“Babeș-Bolyai” University of Cluj-Napoca

Faculty of Economics and Business Administration

Doctoral School of Economics and Business Administration

SUMMARY OF DOCTORAL THESIS

PhD Candidate:

Gabriel Marius Pop

Scientific Advisor:

Prof. univ. dr. habil. Dorina Lazăr

Cluj-Napoca

2025

UNIVERSITATEA BABEŞ-BOLYAI
BABEŞ-BOLYAI TUDOMÁNYEGYETEM
BABEŞ-BOLYAI UNIVERSITÄT
BABEŞ-BOLYAI UNIVERSITY
TRADITIO ET EXCELLENTIA

UBBFSEGA
Facultatea de Ştiinţe Economice şi Gestiunea Afacerilor
ŞCOALA DOCTORALĂ DE ŞTIINŢE ECONOMICE ŞI GESTIUNEA AFACERILOR

ESSAYS ON MONTE CARLO ANALYSIS OF REPEATED GAMES: AXELROD'S ITERATED PRISONER'S DILEMMA AND AUTOCOVARIANCE IN THE ALLISON MIXTURE

– SUMMARY –

PhD Candidate:
Gabriel Marius Pop

Scientific Advisor:
Prof. univ. dr. habil. Dorina Lazăr

Cluj-Napoca
2025

Table of Contents of the Doctoral Thesis

Contents of Summary

Chapter 1

Introduction

Strategic decision-making is essential for understanding complex multi-agent interactions, and game theory is a powerful tool for analysing the complex structure of these behaviors. Among various analytical approaches, repeated games stand out for their ability to explain the evolution of cooperation, conflict, and adaptation, offering deeper insights than traditional static, one-shot models. A significant component of this thesis focuses on the Iterated Prisoner's Dilemma (IPD), which has been extensively explored through Robert Axelrod's computational tournaments. Axelrod's work revealed how agents' strategic choices depend not only on immediate payoffs but also significantly on historical experiences, expectations about future interactions, and adaptive learning over time. An important theoretical question arises within this framework: *Are the agents developed within the framework of Axelrod tournaments, endowed with sophisticated reasoning, capable of distinguishing between the implications of finite versus infinite horizons?* Although we cannot directly test behavior in an infinite horizon, we can introduce uncertainty regarding the number of stages in a supergame. This allows us to observe whether successive generations of agents in Axelrod tournaments behave purely mechanically from a time-span perspective or if they internalize the certainty or uncertainty of the number of stages in the supergames they play. Such an approach allows us to explore whether agents across generations act mechanically, driven purely by immediate incentives, or whether they internalize the uncertainty about game length, thus adapting their strategies in more sophisticated, forward-looking ways. This thesis aims to address these nuanced questions, enriching our understanding of inter-temporal strategic decision-making and providing valuable insights for economic theory, political modelling, behavioral science, and the development of intelligent artificial agents.

This thesis offers a new framework for understanding strategic interaction, emphasizing *uncertainty as a fundamental structural feature* of repeated games. Through a fundamental *transition from fixed-horizon to uncertain-horizon modeling*, this work introduces significant theoretical and operational enhancements applicable to simulation-based experimental contexts. A key contribution is the *design and implementation of an innovative Monte Carlo simulation system*, specifically created for this research, offering a strong methodological and experimentally tested approach to evaluating strategic behavior under unknown game durations. By *integrating formal game-theoretic reasoning with computational simulations*, this research advances a scalable theory of decision-making for dynamic multi-agent systems. It offers new insights into the mechanisms behind cooperation, its vulnerabilities, and the influence of adaptability and environmental volatility. By moving beyond fixed-horizon assumptions and validating a flexible framework for behavioral modeling, the thesis significantly contributes to understanding strategic intelligence under uncertainty.

Chapter 2 offers a solid theoretical and methodological framework for analyzing strategic behavior in repeated games, using the Iterated Prisoner's Dilemma (IPD) as a central model. It begins with core concepts from classical game theory—normal-form representations, dominance relations, and Nash equilibria in pure and mixed strategies—essential for characterizing strategic interactions.

The chapter then explores how repetition changes incentives: while one-shot games predict defection, repeated interactions—especially with uncertain or infinite horizons—can foster cooperation via reciprocity, punishment, and reputation. The Folk Theorem and equilibrium refinements illustrate how patient agents can sustain cooperation even without external enforcement. Belief formation and monitoring structures are also examined for their role in strategy credibility.

An original *Monte Carlo simulation approach* is introduced to assess strategy robustness under uncertainty. The environment uses stochastic halting criteria based on normal distributions, enabling repeated evaluation of agent performance across probabilistic scenarios. Agents are modeled using deterministic logic and stochastic finite-state machines, supporting systematic comparisons of a wide range of strategies within a unified experimental setup.

Chapter 3 investigates how the structure of repeated games—particularly the (un)certainty of their duration—influences strategic decision-making in multi-agent systems. It contrasts finite-horizon games, where backward induction leads to defection, with infinite or uncertain horizons that allow cooperation through discounting, belief formation, and reciprocity.

The chapter integrates recent advances in infinite-horizon dynamic games, including anticipative feedback and information-updating models, to depict agents with adaptive, foresight-driven behavior shaped by memory and environmental signals.

This research *extends Axelrod's tournaments* within a simulation framework that incorporates uncertain game duration via stochastic halting criteria based on normal distributions, better mirroring real-world scenarios without known interaction endpoints. This approach broadens the analytical scope of standard IPD studies.

To support this, a custom Monte Carlo simulation system was developed. *A key contribution is the experimental platform itself*, enabling variable-length matches, automated outcome tracking, and rich data exports. Visual tools—like pie charts, cooperation matrices, and heatmaps—highlight emergent patterns and strategy mismatches.

The simulation also reconstructed Axelrod's Second Tournament under probabilistic horizons. Strategies were evaluated through repeated trials using new performance metrics. *A methodological innovation is the use of discrepancy-based indicators*, such as rank shifts and payoff volatility, allowing empirical classification of strategies as robust, adaptive, or volatile.

Chapter 4 offers an in-depth analysis of the Allison Mixture model as a methodological innovation for studying strategic behavior in repeated games under uncertainty. The model is contextualized within stochastic processes and Parrondo's paradox, showing how probabilistic switching between neutral or suboptimal strategies can generate favorable emergent outcomes. This illustrates how simple structural rules can lead to complex adaptive behavior in ambiguous environments.

A major contribution is the *translation of the Allison Mixture into a strategic framework* for repeated interactions, via a mathematically grounded model of transition dynamics and autocovariance. This captures how memory, signal variability, and asymmetry interact in uncertain strategic settings.

The chapter also presents the *design of a custom simulation infrastructure* for extensive Monte Carlo experiments across varying transition probabilities and interaction lengths. This system enables precise testing of convergence, estimator variance, and reward consistency under dynamic conditions, addressing limitations of earlier low-scale or purely analytical approaches.

The model is further applied in both symmetric and asymmetric game scenarios, showing its value *not only as a statistical construct but as a conceptual tool* for modeling adaptive strategies. Its ability to reproduce autocorrelation and structural dependence where classical strategies fail marks it as a promising instrument for future work in AI and decision theory. Finally, the research provides a *reproducible and extensible simulation architecture*, opening the path for hybrid models that combine probabilistic transitions with strategic heuristics. This framework serves not only as a validation engine for the Allison Mixture but also as a foundation for future experimental designs in the broader field of dynamic strategic interaction. In this broader context, the present research contributes to a reconfiguration of how repeated games are theoretically modeled and empirically evaluated by shifting the analytical focus from fixed-horizon assumptions to probabilistic representations of interaction duration. Through the construction of an integrated Monte Carlo simulation framework and its dual application to both classical strategy tournaments and stochastic processes like the Allison Mixture, the thesis not only tests the structural robustness of known strategies but also introduces new *methodological avenues for capturing volatility, adaptability, and emergent cooperation in uncertain environments*.

This orientation toward uncertainty as a structural element—rather than a boundary condition—allows the research to better reflect the realities of decentralized decision-making and temporally ambiguous settings. The tools and insights developed here hold relevance for future studies in multi-agent systems, behavioral modeling, and the computational foundations of adaptive intelligence. By combining classical game-theoretic insight with scalable experimentation, the thesis lays a conceptual and technical groundwork for more realistic and predictive models of strategic behavior under dynamic and uncertain conditions.

Chapter 2

Summary of Chapter 2 : Theoretical, formal, and methodological foundations of Strategic Decision-Making in Repeated Games and the Iterated Prisoner's Dilemma

Chapter 2 establishes the conceptual, formal, and computational framework required for analyzing strategic behavior in repeated games, using the Iterated Prisoner's Dilemma (IPD) as a central model. This structure is vital for understanding the dynamics of cooperation and conflict in multi-agent systems, both in theoretical and applied contexts. The chapter is structured into six major sections, each contributing uniquely to the development of a rigorous methodology for studying strategic robustness under temporal uncertainty.

Theoretical Concepts and Notations

This opening section lays the groundwork for game-theoretic reasoning by introducing the formal structure of normal-form games, including the mathematical representation of players, strategies, and payoffs. Nash equilibrium is rigorously defined and discussed both in pure and mixed strategies, highlighting the strategic stability it offers across a wide range of environments (Nash, 1950). The discussion moves beyond classical settings by incorporating insights from evolutionary game theory and models with incomplete information. These perspectives allow for the analysis of strategic behavior under bounded rationality and ambiguous signals, enabling a more nuanced understanding of agent dynamics in complex systems.

Strategic Games and the Iterated Prisoner's Dilemma

This section deepens the theoretical and conceptual groundwork for analyzing strategic interactions in environments characterized by interdependent decision-making. The discussion begins by distinguishing between strategic and non-strategic decisions, emphasizing that strategic behavior arises when the outcome of an agent's choice is contingent not only on their own preferences, but also on the concurrent or anticipated actions of others.(Malook, 2024) This interdependence introduces a layer of complexity that necessitates rational expectation formation and adaptive reasoning.

Strategic games provide a formal framework to model such scenarios, wherein agents—whether individuals, institutions, or artificial systems—must account for the strategies and likely responses of co-players. Drawing on both classical sources and contemporary contributions, this section explores how strategic reasoning applies across diverse agent typologies: natural (human), corporate, and artificial intelligence (AI).(Floridi, 2023) Each category is shown to interact with its environment based on distinct epistemic limitations and structural affordances, yet all are subject to the same formal strategic constraints derived from game-theoretic principles.

This foundation is used to objectively assess static, one-shot strategic interaction models. Such models are analytically tractable and essential for understanding equilibrium concepts like dominance and Nash equilibria, but they fail to capture real-world decision-making dynamics, which are often embedded in ongoing relationships, institutional structures, or evolving strategic landscapes. A classic strategic problem, the Prisoner's problem, illustrates the conflict between individual rationality and collective optimality. Without external enforcement or structure, rational agents will defy each other, even though mutual cooperation gives the largest joint benefit.

The transition from one-shot to repeated interaction marks a pivotal theoretical shift. The Iterated Prisoner's Dilemma (IPD) introduces a temporal dimension that fundamentally alters the strategic landscape.(Myerson, 1991) Repetition creates the possibility for history-dependent behavior, allowing strategies to condition their actions on past outcomes. This in turn enables mechanisms such as reciprocity, punishment, forgiveness, and signaling, which can sustain cooperation even in environments governed by self-interest.(Axelrod, 1984) Strategies like Tit-for-Tat, Grim Trigger, and Win-Stay Lose-Shift illustrate how simplicity, memory, and conditionality interact to produce robust cooperative equilibria.

Finally, this section highlights the role of uncertainty and imperfect monitoring in shaping strategic conduct. Reputational concerns, belief formation, and the availability of public or private signals all influence the credibility and enforceability of strategies. The theoretical implications of these factors are discussed in light of real-world analogues, from social norms and institutional design to AI-based agents in digital environments.

Repeated Games: Finite Horizon vs. Infinite Horizon

This section provides a comprehensive analysis of the strategic consequences arising from the temporal structure of repeated games, specifically contrasting games with a finite horizon—where the number of iterations is known and fixed—with those that feature infinite or uncertain horizons. The distinction is not merely technical; it reshapes the strategic landscape by altering how agents assess future payoffs, weigh cooperation versus defection, and anticipate the behavior of their counterparts.

In finitely repeated games, the principle of backward induction imposes a strong theoretical constraint. Since rational players are assumed to anticipate the future and reason recursively, defection becomes the dominant strategy in the final round, which then triggers a cascade of defection in all preceding stages. This logic, grounded in classical game-theoretic rationality, predicts the collapse of cooperation even in games where mutual cooperation would yield higher cumulative payoffs.(Cressman, 1996) While this result is formally robust, it hinges on strict assumptions such as perfect information, common knowledge of rationality, and deterministic expectations about game termination.

Infinite-horizon repeated games, or those with probabilistic continuation (i.e., players are uncertain whether the game continues after each round), offer a more favorable environment for the emergence and stability of cooperation. (Fudenberg & Maskin, 1986) The absence of a known final round removes the backward induction argument, allowing future consequences to influence present behavior. The concept of the "shadow of the future" becomes central: when agents sufficiently value long-term payoffs—typically modeled via a high discount factor—they are incentivized to maintain cooperation to avoid triggering future punishment.

Beyond theoretical conditions, recent refinements incorporate more realistic features such as imperfect monitoring, limited memory, and asymmetric information. For example, models with public signals (observable by all players) allow for punishment mechanisms to be enforced credibly, while private or noisy signals introduce complexity in belief formation and make cooperative equilibria more fragile. Researchers such as Hörner and Olszewski (2009) have shown that cooperation remains possible even under imperfect observability, as long as agents possess sufficient memory and discount future payoffs appropriately.

To capture the full implications of time structure, the section also contrasts the metrics used for payoff evaluation in finite and infinite games. While finite games typically sum stage payoffs or use cumulative utility, infinite-horizon models employ average payoffs or discounted utilities. These evaluation models influence the incentives for cooperation and defection, shaping strategic design and equilibrium behavior.

In summary, this section underscores that the temporal horizon of repeated games is a first-order determinant of strategic possibilities. Whereas finite games constrain cooperation through backward reasoning, infinite or uncertain horizons enable long-term strategies that reward cooperation and punish defection. This distinction is not only theoretically significant but also critical for designing robust agents in real-world multi-agent environments, where the duration and structure of interaction are rarely fixed or known in advance.

Strategies from Axelrod's Tournaments

This section offers a detailed examination of the empirical and theoretical contributions arising from Robert Axelrod's influential tournaments on the Iterated Prisoner's Dilemma (IPD).(Axelrod, 1984) These experiments marked a turning point in the study of cooperation by demonstrating that strategic behavior conducive to mutual benefit can emerge even in environments governed by individual rationality and self-interest. More than a methodological innovation, Axelrod's tournaments served as a bridge between theoretical game models and agent-based simulations, providing a robust empirical framework for testing hypotheses about strategy resilience and adaptability in repeated interactions.

Axelrod's first tournament (1980), which invited researchers from diverse fields to submit algorithmic strategies for the IPD, was conducted as a round-robin competition: each strategy played against all others—including itself—across a series of

200-round matches. The results were striking. Despite the wide variety in strategic complexity, memory usage, and behavior rules, the simplest strategy submitted—*Tit for Tat* (TFT), proposed by Anatol Rapoport—emerged as the overall winner. TFT's rule was elegant: begin with cooperation and then mimic the opponent's previous move. This strategy was not only effective in generating high payoffs but also robust against exploitation and retaliation.

The first tournament revealed several important principles. Strategies that combined niceness (never defecting first), retaliatory capacity (responding to defection), forgiveness (returning to cooperation), and clarity (being easily interpretable by opponents) performed better in repeated settings.(Axelrod & Hamilton, 1981) The emergent success of TFT led Axelrod to organize a second, more complex tournament involving 63 strategies, submitted by researchers from six countries. This second iteration introduced an evolutionary dimension: poorly performing strategies were removed and replaced by copies of better-performing ones, simulating population dynamics over generations.

Interestingly, the second tournament also revealed the limits of TFT and opened the field to more nuanced strategy design. While TFT remained highly competitive, it was occasionally outperformed in specific contexts by variations such as *Generous Tit for Tat* (GTFT), which occasionally forgives defection, or more complex strategies like *Tester* and *Tranquilizer*, which used probing moves to detect exploitable opponents. Moreover, the tournament structure provided fertile ground for the emergence of *extortionate* behaviors later formalized as *Zero-Determinant (ZD)* strategies. These strategies, although not explicitly included at the time, anticipated the idea of unilaterally enforcing linear payoff relationships.

From a theoretical standpoint, Axelrod's tournaments demonstrated that cooperation could be an evolutionarily viable outcome, even without centralized enforcement or communication between agents. The results challenged the dominance of strictly self-interested strategies (like Always Defect) and showed that behavior rules grounded in reciprocity and proportionality can sustain mutual benefit over long horizons.

In summary, this section illustrates how Axelrod's tournaments provided both a methodological blueprint and a theoretical paradigm for studying cooperation in repeated games. The insights derived from the success and failure of different strategies informed the subsequent development of evolutionary game theory, agent-based modeling, and reinforcement learning systems. As such, the Axelrod framework remains a cornerstone in the analysis of strategic behavior in multi-agent systems, bridging normative theory and empirical dynamics in a way that few other experiments in game theory have achieved.

Monte Carlo Simulation Methodology

This section introduces the computational simulation framework employed to evaluate strategic behavior in the Iterated Prisoner's Dilemma (IPD) under uncertainty.The simulation methodology presented here builds upon the framework and experimental approach proposed on our article: Axelrod First Tournament: examining certainty versus uncertainty about the end stage in repeated games.(Milencianu & Pop, 2023) Building on both classical insights and recent empirical approaches, the methodology leverages Monte Carlo simulation to overcome analytical limitations and systematically explore how strategies perform across stochastic interaction structures.

At its core, Monte Carlo simulation involves the generation of a large number of repeated random samples to estimate expected outcomes. In the context of repeated games, it enables the approximation of average payoffs, behavioral stability, and convergence tendencies under variable interaction lengths and opponent combinations. Unlike deterministic modeling, which assumes fixed game duration, this framework employs probabilistic termination rules, capturing more realistic conditions where agents do not know in advance when an interaction will end.(Metropolis & Ulam, 1949)

Each simulation trial involves pairing strategies in a round-robin format, where each strategy plays a probabilistically terminated IPD game against every other strategy in the pool. For robustness, each pairwise interaction is repeated multiple times (e.g., 1000 iterations) to ensure reliable convergence of average performance metrics. For every match, the sequence of decisions is determined by the internal logic of each agent—be it deterministic or stochastic—based on historical moves and the structure of the strategy (e.g., finite state machine, memory-one rule, probabilistic transition function).

The simulation environment is implemented in Python, leveraging efficient vectorized operations via NumPy. It supports both sequential simulations (for agent learning or adaptation tracking) and non-sequential sampling (for statistical summaries). Two schemes are employed:

- **Sequential Simulation:** Generates full play histories across rounds, enabling dynamic analysis of behavioral evolution within matches.

- **Non-Sequential Sampling:** Focuses on outcome distributions without recording step-wise transitions, thus reducing computational overhead.

Additionally, agents are modeled using either deterministic rules or stochastic finite-state automata, allowing for the comparison of strategies that differ in memory, randomness, and responsiveness. Memory-1 strategies, for example, base their actions solely on the outcome of the previous round, leading to a compact strategy space of 16 possible configurations. More complex strategies, such as those implemented as finite-state machines with multiple internal states and transitions, exhibit richer behavioral diversity but may also incur higher volatility.

To evaluate performance, each simulation produces both payoff-based and behavior-based metrics. These include:

- Average payoff per agent per match
- Frequency vectors of wins, losses, and ties
- Sensitivity to uncertainty (via payoff deviation across deterministic vs. stochastic horizons)

(Milencianu & Pop, 2023)

In methodological terms, the simulation framework developed here represents an extension of Axelrod's original experimental design. While Axelrod's tournaments used a fixed number of rounds (e.g., 200), the current approach generalizes the interaction length to be drawn from probability distributions (e.g., geometric, normal, uniform). This enables the comparison of identical strategies under conditions of certainty and uncertainty, facilitating controlled experimental investigation into temporal robustness. The simulation procedure follows a structured sequence of steps, as outlined below. For each ordered strategy pair (s_i, s_j), the simulation executes K independent trials. In each trial, the number of rounds T is sampled from a geometric distribution with continuation probability ω, and both strategies interact over T rounds of the Prisoner's Dilemma. Agent decisions are determined by their internal logic (e.g., deterministic rule, memory-based response, probabilistic transition), and the resulting payoffs are recorded and aggregated.

In conclusion, the Monte Carlo simulation methodology detailed in this section forms the empirical backbone of the thesis. It supports systematic testing of a wide array of strategies across probabilistic time structures and offers a rigorous platform for investigating how uncertainty affects strategic conduct. Beyond its utility for the Iterated Prisoner's Dilemma, the approach is broadly applicable to multi-agent decision environments where temporal unpredictability and decentralized adaptation play key roles.

Performance Metrics and Strategic Evaluation

This section presents the evaluative framework used to analyze and compare strategies in repeated games under both deterministic and uncertain time horizons. Rather than relying solely on average payoff—an often insufficient indicator of strategic robustness—the evaluation incorporates a multidimensional metric system designed to capture not only outcome performance, but also behavioral stability, volatility, and context sensitivity. This framework was applied systematically across all strategy match-ups in the simulation environment described in the previous section.(Milencianu & Pop, 2023)

For each strategy pair, the simulation produces two key empirical vectors: the frequency vector of wins, losses, and ties; and the mean payoff vector, averaged across all simulation runs. These are computed separately under both fixed and uncertain game lengths, allowing for direct comparison. Behavioral differences are captured using discrepancy metrics, which compute the absolute differences in observed outcomes (e.g., changes in win rate or payoff) between the two temporal conditions. A large discrepancy suggests that the strategy is structurally sensitive to the horizon condition, while small discrepancies may indicate robustness or mechanical consistency.

Another key aspect of the evaluative system is the classification of strategies based on discrepancy-driven typologies. Empirical results show that strategies tend to fall into three broad categories:

- **Robust strategies**, such as Tit-for-Tat and Grudger, which perform consistently across both deterministic and probabilistic horizons.
- **Adaptive strategies**, which exhibit moderate discrepancies but maintain rank and relative advantage across conditions, suggesting strategic plasticity.
- **Volatile strategies**, which display large shifts in payoff, win/loss patterns, or relative rankings between the two environments, often due to overfitting or structural rigidity.

To identify such patterns, rank-based indicators were used in addition to payoff measures. For instance, the *Rank Shift Index* tracks changes in a strategy's position within the global performance ranking between the fixed and uncertain conditions. Combined with payoff-based discrepancy values, this allows for the detection of strategies that are not just low-performing, but structurally unstable—potentially useful in adversarial settings, but unreliable in dynamic systems.

All performance metrics were encoded in matrix structures (payoff matrices, discrepancy matrices, volatility matrices), enabling both quantitative aggregation and qualitative visualization. Heatmaps, pie charts, and cooperation matrices were used to display global and local behavioral patterns, providing visual confirmation of the statistical findings.

In conclusion, the performance evaluation system presented in this section moves beyond traditional measures of success and introduces a multi-dimensional, discrepancy-sensitive approach tailored to environments with temporal uncertainty. It supports the thesis's broader goal of understanding strategic resilience, identifying emergent behaviors, and informing the development of adaptive agents capable of robust decision-making in stochastic multi-agent systems.

Conclusion

Chapter 2 synthesizes classical game-theoretic models with modern simulation techniques to offer a unified approach for analyzing strategic decision-making under temporal uncertainty. Through theoretical formalism, empirical validation, and methodological innovation, the chapter lays a rigorous foundation for the analysis of strategic robustness. The Monte Carlo-based simulation framework and multidimensional evaluation metrics provide the necessary tools for investigating emergent dynamics, evaluating strategic complexity, and identifying reliable behaviors in multi-agent systems.

Chapter 3

Summary of Chapter 3 : Experimental analysis of strategy behaviour in Determined vs. Uncertain Conditions

The content of this chapter has been partially published in: The Second Axelrod Tournament: A Monte Carlo Exploration of Uncertainty About the Number of Rounds in Iterated Prisoner's Dilemma. Studia Universitatis Babes,-Bolyai Oeconomica, 70(1), 67–82 (G. Pop et al., 2025)

Theoretical Insights into Determined and Uncertain Environments

This section explores how the strategic landscape shifts depending on whether the game is governed by a fixed horizon or by conditions of temporal ambiguity.

In determined, finite-horizon environments, players are fully aware of the number of rounds the game will last. This shared knowledge enables backward reasoning, a cornerstone of classical game theory.(Fudenberg & Tirole, 1991) Rational agents, anticipating that cooperation cannot be enforced in the final round, deduce that defection is the optimal final move. Knowing this, they project defection backward through each prior stage, ultimately concluding that mutual defection is the only rational path from the outset. This logic, while elegant and mathematically precise, often fails to align with observed behavior in real-world scenarios, where cooperation frequently emerges and persists even in finite settings. The backward induction solution presumes idealized conditions: perfect information, common knowledge of rationality, and the absence of noise or ambiguity.(Myerson, 1991)

Uncertain environments challenge these assumptions. In many real-world strategic contexts, agents do not know precisely when interaction will cease. Instead, each stage is followed by a continuation whose probability is unknown or defined probabilistically. This creates what is often referred to as the "shadow of the future"—the idea that future consequences shape present decisions, even in the absence of strict enforcement mechanisms.(Bó, 2002) Under uncertainty, agents have reason to invest in reciprocity, build reputations, and adopt strategies that condition current behavior on anticipated future interaction. Unlike the finite-horizon setting, where the last round defines strategic expectations, uncertainty allows cooperation to be enforced through implicit expectations, relational contracts, and informal norms.

From a systemic point of view, the uncertain environment better reflects how many socio-economic systems are set up. Agents don't often know how many times they will interact in marketplaces, negotiations, alliances, or online platforms. But over time, patterns of cooperation and punishment do start to show up and stay the same. This means that uncertainty is not just a problem to be fixed; it is a part of the structure that allows for more flexible and socially beneficial outcomes. The uncertain horizon promotes more complex interactions, such as the formation of norms, the growth of cooperation, and the establishment of stable behavioral equilibria, by removing the artificial limit of a definite endpoint.(Leyton-Brown & Shoham, 2008)

In light of these theoretical considerations, any empirical investigation into strategy robustness must take seriously the implications of temporal uncertainty. Evaluating strategies solely under fixed-horizon conditions would miss the crucial mechanisms by which cooperation emerges and endures. It would also underestimate the strategic capacities of agents who adapt behavior based not only on historical moves but also on projected futures and evolving expectations. This chapter builds on the conceptual foundations outlined here to develop a simulation-based methodology capable of testing strategic performance across both determined and uncertain settings.

Infinite-Horizon Dynamic Games and the Iterated Prisoner's Dilemma

This section integrates developments from infinite-horizon dynamic games into the study of IPD, with particular attention to anticipative control, robustness under uncertainty, and adaptive learning. These theoretical enhancements extend the explanatory power of the IPD framework by formalizing how agents can plan over uncertain futures, update beliefs dynamically, and design strategies that are resilient to variation in horizon length or opponent behavior.

One significant contribution in this regard is the introduction of *anticipative feedback mechanisms*, as proposed in the context of infinite-horizon Stackelberg games.(Chen & Zadrozny, 2002) Unlike the backward-induction logic of finite settings or the time-invariant strategies of classical equilibria, anticipative feedback structures enable players—particularly leaders in hierarchical games—to incorporate expectations about the future responses of opponents directly into their current decisions. When adapted to the IPD, this anticipatory reasoning can model generous leadership strategies: a player may initiate cooperation not because of immediate gain but because they foresee that such a signal alters the opponent's best response in future interactions. This aligns with phenomena observed in decentralized systems, where early cooperative gestures can shape long-term norms and expectations.

In parallel, research on *guaranteed-cost strategies*, developed by Gyurkovics and Takács (2005), introduces a structured framework for analyzing bounded-risk play in uncertain, adversarial settings.(Gyurkovics & Takács, 2005) These strategies ensure that players incur no more than a specified maximum cost, regardless of their opponent's actions or systemic noise. In the IPD context, such strategies embody cautious cooperation: agents are willing to cooperate, but only if the long-run risk of exploitation is provably constrained. This perspective adds a robustness layer to traditional retaliatory strategies like Grim Trigger or Tit for Tat, reframing them as instances of bounded-regret behavior in stochastic dynamic systems.

Another theoretical innovation, from D. Yeung and Petrosian stems from *information updating frameworks*, which model how players revise their expectations over time based on observed outcomes.(D. W. K. Yeung & Petrosyan, 2017) This breaks with the assumption of fixed beliefs or static payoff structures and instead models IPD interactions as evolving games with endogenous strategy revision. In environments characterized by ambiguous horizons or noisy signals, this mechanism allows players to gradually refine their strategic posture. In particular, belief-based triggers—where cooperation is sustained as long as the inferred probability of the opponent's cooperation exceeds a threshold—offer a rational foundation for trust-building even in the absence of full observability or deterministic horizons.

With the inclusion of these theoretical frameworks, the Iterated Prisoner's Dilemma (IPD) moves beyond its traditional role as a static benchmark, becoming instead a valuable framework for exploring dynamic strategic intelligence. The anticipative feedback approach demonstrates how forward-looking strategies can sustain long-term cooperation; guaranteed-cost strategies ensure robustness in the face of unpredictable adversaries; and information-updating mechanisms address adaptive behavior, bounded rationality, and evolutionary plausibility. Collectively, these contributions provide a comprehensive view of how strategic agents navigate uncertainty, offering insights into the practical application of theoretical models to real-world strategic interactions.

Experimental Setup and Results from Axelrod's First Tournament

This section builds directly on methodological foundations and findings of :Axelrod first tournament: Examining certainty versus uncertainty about the end stage in repeated games (Milencianu & Pop, 2023), providing a detailed presentation of the experimental architecture, evaluation procedures, and performance results.

The empirical investigation of cooperation in repeated games gained unprecedented momentum with the organization of Axelrod's First Tournament, a landmark experiment designed to test how different strategies perform in the Iterated Prisoner's Dilemma (IPD) under controlled simulation conditions. This section presents a structured overview of the experimental setup, key methodological choices, and emergent results from this foundational study, highlighting its influence on both the theoretical and computational modeling of strategic behavior.

The game configuration featured a deterministic horizon of 200 rounds per match, a choice that was both pragmatic—facilitating bounded simulations—and theoretically significant. With the number of rounds fixed and known to all agents, the tournament embedded a structural incentive for defection as the final stage approached. This setting placed cooperative strategies under strong pressure, offering an empirical test of whether and how cooperation could emerge despite the logic of backward induction.

Strategies varied widely in complexity and design philosophy. Some employed simple rule-based heuristics, such as *Always Defect*, *Always Cooperate*, or *Tit for Tat* (TFT). Others adopted more intricate logics incorporating memory, randomization, or probing behavior. Despite this heterogeneity, the tournament revealed robust empirical regularities that reshaped the understanding of cooperation.

The most striking outcome was the success of TFT, a minimalistic yet powerful strategy that began by cooperating and then mimicked the opponent's last move. TFT's performance was remarkable not only in terms of raw payoff but also in its ability to promote cooperation without being vulnerable to exploitation.

Beyond the success of TFT, the tournament offered valuable data on how different strategy types fared under repeated interaction. Strategies that were "nice"—never initiating defection—tended to perform better overall, particularly when paired with similar opponents. Aggressive or deceptive strategies, while potentially successful in specific matchups, often incurred losses when facing equally retaliatory counterparts. This result provided empirical support for the hypothesis that cooperation can be evolutionarily stable in environments characterized by repeated interaction and bounded rationality.

Methodologically, Axelrod's First Tournament introduced the use of simulations as a rigorous tool for comparative strategy evaluation. Rather than relying solely on analytical equilibria, the tournament format facilitated the observation of emergent phenomena such as cycles of cooperation and defection, lock-in effects, and strategic miscoordination. It also highlighted the role of initial conditions and early interaction patterns in determining long-term outcomes—a feature particularly salient in long-horizon games.

In sum, Axelrod's First Tournament demonstrated that cooperative behavior can emerge and persist even in environments where defection is the equilibrium prediction under classical assumptions. Through a transparent and replicable experimental design, it challenged the dominance of strictly self-interested models and introduced empirical metrics—such as average payoff, robustness to noise, and response to provocation—that remain central to the evaluation of strategies in repeated games. The insights generated continue to inform the design of strategic agents in decentralized systems, from economics to artificial intelligence.

Experimental Setup

This section presents the customized experimental infrastructure developed to evaluate strategic behavior in Iterated Prisoner's Dilemma tournaments under both deterministic and uncertain temporal conditions. While the Axelrod-Python library provided a robust foundation for simulating repeated interactions among agents, its original implementation lacked key features required for the analytical goals of this study—most notably, the ability to extract detailed match-level statistics such as the explicit identification of winners and the dynamic control of match durations via stochastic processes.

To address these limitations, a modular extension of the Axelrod tournament engine was implemented in Python. The core of this extended system enabled randomized control over both the number of matches played (`mean_m`, `dev_m`) and the number of turns per match (`mean_t`, `dev_t`), simulating infinite-horizon interaction structures using probabilistic time distributions. Furthermore, the number of independent simulation runs was parameterized to ensure statistical robustness across repeated tournament executions.

The implementation included key modules for defining agent behavior, orchestrating match sequences, and aggregating outcomes across stochastic trials. Each strategy played against every other, including itself, with match lengths sampled from a normal distribution, ensuring that players faced varying interaction spans without prior knowledge of termination. Results were stored as cumulative win/loss/draw statistics and normalized through automated aggregation routines. Dedicated CSV export functionality ensured structured data handling, and versioned output files facilitated reproducibility and long-term experiment tracking.

Visualization was an integral component of the experimental pipeline. Using `matplotlib` and `pandas`, the simulation outputs were transformed into intuitive visual summaries. Each strategy's performance profile was represented using pie charts—displaying win-loss-draw ratios across opponents—and cooperation matrices, which revealed deeper behavioral tendencies and symmetry or asymmetry in decision-making.

This custom tournament system not only enhanced flexibility in configuring game environments but also enabled the integration of stochastic match lengths, visual diagnostics, and expanded metrics. It allowed for fine-grained control over simulation parameters and supported rigorous comparative analyses between deterministic and uncertain scenarios. In doing so, the experimental setup provided a solid empirical foundation for subsequent sections, which analyze the strategic stability, volatility, and adaptive capacities of participating agents under varying informational and structural constraints.

Graphical Analysis of Strategy Sensitivity in the First Axelrod Tournament

This section presents a graphical and statistical examination of how strategies in the First Axelrod Tournament respond to structural changes in game duration—specifically, the contrast between determined (fixed-length) and uncertain (stochastically terminated) interaction environments. The aim is to provide empirical insight into the extent to which strategies are sensitive to temporal uncertainty, beyond what average payoffs or win rates alone can reveal. One of the core diagnostic tools is the *Discrepancy Matrix*, which captures the absolute payoff differences between the deterministic and stochastic tournament outcomes. This matrix provides a compact representation of how the strategic environment—defined solely by the certainty or uncertainty of match termination—affects bilateral interactions. Strategies exhibiting consistently low discrepancies across all match-ups are considered temporally robust, whereas those with large deviations are flagged as sensitive or potentially overfitted to specific game structures.

In addition to payoff variance, the *Cooperation Matrix* visualizes the frequency and symmetry of cooperative actions between all strategy pairs. This allows the identification of stable cooperation patterns, asymmetrical exploitation, or mutual defection regimes. Notably, certain strategies that perform well in the deterministic setting display marked decreases in cooperative engagement when match durations become uncertain, signaling a shift in strategic behavior prompted by ambiguity over the game horizon.

Taken together, the graphical analysis reveals that strategies vary not only in absolute performance but also in structural sensitivity. Some exhibit robust cross-context consistency (e.g., Tit for Tat, Grudger), while others show instability or performance degradation under uncertainty (e.g., more aggressive or probe-based strategies). These observations contribute to a deeper classification framework that accounts for volatility, adaptability, and strategic responsiveness to environmental ambiguity.

Experimental Setup and Results from Axelrod's Second Tournament

Axelrod's Second Tournament introduced several critical modifications to the initial experimental design, making it a landmark in the empirical study of strategic behavior under uncertainty. Building on the success and limitations of the first round-robin competition, this second iteration aimed to test the evolutionary viability and temporal resilience of strategies when exposed to a more dynamic and uncertain interaction structure.(Axelrod, 1984)

A substantial part of the data and comparative analysis presented in this section has been formally disseminated in our article *The Second Axelrod Tournament: A Monte Carlo Exploration of Uncertainty About the Number of Rounds in Iterated Prisoner's Dilemma* (G. Pop et al., 2025)

A major innovation was the implementation of an uncertain game horizon. Rather than fixing the number of rounds in advance, each interaction in the second tournament continued with a fixed probability ω, introducing stochastic termination and effectively simulating an indefinite game with probabilistic continuation. This adjustment neutralized backward induction and more closely approximated real-world strategic settings where agents cannot predict the exact endpoint of interactions.

The tournament featured 63 strategies, including both refinements of earlier submissions and entirely new entrants. Participants ranged from simple deterministic rules (e.g., Always Cooperate, Tit-for-Tat) to more complex designs featuring probing, randomization, or finite state machine logic. Strategies were blind to the identity or source code of their opponents and could only base their decisions on prior moves, thereby preserving the integrity of the interaction histories.

Results from the second tournament confirmed the competitive strength of reciprocal and forgiving strategies, particularly *Tit-for-Tat* and *Generous Tit-for-Tat*. These strategies demonstrated robustness across uncertain time structures, high payoff consistency, and minimal sensitivity to random termination. In contrast, more aggressive or exploitative strategies often suffered from early retaliation and failed to establish mutually beneficial equilibria.

It is important to note that the uncertain horizon revealed new aspects of strategic resilience. Some strategies that did really well in the first tournament had a lot of ups and downs when the lengths of interactions were unpredictable. On the other hand, other people were better able to adjust to the changing environment by using forgiveness, stochastic probing, or mixed-memory heuristics to stay competitive.

The rise of this kind of behavior shows how important it is to judge tactics not just by their overall return, but also by how well they can adapt to the environment's structure.

From a methodological perspective, the second tournament validated the use of probabilistic termination as a mechanism for eliminating artificial endgame behavior and generating richer patterns of cooperation and competition. It also served as a prototype for subsequent simulation-based research in repeated games, including the Monte Carlo framework developed in this thesis.

Results and Discussion

The empirical results derived from the extended tournament simulations provide a multifaceted view of how strategy performance is influenced by temporal uncertainty and evolutionary pressures. By systematically comparing outcomes from deterministic and stochastic settings, this section offers a comprehensive evaluation of strategy robustness, volatility, and adaptability.

One of the central findings is that strategies which perform well under fixed-horizon conditions do not necessarily maintain their advantage when the interaction length becomes uncertain. For example, while simple reciprocal strategies such as *Tit-for-Tat* remain competitive in both environments, their payoff rankings and behavioral patterns exhibit measurable deviations. This suggests that temporal uncertainty introduces a distinct structural shift in the strategic landscape—one that rewards flexibility and penalizes overly rigid or exploitative behavior.

The data also support the emergence of a typology of strategic profiles:

- **Robust strategies**, characterized by high average payoffs and low variance across both deterministic and stochastic conditions.
- **Adaptive strategies**, which adjust their behavior to temporal shifts, maintaining performance via probabilistic decision-making or forgiveness mechanisms.
- **Volatile strategies**, which are highly sensitive to game horizon changes and often exhibit overfitting to one specific context.

Key behavioral metrics—including win/loss ratios, cooperation frequencies, and payoff discrepancies—were synthesized into matrix-based visualizations. These tools revealed emergent behavioral asymmetries not apparent in aggregate payoff data. For instance, certain strategies showed cooperative tendencies under fixed settings but shifted to defect-heavy behavior under uncertainty, suggesting a reliance on endgame predictability for enforcing discipline or extracting gains.

From an evolutionary standpoint, strategies that incorporated stochastic elements or simple adaptive rules tended to persist longer in simulations with evolutionary feedback mechanisms. This indicates a fitness advantage for behavioral plasticity, especially in environments lacking precise temporal boundaries. Importantly, the probabilistic horizon model neutralized backward induction effects, creating room for trust-building and contingent cooperation to emerge organically.

The discussion additionally examines the methodological contributions of the simulation framework. The adaptation of Axelrod's model to include random match lengths, automated repetition, and extensive data tracking facilitated a more detailed analysis compared to the original studies. These innovations enhanced the understanding of how structural uncertainty influences strategic success, considering both immediate outcomes and long-term evolutionary sustainability.

The results highlight the necessity of incorporating uncertainty in the design and assessment of strategies for repeated games. Successful agents in temporally ambiguous environments typically integrate reciprocity, adaptability, and noise tolerance. This has significant implications for game theory, behavioral economics, and the development of resilient multi-agent systems in artificial intelligence and decentralized coordination contexts.

Chapter 4

Summary of Chapter 4: Autocovariance in the Allison Mixture – A Monte Carlo analysis of strategic decision-making in Repeated Games

Introduction to the Allison Mixture and Parrondo's Paradox

The results presented in this section build upon the empirical foundations introduced in our article *Autocovariance in the Allison Mixture: A Monte Carlo Analysis of Strategic Decision-Making in Repeated Games*(G. M. Pop et al., 2025), which outlined the theoretical motivations and demonstrated the applicability of the Allison Mixture formulation through selected simulation outcomes.

This section presents the Allison Mixture as a probabilistic framework for analyzing emergent strategic behavior in repeated games characterized by uncertainty. The model is situated at the convergence of stochastic process theory and behavioral game dynamics, drawing conceptual inspiration from Parrondo's paradox, which illustrates how alternating between losing strategies may result in a winning outcome. The Allison Mixture applies this paradoxical intuition to decision-making in strategic contexts, suggesting that adaptive combinations of individually neutral or suboptimal strategies may result in enhanced overall performance.

From a theoretical standpoint, the Allison Mixture challenges the conventional emphasis on static optimization and equilibrium-based rationality. Instead, it highlights how probabilistic transitions between behavioral modes—each suboptimal in isolation—can collectively produce robust and context-sensitive behaviors.(Harmer

& Abbott, 2002) This dynamic switching mechanism is modeled via a mixture of finite-state processes, governed by transition probabilities that encode the degree of structural uncertainty and memory dependence within the system.

Parrondo's paradox provides the foundational logic for this mechanism. Originally derived from physical systems exhibiting Brownian ratchet behavior, the paradox has been applied across various domains, from evolutionary biology to economics and statistical mechanics. Within game-theoretic contexts, it suggests that the strategic alternation between deterministic and stochastic rules can yield outcomes that are not achievable through any pure strategy alone.(Parrondo et al., 2000) The Allison Mixture captures this logic by embedding stochastic switching directly into the strategic fabric of repeated interactions, thereby generating behaviors that adapt dynamically to local conditions, opponent responses, and informational noise.

The significance of this approach lies in its capacity to model bounded rationality, signal-driven adaptation, and autocorrelated behavior without requiring complete information or infinite memory. By combining the probabilistic logic specific to Parrondo-type systems with formal representations of strategic interactions, the Allison Mixture offers a novel framework for investigating how cooperation, coordination, and competition can emerge within environments marked by ambiguity, asymmetry, and stochastic transitions. This theoretical foundation serves as a basis for developing simulation-based methods designed to empirically evaluate the mixture model in complex repeated games.

Allison Mixture

This section defines the Allison Mixture as a formal modeling construct that facilitates behavioral alternation between two complementary strategic structures. This section highlights that the alternation is not simply mechanical; rather, it is intended to produce autocorrelation in behavioral output, indicating a statistical dependence between successive decisions that adds temporal structure to the agent's behavior.(Gunn et al., 2014)

The Allison Mixture is defined by two fundamental properties. This approach integrates two baseline strategies into an integrated behavioral sequence via a probabilistic mechanism that determines the switching between them. (Harmer & Abbott, 1999) Secondly, it generates structured behavioral sequences whose characteristics cannot be simplified to those of either component independently. The alternation produces emergent patterns that represent the internal regularities of the mixture rule.

The section further clarifies that this construction does not require a learning algorithm, adaptation based on feedback, or memory of past plays. Instead, the behavioral correlation emerges from the structure of the alternation itself. This design makes the Allison Mixture especially suitable for simulation environments that aim to isolate the role of internal structural asymmetries in repeated games.

Application of Allison Mixture in Entry-Deterrence Game

This section discusses the application of the Allison Mixture in the context of an Entry-Deterrence game, in order to examine the impact of probabilistic alternation between base strategies on equilibrium behavior in an asymmetric strategic environment. The experimental setup examines the interaction between an Incumbent and a Potential Entrant, with payoffs structured to create strategic tension between deterrence and entry responses.(G. M. Pop et al., 2025)

Simulation results focus on how variations in the switching parameter of the mixture influence the deterrent effectiveness of the Incumbent. The analysis documents fluctuations in payoff structure, entry frequency, and conditional responses, confirming that Allison-type alternation introduces systematic behavioral patterns. These patterns differ substantially from those generated by purely random or deterministic strategies, suggesting that the mixture embeds structural dependencies that propagate over the course of interaction.

Furthermore, the section highlights that the Entry-Deterrence context serves as a testbed for evaluating the strategic expressiveness of the Allison Mixture beyond symmetric games. By incorporating the mixture mechanism into an asymmetric payoff environment, the analysis offers insight into how internally structured alternation may influence beliefs, provoke strategic miscoordination, or generate reputation-like effects, even in the absence of explicit signaling mechanisms.

Experimental Methodology for Monte Carlo Simulations Applied to the Allison Mixture

This section outlines the simulation methodology used to assess the strategic behavior of agents influenced by the Allison Mixture in repeated games. The analysis uses a

Monte Carlo simulation protocol designed to evaluate the statistical and behavioral characteristics of mixture-based strategies under varying interaction conditions. The simulation system facilitates the repeated execution of situations between agents, with their behavior governed by structured alternation rules.

The simulation framework is implemented in Python and configured to run large-scale experiments, allowing robust statistical conclusions through repeated sampling. Agent behavior is encoded via transition systems that reflect the alternation dynamics specified by the Allison Mixture. Results from these simulations are aggregated into summary statistics, including average payoffs, decision sequences, and structural descriptors of correlation patterns.

Experimental Methodology for Monte Carlo Simulations Applied to Allison Mixture

This part explains the official experimental protocol used to run Monte Carlo simulations to test how agents made with the Allison Mixture act. The method is meant to look at how structural mixing rules affect how agents act in different simulated game scenarios. This part is all about how to find emergent statistical characteristics in decision sequences and the payoff distributions that go with them.

The experimental design employs a repeated random sampling scheme, wherein each simulation trial consists of a large number of iterations involving repeated games governed by the same structural alternation rules. Each agent using the Allison Mixture alternates between two base strategies, and the simulation records the outputs over varying match lengths. These simulations are intended to produce robust distributions from which average tendencies, variance, and autocorrelation properties can be inferred.

Agent behavior is implemented through programmable logic capable of enforcing the switching mechanism specified in the Allison Mixture. The internal state of each agent transitions based on probabilistic switching rules, and their resulting moves are recorded round by round. The simulation data is then analyzed for regularities such as phase patterns, periodicity, and fluctuation ranges.

Experimental Results Obtained through Monte Carlo Simulations

This section offers empirical findings derived from Monte Carlo simulations utilizing agents driven by the Allison Mixture model. The analysis examines essential statistical metrics obtained from comprehensive repeated-game experiments, intending to evaluate how probabilistic switching structurally affects agent behavior and payout stability.

The analysis highlights that, under a wide range of parameter configurations, Allison-type agents produce sequences with visible regularities and internal structure, departing from purely stochastic behavior. Notably, the simulations reveal emergent autocorrelation within action patterns, suggesting that the alternation mechanism embedded in the mixture introduces memory-like effects even when the base strategies are individually memoryless.(G. M. Pop et al., 2025)

The section also documents heatmaps and summary statistics, which illustrate that mixture-induced behavior diverges systematically from purely random or deterministic strategies. These patterns underscore the model's capacity to embed functional unpredictability while retaining coherence over time.

The experimental results confirm that the Allison Mixture generates distinctive behavioral characteristics, such as autocorrelation, payout consistency, and environment-sensitive adaptability—attributes that differentiate it from traditional strategies in repeated game scenarios.

Conclusions on Autocovariance Analysis in Allison Mixture through Monte Carlo Simulations for Strategic Decision-Making

Building on the simulation data, this section highlights the main results of the autocovariance analysis, evaluating whether the probabilistic alternation of strategies leads to temporal patterns in behavior and what that means for the stability of strategic responses.

The Monte Carlo simulations demonstrated a persistent non-zero autocovariance in decision series produced by Allison-type agents, especially under conditions of balanced switching probability. This indicates that while individual strategies may not directly encode memory, their systematic alternation generates temporal regularity that can be utilized in repeated games. Autocovariance indicates an underlying structure in the decision-making process, permitting agents to demonstrate behavior that is not entirely random or completely deterministic.

In conclusion, our research suggests substantial opportunities for applying the Allison Mixture concept beyond game theory, extending into interdisciplinary fields such as biology, finance, and computer science, where the analysis and interpretation of complex data sequences exhibiting relevant autocovariance are critical. Future studies should further explore these directions, deepening empirical and theoretical analyses to maximize the utility and robustness of the Allison Mixture model in advanced strategic decision-making contexts.

Chapter 5

General Conclusions and Future Research Directions

This thesis provides a thorough framework for examining strategic behavior in repeated games with various degrees of uncertainty about game duration. This work seeks to connect classical game-theoretic principles with modern agent-based decision-making and evolutionary robustness through the combination of formal theoretical foundations, algorithmic modeling, and Monte Carlo simulation.

The first chapter has given the formal and conceptual framework needed to replicate strategic interaction in repeated settings. Drawing on fundamental work in non-cooperative game theory, the chapter has underlined the significance of Nash equilibrium, mixed strategies, and expected utility as main instruments for projecting rational conduct. Applying these ideas to dynamic environments, such the Iterated Prisoner's Dilemma (IPD), demonstrated that considering future outcomes alters the reasoning of strategic interaction significantly. Analysis of infinite-horizon models and the consequences of the folk theorem showed that, given agents value future payoffs sufficiently, cooperation can result as an equilibrium outcome.

The investigation of discrepancy metrics provides a principled foundation for the identification of structurally strong strategies, therefore facilitating not only performance comparison in stable conditions but also strategic flexibility evaluation under temporal unpredictability. The transition from fixed-horizon to uncertain-horizon game environments demonstrates that behavioral resilience is not an absolute superior but rather depends on the capacity of a strategy to sustain constant performance throughout changing conditions. This realization helps to reinforce the general conclusion that evaluating the long-term adaptation and sustainability of decision rules in repeated games depends critically on sensitivity to the temporal structure of interaction.

Deeper understanding of agent adaptation under uncertainty is made possible by including structural performance measurements into the study of recurrent interactions. This approach does not merely interpret simulation outcomes but establishes a principled framework for the design of adaptive agents capable of reconfiguring their strategies in response to volatility, asymmetry, and dynamic game conditions. Therefore, it creates fresh directions for investigation in resilient policy design, evolutionary game theory, and reinforcement learning.

These theoretical models were empirically investigated in the second chapter under experimental baselines developed from the famous Axelrod competitions. Comparative simulation of strategy performance in contexts with known and unpredictable number of rounds revealed essential behavioral differences. Sometimes more adaptive or forgiving variations outperformed strategies like Tit for Tat, which performed consistently in deterministic environments under uncertainty. On the other hand, strict plans devoid of probabilistic elements often did not work in settings marked by horizon uncertainty. Using discrepancy measures—such as payback volatility and variance in win-loss ratios—the statistical research indicated which agents were particularly sensitive to structural changes and which were strong under all situations.

Importantly, the analysis demonstrated that agents with deterministic behavior often exhibited emergent sensitivity to game duration, even without explicit temporal reasoning. This finding supports the idea that structural features of repeated interaction can induce complex outcomes in agent behavior, particularly when agents rely on conditional cooperation or memory-based responses.

Furthermore, this study highlights the importance of constructing evaluation frameworks that capture the stochastic nature of real-world decision-making. The contrast between fixed-horizon and uncertain-horizon simulations underscores how rigid experimental assumptions can obscure or distort the understanding of what constitutes optimal strategic behavior.

An further significant insight is the potential role of hybrid methodologies that integrate adaptive learning systems competent in handling signal noise, not known intentions, or unexpected environmental alterations with the resilience of reactive heuristics, such as *Grudger* or *Tit for Tat*. These hybrid designs may combine the trust-enhancing advantages of collaboration with the defensive strategies necessary for hostile contexts, so functioning as more precise representations of actual decision-making entities.

This study reaffirms the educational and methodological value of simulated tournaments as instruments for evaluating strategic interaction. Inspired by Axelrod's original design, such competitions remain relevant as controlled experimental environments in which hypotheses regarding agent behavior—whether artificial or human—can be tested, refined, and generalized.

Furthermore, the findings highlight that addressing uncertainty in strategic contexts requires more than algorithmic sophistication. It demands a clear conceptual understanding of which behavioral traits—such as adaptability, forgiveness, or learning capacity—truly underpin strategic success, laying the groundwork for agents that are not only performant, but also interpretable, resilient, and ethically aligned.Moreover, the results emphasize the evolving role of uncertainty as a selective force in strategic environments. Predictability, while beneficial under deterministic assumptions, becomes a liability in volatile contexts, where adaptability constitutes a critical asset. Agents that rigidly follow predefined rules may perform well in idealized settings, yet they often fail when exposed to the complexity and variability inherent in real-world interactions.

The results suggest that future design of strategies should consider the dynamic interaction between memory depth and response variance. While shallow-memory agents may react fast to local changes, they generally ignore cumulative dynamics; conversely, agents with deeper recall and evaluative systems are more suited to identify recurrent behavioral patterns. Real-time calibration of this trade-off might open interesting directions for building more intelligent and responsive decision-making systems.

In Chapter 3, the research focused on investigating the Allison Mixture formula, starting from the theoretical premise that, while elegant in its probabilistic structure, it necessitates systematic empirical validation within complex and dynamically evolving environments. Particular attention was given to contexts involving stochastic transitions, where the inherent uncertainty challenges traditional modeling assumptions. The principal objective was to evaluate the formula's viability in practical Monte Carlo simulation settings and to explore its potential for informing the design of advanced artificial intelligence agents capable of anticipating and responding to opponents' strategic behavior in repeated interactions.

To accomplish this, we built a strong experimental design in which repeated simulations were carried out throughout a wide spectrum of parameters and interaction distances. To guarantee the statistical dependability of the outcomes, every game configuration was run hundreds of times, thereby reducing the impact of random fluctuations and improving the correctness of computed performance measurements. This configuration made it possible to estimate important indicators including the mean and standard deviation of autocovariance and autocorrelation throughout several regimes consistently.

Beyond validating the theoretical formulation, the analysis revealed that the Allison Mixture holds considerable promise for designing adaptive strategic agents. In environments where conventional techniques often fail to capture subtle interdependencies and evolving behavioral signals, the Allison Mixture facilitates the emergence of higher-order structures by leveraging controlled randomness. Its ability to synthesize patterns through probabilistic alternation positions it as a valuable tool for modeling and navigating uncertainty in multi-agent systems. As such, this study provides both a methodological and conceptual contribution to the strategic modeling literature, offering a data-driven foundation for future applications in artificial intelligence and computational game theory.

The application of Monte Carlo methods throughout the thesis has proven essential for managing the high-dimensional, stochastic nature of repeated games. This computational approach enabled the approximation of expected outcomes and the measurement of strategic volatility across thousands of simulated interactions. The findings suggest that simulation is not merely a tool for empirical illustration, but a fundamental component in the study of strategic dynamics where analytical solutions are intractable.

The methodological rigor underlying the Allison Mixture simulations—anchored in large-scale iteration and high-performance computation—ensures that the observed outcomes are not artifacts of noise, but statistically sound reflections of the modeled dynamics, thereby reinforcing the empirical validity of the approach.

Despite the encouraging outcomes, the investigation revealed limitations regarding the generalizability of the Allison Mixture framework. In particular, the performance of the formula was shown to be sensitive to the structural characteristics of the underlying game, limiting its immediate applicability across diverse strategic settings. While many scenarios aligned with theoretical predictions, the absence of consistent results across all conditions indicates the need for further exploration. Future research should expand the scope of analysis by systematically varying initial parameters and testing the Allison Mixture in a broader array of game structures to better delineate the boundaries of its effectiveness.

These results suggest that the Allison Mixture framework holds promise for enhancing adaptive decision-making in dynamic environments, offering a pathway toward strategies that can better anticipate and respond to complex patterns of interaction.

All taken together, the thesis advances three main aspects of the body of knowledge on repeated games. First, it underlines how cooperation and strategic complexity develop under relaxed assumptions about rationality and information, therefore confirming the validity of classical strategic thinking. Second, it develops a simulation-based approach for assessing agent performance under both deterministic and stochastic horizons. At last, it presents a methodology using measures derived from real interaction data for evaluating structural robustness and adaptation.

A key contribution of this research lies in the formal integration of uncertainty into the strategic evaluation of repeated games, an element often overlooked in classical analyses. By developing and applying a robust Monte Carlo framework to both the Axelrod tournaments and the Allison Mixture model, the research demonstrates how probabilistic horizon structures reshape agent behavior, expose vulnerabilities, and reveal emergent adaptive capacities. This dual exploration not only bridges experimental game theory with stochastic modeling, but also provides a replicable methodology for testing strategy resilience under real-world-like unpredictability. The proposed framework introduces a flexible simulation protocol that can accommodate both deterministic and stochastic structures, enabling more nuanced and context-sensitive assessments of agent performance. Ultimately, the research opens new directions for designing intelligent agents capable of learning, adapting, and sustaining cooperation in dynamic and uncertain environments.

This study introduces a unified experimental framework capable of systematically contrasting agent behavior under fixed and uncertain interaction lengths, offering new insights into strategic adaptability and volatility. By extending this framework to cover both classical behavioral models from the Axelrod tournaments and the statistical dynamics of the Allison Mixture, the research provides a methodological blueprint for analyzing how uncertainty alters incentive structures and coordination outcomes. This dual application underscores the originality of the approach, which goes beyond static equilibrium analysis to uncover how structural features of interaction—such as round unpredictability—drive long-term cooperation, strategic drift, or convergence to robust norms.

Selective Bibliography

Abbott, D. (2010). Asymmetry and disorder: A decade of parrondo's paradox. *Fluctuation and Noise Letters*, *9*(1), 129–156.

Abreu, D., & Pearce, D. (2007). Bargaining, reputation, and equilibrium selection in repeated games with contracts. *Econometrica*, *75*(3), 653–710. https://doi.org/10.1111/j.1468-0262.2007.00762.x

Akin, E. (2016). The iterated prisoner's dilemma: Good strategies and their dynamics. In A. Idris (Ed.), *Ergodic theory* (pp. 77–107). De Gruyter.

Allison, A., Pearce, C., & Abbott, D. (2007). Finding keywords amongst noise: Automatic text classification without parsing. *Proceedings of SPIE - The International Society for Optical Engineering*, *6601*. https://doi.org/10.1117/12.724655

Asmussen, S., & Glynn, P. W. (2007). *Stochastic simulation: Algorithms and analysis* (Vol. 57). Springer. https://doi.org/10.1007/978-0-387-69033-4

Axelrod, R. (1984). *The evolution of cooperation*. Basic Books.

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396. https://doi.org/10.1126/science.7466396

Ben-Porath, E. (1990). The complexity of computing a best response automaton in repeated games with mixed strategies. *Games and Economic Behavior*, *2*(1), 1–12.

Binmore, K. (1992). *Fun and games: A text on game theory*. D. C. Heath; Company.

Bó, P. D. (2002). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *Working Paper, Brown University, No. 2002-21*.

Brunauer, R., Löcker, A., Mayer, H. A., Mitterlechner, G., & Payer, H. (2007). Evolution of iterated prisoner's dilemma strategies with different history lengths in static and cultural environments. *Proceedings of the 2007 ACM Symposium on Applied Computing (SAC)*, 2058–2065. https://doi.org/10.1145/1244002.1244163

Bukhori, S., Hariadi, M., Purnama, I. K. E., & Purnomo, M. H. (2013). Parrondo's paradox based strategies in the serious game of rtgs using forest fire model. *The International Arab Journal of Information Technology*, *10*(5), 428–435.

Chen, B., & Zadrozny, P. A. (2002). An anticipative feedback solution for the infinite-horizon, linear-quadratic, dynamic, stackelberg game. *Journal of Economic*

Dynamics and Control, *26*(9–10), 1397–1416. https://doi.org/10.1016/S0165-1889(01)00077-X

Cheong, K. H. (2019). Adaptive therapy and parrondo's paradox: A model-based perspective. *BioSystems*, *178*, 10–17.

Cressman, R. (1996). Evolutionary stability in the finitely repeated prisoner's dilemma game. *Journal of Economic Theory*, *68*, 234–248.

Doucet, A., de Freitas, N., & Gordon, N. (Eds.). (2001). *Sequential monte carlo methods in practice*. Springer. https://doi.org/10.1007/978-1-4757-3437-9

Eckhardt, R. (1987). Stan ulam, john von neumann, and the monte carlo method. *Los Alamos Science*, *15*, 131–137.

Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.

Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, *54*(3), 533–554.

Fudenberg, D., & Tirole, J. (1991). *Game theory*. MIT Press.

Gunn, L. J., Allison, A., & Abbott, D. (2014). Allison mixtures: Where random digits obey thermodynamic principles. *Fluctuation and Noise Letters*, *13*(03), 1450015. https://doi.org/10.1142/S0219477514500152

Gunn, L. J., Chapeau-Blondeau, F., Allison, A., & Abbott, D. (2016). Towards an information-theoretic model of the allison mixture stochastic process. *Journal of Statistical Mechanics: Theory and Experiment*, *2016*(5), 054041. https://doi.org/10.1088/1742-5468/2016/05/054041

Gyurkovics, É., & Takács, T. (2005). Guaranteeing cost strategies for infinite horizon difference games with uncertain dynamics. *International Journal of Control*, *78*(8), 587–599. https://doi.org/10.1080/00207170500093978

Harmer, G. P., & Abbott, D. (1999). Parrondo's paradox. *Statistical Science*, *14*(2), 206–213. http://www.jstor.org/stable/2676740

Harmer, G. P., & Abbott, D. (2002). A review of parrondo's paradox. *Fluctuation and Noise Letters*, *2*(02), R71–R107.

Imhof, L. A., & Nowak, M. A. (2007). Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1621), 173–179. https://doi.org/10.1098/rspb.2006.3720

Koh, J. M., & Cheong, K. H. (2020). Generalized solutions of parrondo's games. *Advanced Science*, *7*(24), 2001126.

Lesne, A. (2014). Shannon entropy: A rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics.

Mathematical Structures in Computer Science, *24*(3), e240311. https://doi.org/10.1017/S0960129512000783

Leyton-Brown, K., & Shoham, Y. (2008). *Essentials of game theory: A concise multidisciplinary introduction*. Morgan; Claypool.

Lindsten, F., & Schön, T. B. (2013). Backward simulation methods for monte carlo statistical inference. *Foundations and Trends® in Machine Learning*, *6*(1), 1–143.

Mailath, G. J., & Samuelson, L. (2006). *Repeated games and reputations: Long-run relationships*. Oxford University Press.

Malook, S. (2024). The prisoner's versus pardoner's dilemmas: A juxtaposition of two strategic decision-game theoretic approaches in social sciences. *Journal of Social & Organizational Matters*, *3*(3), 52–74. https://doi.org/10.56976/jsom.v3i3.21

Metropolis, N., & Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, *44*(247), 335–341.

Milencianu, M., & Pop, G. (2023). Axelrod first tournament: Examining certainty versus uncertainty about the end stage in repeated games. *Virgil Madgearu Review of Economic Studies and Research*, *16*, 57–73. https://doi.org/10.24193/RVM.2023.16.99

Mittal, V., & Huang, Y.-P. (2024). Parrondo's paradox in quantum walks with inhomogeneous coins. *arXiv preprint arXiv:2407.16558*. https://arxiv.org/abs/2407.16558

Myerson, R. B. (1991). *Game theory: Analysis of conflict*. Harvard University Press.

Nachbar, J. H. (2005). Beliefs and learning in repeated games. *Econometrica*, *73*(2), 459–480. https://doi.org/10.1111/j.1468-0262.2005.00578.x

Nash, J. (1950). Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA*, *36*, 48–49.

Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, *364*, 56–58. https://doi.org/10.1038/364056a0

Parrondo, J. M. R., Harmer, G. P., & Abbott, D. (2000). New paradoxical games based on brownian ratchets. *Physical Review Letters*, *85*(24), 5226–5229.

Pop, G., Milencianu, M., & Pop, A. (2025). The second axelrod tournament: A monte carlo exploration of uncertainty about the number of rounds in iterated prisoner's dilemma. *Studia Universitatis Babeș-Bolyai Oeconomica*, *70*(1), 67–82. https://doi.org/10.2478/subboec-2025-0004

Pop, G. M., Tătar, F.-A., Vlădărean, M.-M., & Marian, C. G. (2025). Autocovariance in the allison mixture: A monte carlo analysis of strategic decision-making in repeated games. *Virgil Madgearu Review of Economic Studies and Research*, *18*(1), 143–164. https://doi.org/10.24193/RVM.2025.18.130

Rubinstein, A. (1998). *Modeling bounded rationality*. MIT Press. https://mitpress.mit.edu/9780262681001/modeling-bounded-rationality/

Schelling, T. C. (1980). *The strategy of conflict*. Harvard University Press. https://www.hup.harvard.edu/books/9780674840317

Stewart, A. J., & Plotkin, J. B. (2012). Extortion and cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, *109*(26), 10134–10135. https://doi.org/10.1073/pnas.1208087109

Stewart, A. J., & Plotkin, J. B. (2013). From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, *110*(38), 15348–15353.

von Neumann, J., & Morgenstern, O. (1974). *Theory of games and economic behavior* (3rd). Princeton University Press.

Weinstein, J., & Yildiz, M. (2016). Reputation without commitment in finit ely repeated games. *Theoretical Economics*, *11*(1), 157–185.

Yeung, D., & Petrosian, O. (2017). Infinite horizon dynamic games: A new approach via information updating. *International Game Theory Review*, *19*(4), 1750026. https://doi.org/10.1142/S0219198917500268

Yeung, D. W. K., & Petrosyan, L. A. (2017). Infinite horizon dynamic games: A new approach via information updating. *Dynamic Games and Applications*, *7*(4), 581–608. https://doi.org/10.1007/s13235-017-0212-7