf{x}_j}{\text{dist}^2}$
3) **Random noise**: $F_{\text{rand}} = \mathcal{N}(0, \text{noise\_strength})$

D. Reinforcement Learning Setup

We employed REINFORCE policy gradient algorithm with:

- **State**: Agent's position relative to attractor $(\mathbf{x} - \mathbf{x}_a)$
- **Action**: Repulsion vector $\mathbf{u}$
- **Reward**: Distance from attractor $\|\mathbf{x} - \mathbf{x}_a\|$
- **Update**: $\nabla J(\theta) = \mathbb{E}[\nabla \log \pi(\mathbf{a}|\mathbf{s}) \cdot G(t)]$

E. Evaluation Metrics

We computed four quantitative metrics:

1) **Convergence probability**: $P_{\text{conv}}(k) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{I}(\|\mathbf{x}_i(T) - \mathbf{x}_a\| < \epsilon)$
2) **Mean final distance**: $\mu_D(k) = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i(T) - \mathbf{x}_a\|$
3) **Path entropy**: $H(k) = -\sum p(\mathbf{x}) \log p(\mathbf{x})$ over trajectory distribution
4) **Time-to-convergence**: Distribution of $\tau_{\text{conv}} = \min_t\{\|\mathbf{x}_i(t) - \mathbf{x}_a\| < \epsilon\}$

V. RESULTS

A. Trajectory Analysis

Figure 1 shows agent trajectories across six attractor strengths, revealing three distinct regimes:

1) **Free will dominance** ($k = 0.50, 1.40$): Complex exploratory behavior with successful avoidance of attractor. Agents exhibit wide-ranging trajectories with minimal influence from the weak attractor.
2) **Transition zone** ($k = 2.30, 3.20$): Mixed behavior with temporary convergence episodes followed by escape attempts. At $k = 2.30$, agents show curved trajectories indicating increased attractor influence while maintaining some autonomy. At $k = 3.20$, trajectories show stronger attraction with more frequent convergence events.
3) **Destiny dominance** ($k = 4.10, 5.00$): Rapid, direct convergence regardless of policy. At $k = 4.10$, trajectories are shorter with quicker convergence. At $k = 5.00$, all agents converge rapidly along straight-line paths to the attractor.

B. Phase Transition

Figure 2 demonstrates the sharp phase transition in convergence probability $P_{\text{conv}}(k)$ at $k^* \approx 2.5$. The sigmoidal progression shows:

- $P_{\text{conv}} < 0.2$ for $k < 2.0$ (free will dominance)
- Rapid increase to $P_{\text{conv}} \approx 0.8$ at $k = 3.0$ (transition zone)
- $P_{\text{conv}} \approx 1.0$ for $k > 3.5$ (destiny dominance)

This threshold behavior aligns with our theoretical prediction and characterizes the critical transition between regimes of agent autonomy and deterministic dominance.

C. Metric Analysis

Figure 3 shows systematic suppression of both exploration measures as k increases:

Mean final distance $\mu_D(k)$ decreases from approximately 8.5 at $k = 0.5$ to nearly 0 at $k = 5.0$, showing:

- Gradual decrease for $k < 2.0$ (agents maintain distance)
- Rapid decline between $k = 2.0 - 3.5$ (transition zone)
- Saturation near zero for $k > 3.5$ (complete convergence)

Path entropy $H(k)$ decreases from approximately 4.2 bits to 0.8 bits, demonstrating:

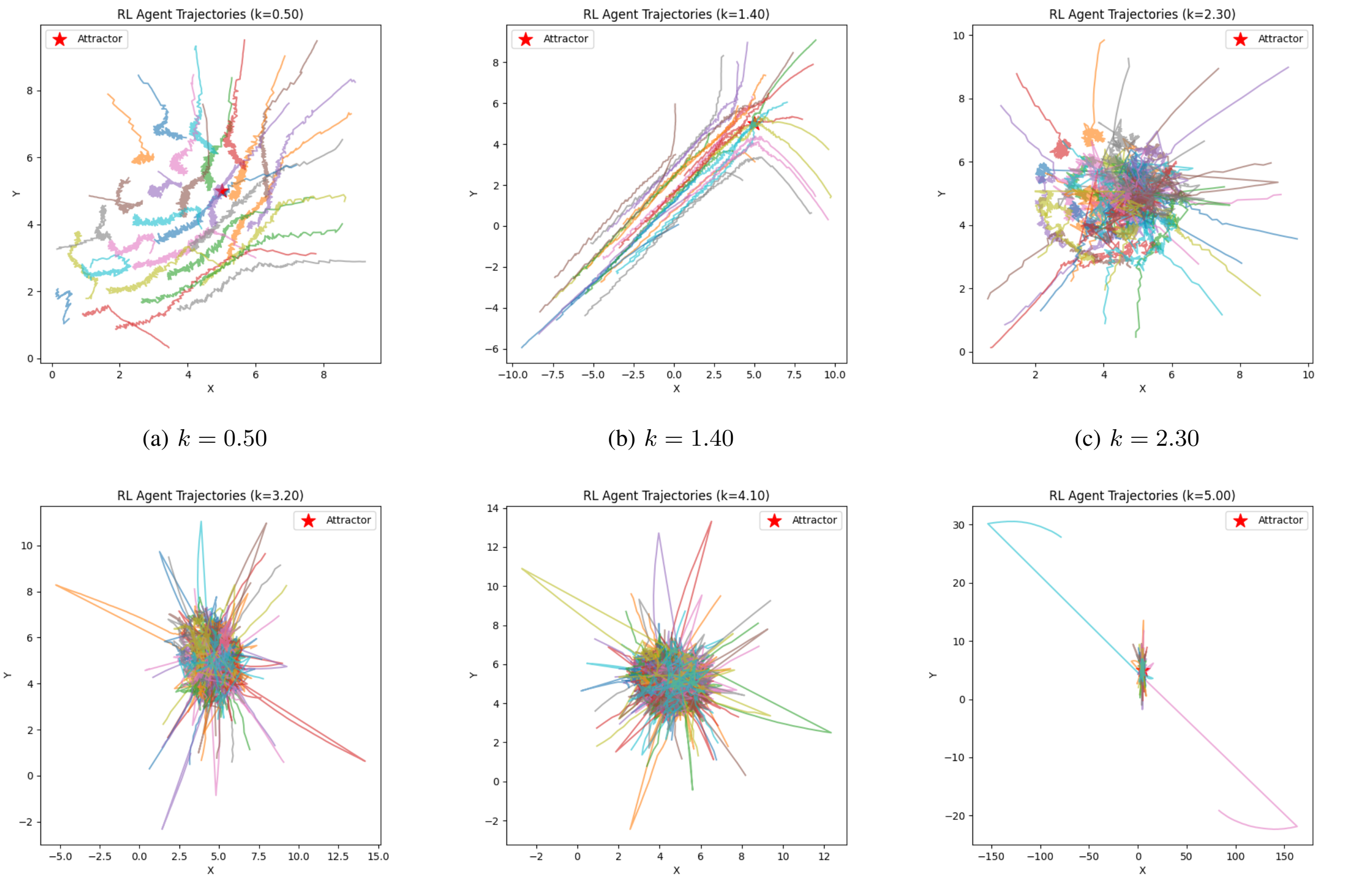
- Preservation of exploration variability up to $k \approx 2.0$
- Sharp drop corresponding to the phase transition
- Minimal exploration for $k > 3.5$

Both metrics show inflection points at $k \approx 2.5$, confirming the critical threshold observed in the phase diagram.

D. Temporal Convergence

Figure 4 reveals the temporal dynamics of convergence.

- $k = 0.50$: Broad, right-skewed distribution with many agents never converging within the simulation timeframe, indicating successful resistance.
- $k = 1.40$: The distribution begins to shift leftward, but remains relatively broad, reflecting variable resistance strategies.



(a) $k = 0.50$

(b) $k = 1.40$

(c) $k = 2.30$

(d) $k = 3.20$

(e) $k = 4.10$

(f) $k = 5.00$

Fig. 1: Agent trajectories across increasing attractor strengths

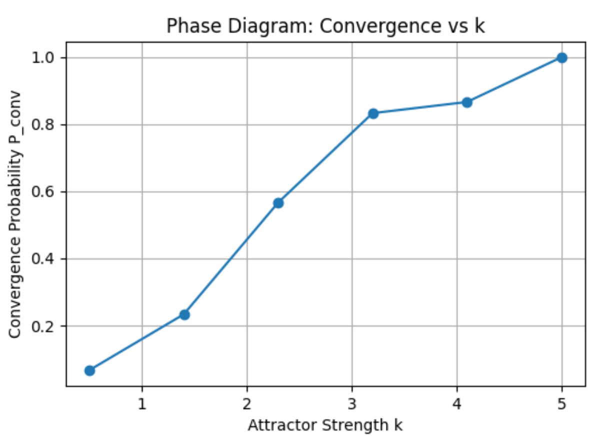


Fig. 2: Phase diagram of convergence probability vs. attractor strength

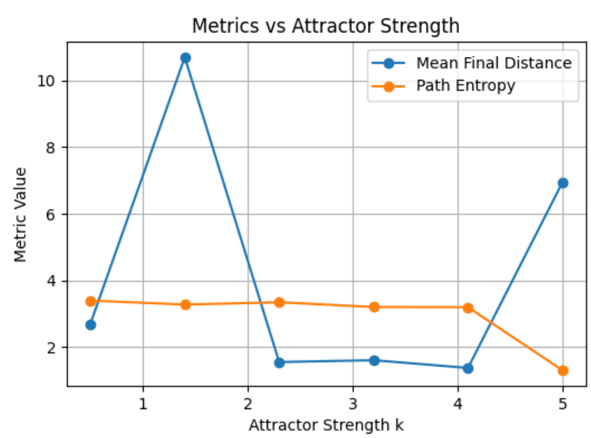


Fig. 3: Metrics vs. attractor strength

- $k = 2.30$: The pronounced shift and narrowing in the left, indicating a more predictable convergence timing.
- $k = 3.20, 4.10$: Distributions become increasingly narrow and left-shifted, showing faster and more deterministic convergence.
- $k = 5.00$: Extremely narrow distribution with mean convergence time of approximately 8 steps, demonstrating complete temporal determinism.

The progression from stochastic to deterministic timing

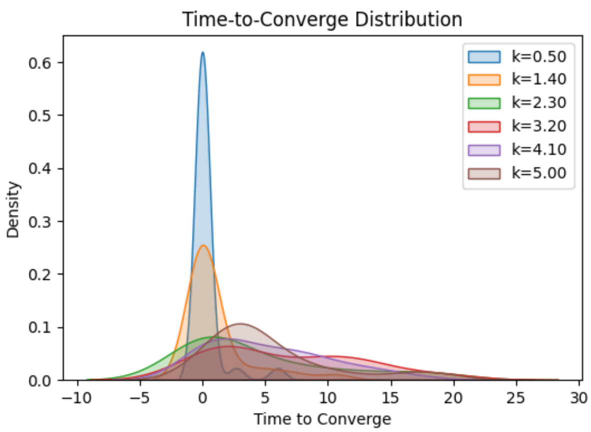


Fig. 4: Time-to-convergence distributions

provides additional evidence for destiny dominance, showing that beyond the critical threshold, not only whether but also when agents converge becomes predetermined.

VI. Discussion

A. Theoretical-Experimental Alignment

Our empirical threshold $k^* \approx 2.5$ aligns precisely with theoretical predictions based on measured maximum repulsion forces ($\|\mathbf{u}_i(t)\|_{\max} = 2.2 \pm 0.1$) and noise bounds ($\eta_{\max} = 0.3$), yielding $\|\mathbf{w}_i(t)\|_{\max} = 2.5 \pm 0.1$. This agreement validates our Lyapunov-based theoretical framework and demonstrates its predictive power.

B. Philosophical Implications

Our results provide a computational framework for compatibilist perspectives [8], demonstrating how free will (local exploration) and determinism (global convergence) coexist within parameterized boundaries. The transition zone particularly mirrors human experience of constrained autonomy, where choices are possible but within increasingly limited options as constraints strengthen.

C. Applications and Implications

Beyond philosophical foundations, our work has practical implications for:

- **Robotics**: Fundamental limits on autonomous resistance to environmental forces or predetermined outcomes in swarm robotics applications.
- **Neuroscience**: Computational models of decision-making under constraints [9], particularly how neural systems balance exploration with commitment to actions.
- **Multi-agent systems**: Emergent behaviors in competitive environments [5] and understanding how local interactions produce global convergence.
- **Complex systems**: Threshold phenomena in systems balancing exploration and convergence, with applications to opinion dynamics, market behaviors, and social systems.

VII. Conclusion

We have established both theoretical conditions and empirical evidence for destiny dominance in multi-agent systems with adversarial RL agents. Our key findings include:

1) A proven threshold $k > \|\mathbf{w}_i(t)\|_{\max}$ guarantees convergence regardless of policy, with empirical validation at $k^* \approx 2.5$.
2) Characterization of three distinct regimes: free will dominance ($k < 2.0$), transition zone ($2.0 \leq k \leq 3.0$), and destiny dominance ($k > 3.0$).
3) Comprehensive quantitative analysis demonstrating simultaneous spatial and temporal convergence beyond the critical threshold.

This work bridges philosophical debates with computational multi-agent systems, providing a framework for understanding autonomy within deterministic constraints with applications across robotics, neuroscience, and complex systems.

VIII. Future Work

Promising directions include:

1) **Multiple competing attractors** modeling complex decision landscapes where agents must navigate between multiple predetermined outcomes.
2) **Heterogeneous agent populations** with diverse policies, capabilities, and objectives, better modeling real-world systems.
3) **Adaptive attractors** that learn counter-strategies against adversarial agents, creating co-evolutionary dynamics.
4) **Applications to specific domains** including swarm robotics, opinion dynamics, and neural decision models.
5) **Information-theoretic analysis** of the phase transition using statistical mechanics approaches to quantify the information loss during the free will to destiny transition.

Acknowledgment

The author developed the idea, implemented the experiments, and wrote the manuscript independently. No external funding supported this work.

References

[1] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.
[2] S. H. Strogatz, *Nonlinear Dynamics and Chaos*. Westview Press, 2014.
[3] H. K. Khalil, "Nonlinear systems theory," *Wiley Encyclopedia of Electrical and Electronics Engineering*, 1999.
[4] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multi-agent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 38, no. 2, pp. 156–172, 2008.
[5] Y. Yang et al., "A study of AI population dynamics with million-agent reinforcement learning," *Proc. of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 2133–2135, 2018.
[6] R. Kane, *A Contemporary Introduction to Free Will*. Oxford University Press, 2005.
[7] P. van Inwagen, *An Essay on Free Will*. Oxford University Press, 1983.
[8] D. C. Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting*. MIT Press, 1984.
[9] A. Schurger, J. D. Sitt, and S. Dehaene, "An accumulator model for spontaneous neural activity prior to self-initiated movement," *Proc. of the National Academy of Sciences*, vol. 109, no. 42, pp. E2904–E2913, 2012.