Quantum-Fuzzy Relational Dynamics in the Iterated Prisoner's Dilemma: Competitive Prediction with Stronger Recovery, Coupling, and Resilience Effects

Murilo Salem[1], **Daniel Pontes**[1], **Luísa Böhm**[1]
Henrique dos Reis [1], **Anderson Priebe Ferrugem**[1]

[1]Centro de Desenvolvimento Tecnológico - Universidade Federal de Pelotas (UFPel)
96010-610 – Pelotas – RS – Brazil

{mcsalem, dhspbarretos, lcbohm, hdreis, ferrugem}@inf.ufpel.br

Abstract. *Modeling socially interdependent decision processes remains a major challenge in computational intelligence, particularly when behavior is influenced by emotional ambiguity, inter-agent coupling, and recovery after disruption. This paper introduces a quantum-fuzzy benchmark for emotion-aware relational dynamics in the Iterated Prisoner's Dilemma. Emotional variables are represented through fuzzy membership degrees and encoded into quantum-fuzzy relational states, enabling controlled analysis of cooperation, betrayal, recovery, coupling, and robustness under perturbation. We compare three settings: a classical fuzzy baseline, a quantum-fuzzy model, and a classifier-augmented quantum-fuzzy variant. On the predictive benchmark, the quantum-fuzzy model is competitive and slightly outperforms the classical baseline, achieving the best F1-score (0.8153 vs. 0.8039) and accuracy (0.7672 vs. 0.7610), while the augmented variant performs similarly. The main gains, however, appear in relational dynamics. Quantum-fuzzy models recover cooperation much faster after betrayal, reduce collapse rates, respond more strongly to inter-agent coupling, and show higher resilience under relational shocks, with recovery time dropping from 5.91 to 2.28 and 1.43, and resilience increasing from 0.0466 to 0.2256. Overall, the results suggest that quantum-fuzzy encoding contributes less through predictive gains and more through improved relational adaptation in emotionally interdependent environments.*

1. Introduction

Decision-making in socially interdependent environments depends not only on immediate incentives, but also on relational memory, disruption, reciprocity, and recovery over repeated interaction [Axelrod and Hamilton 1981, Nowak and Sigmund 1992, Rodrigues et al. 2017]. This makes such settings challenging for computational intelligence, particularly when the objective is not merely to predict isolated actions, but to model how cooperative and conflictive patterns emerge, deteriorate, and reorganize over time [Sandholm 2010]. Classical fuzzy systems are well suited to represent gradual and ambiguous emotional states [Zadeh 1965, Zadeh 1996], yet they may remain limited when the core phenomenon depends strongly on inter-agent dependence rather than only on local decision tendencies.

In this context, fuzzy logic provides an interpretable foundation for modeling affective variables such as trust, stress, empathy, and frustration, which are naturally continuous and overlapping rather than crisp [Kosko 1990]. However, purely fuzzy formulations do not explicitly encode joint relational structure between agents. This motivates the use of a quantum-fuzzy formulation, not as a claim of universal predictive superiority, but as a compact representational mechanism for coupling-sensitive interaction dynamics [Busemeyer and Bruza 2012]. The central question of this study is therefore whether quantum-fuzzy encoding can improve the modeling of relational adaptation without sacrificing predictive competitiveness.

To investigate this question, we adopt the Iterated Prisoner's Dilemma (IPD) as a controlled benchmark for socially inspired decision-making [Axelrod and Hamilton 1981]. The IPD offers a compact and reproducible setting in which cooperation, betrayal, collapse, and recovery can be studied under repeated interaction [Nowak 2006], while avoiding stronger behavioral claims that would be difficult to validate in real human settings. Within this benchmark, we compare three configurations: a classical fuzzy baseline, a quantum-fuzzy decision model, and a quantum-fuzzy representation augmented with a downstream classifier. This design allows us to distinguish between the contribution of fuzzy emotional inference, the effect of quantum-fuzzy relational encoding, and the incremental value of an additional learned policy.

The main premise of the paper is that the contribution of the quantum-fuzzy formulation should be evaluated not only through pointwise prediction, but also through its ability to represent relational dependence under coupling, disruption, and recovery. The contributions of this paper are threefold: *(i)* we define a compact and reproducible benchmark for emotion-aware relational dynamics in the IPD; *(ii)* we compare classical fuzzy and quantum-fuzzy models across predictive and dynamic relational criteria, including betrayal recovery, coupling sensitivity, robustness under shocks, and temporal organization; and *(iii)* we show that the strongest gains of the quantum-fuzzy framework emerge in relational adaptation rather than in broad improvements in isolated action prediction.

Within the Brazilian SBC ecosystem, prior work has explored fuzzy and multiagent approaches for socially grounded interaction, including BDI-fuzzy agents for subjective social exchanges, trust transfer based on reputation and dependence, and fuzzy reputation models in multiagent settings [Farias et al. 2013, Rojas et al. 2014, Rodrigues et al. 2017]. These studies reinforce the relevance of fuzzy representations for socially interdependent interaction processes, but they do not investigate quantum-fuzzy state encoding or controlled benchmarks centered on recovery, coupling sensitivity, and relational resilience.

2. Proposed Framework

The proposed framework evaluates whether a quantum-fuzzy representation provides a more expressive model of socially interdependent decision dynamics than a classical fuzzy baseline. All compared strategies share the same emotional state variables and the same relational update process in the Iterated Prisoner's Dilemma (IPD), differing only in the decision layer. This design allows the benchmark to isolate representational effects from differences in the underlying interaction dynamics.

2.1. Emotional State Representation

Each agent is represented by an emotional state vector

$$\mathbf{e}_i^{(t)} = [\tau_i^{(t)}, \sigma_i^{(t)}, \eta_i^{(t)}, \phi_i^{(t)}],$$

where τ, σ, η, and ϕ denote trust, stress, empathy, and frustration, respectively, with all variables normalized to $[0, 1]$.

Each variable is fuzzified into three linguistic terms—*low*, *medium*, and *high*—using trapezoidal and triangular membership functions. Fuzzy inference follows a Mamdani scheme and produces two latent scores, cooperation and conflict, which are defuzzified by the centroid method. Representative rules include: *(i)* high trust and high empathy imply high cooperation, *(ii)* high stress and high frustration imply high conflict, and *(iii)* low trust and high frustration imply high conflict. The final fuzzy score is defined as

$$s_{\text{final}} = \text{clip}\Big(0.5 + 0.5(s_{\text{coop}} - s_{\text{conf}}), 0, 1\Big),$$

where s_{coop} and s_{conf} are the defuzzified cooperation and conflict scores.

2.2. Classical Fuzzy Baseline

The classical baseline maps the current emotional state and a compact history summary into a cooperation probability. In addition to the fuzzy score, the model incorporates recent interaction information, including previous actions, mutual cooperation/defection balance, and cumulative payoff asymmetry, aggregated into a history signal $h^{(t)}$.

The final decision is computed through a logistic transformation:

$$\ell_{\text{class}}^{(t)} = 4.0(s_{\text{final}}^{(t)} - \theta) + \lambda_h h^{(t)},$$

$$P_{\text{class}}(C \mid \mathbf{e}_i^{(t)}, h^{(t)}) = \sigma(\ell_{\text{class}}^{(t)}),$$

where $\theta = 0.5$ is the cooperation threshold and $\sigma(\cdot)$ is the sigmoid function. This model serves as the main interpretable reference throughout the benchmark.

2.3. Quantum-Fuzzy Encoding

The quantum-fuzzy strategy preserves the fuzzy emotional layer, but maps its outputs into a compact quantum-inspired relational representation. For each agent, cooperation and conflict tendencies are derived from the fuzzy scores and encoded as single-qubit states:

$$|\psi(\mu)\rangle = \sqrt{1-\mu}\,|0\rangle + \sqrt{\mu}\,|1\rangle, \qquad \theta(\mu) = 2\arcsin(\sqrt{\mu}).$$

The implemented circuit uses four qubits: cooperation and conflict qubits for each of the two agents. Inter-agent dependence is introduced through controlled rotations modulated by a coupling parameter κ, allowing the representation to capture interaction-sensitive relational structure [Nielsen and Chuang 2010]. In the reported experiments, the default implementation used statevector simulation, continuous coupling, and a four-qubit encoding, consistent with standard quantum circuit simulation frameworks [Qiskit Development Team 2021].

After state construction, the model extracts local, partner, and joint decision features, which are combined with the fuzzy score and the history signal to produce the quantum-fuzzy raw score. This formulation follows quantum-inspired approaches to decision modeling, where interaction effects are encoded in composite state representations rather than treated as independent variables [Busemeyer and Bruza 2012].

$$r_{\mathrm{QF}}^{(t)} = w_\ell(a_{\mathrm{coop}} - a_{\mathrm{conf}}) + w_p(b_{\mathrm{coop}} - b_{\mathrm{conf}}) + w_j(j_{\mathrm{coop}} - j_{\mathrm{conf}}) + 0.70(s_{\mathrm{final}}^{(t)} - 0.5) + w_h h^{(t)}.$$

The final cooperation probability is then given by

$$P_{\mathrm{QF}}(C \mid \mathbf{e}_i^{(t)}, h^{(t)}) = \sigma(r_{\mathrm{QF}}^{(t)}).$$

This formulation is intended as a representational extension of the classical fuzzy model, rather than as a claim of hardware-level quantum advantage. Its role is to encode interaction-dependent relational structure in a compact and coupling-sensitive way.

2.4. Dynamic Relational Update

All strategies operate over the same repeated relational process. After each round, the agents' emotional states are updated according to interaction markers derived from the IPD outcome, including mutual cooperation, betrayal, conflict, reciprocity, hostility, and other defection. In generic form, each emotional variable is updated as

$$x_i^{(t+1)} = \mathrm{clip}_{[0,1]}\Big(x_i^{(t)} + \sum_k \omega_k\, m_k^{(t)}\Big),$$

where $x_i^{(t)} \in \{\tau_i^{(t)}, \sigma_i^{(t)}, \eta_i^{(t)}, \phi_i^{(t)}\}$, $m_k^{(t)}$ denotes interaction markers at round t, and ω_k are fixed update coefficients.

Operationally, mutual cooperation tends to increase trust and empathy while reducing stress and frustration; betrayal and repeated conflict decrease trust and increase frustration and stress; reciprocity partially restores empathy and supports recovery. Such interaction-driven state updates are consistent with established models of repeated games and adaptive multi-agent systems, where agent states evolve as a function of interaction outcomes over time [Axelrod and Hamilton 1981, Nowak 2006, Sandholm 2010].

The full update coefficients are fixed across all experiments and are omitted here for brevity, since they are not the main object of comparison. Their role is to provide a common relational substrate so that differences in recovery, collapse, coupling sensitivity, and robustness can be attributed to the decision representation rather than to different state-update dynamics, following standard experimental design principles in sequential decision processes [Sutton and Barto 2018].

3. Experimental Protocol

The experimental protocol was designed to test whether the proposed quantum-fuzzy formulation provides benefits that extend beyond pointwise action prediction and become more evident at the level of relational dynamics. To this end, the benchmark was organized in two layers. The first layer preserves the original predictive evaluation introduced in v1, thereby ensuring continuity with a standard supervised decision benchmark. The second layer extends the evaluation to a set of controlled dynamic scenarios designed to probe interaction-dependent phenomena, including betrayal recovery, coupling sensitivity, robustness under relational shocks, and temporal organization of behavioral regimes.

This two-level design is important because it enables a direct comparison between local predictive behavior and system-level relational adaptation, a distinction that is central in repeated-game analysis and multi-agent systems where dynamics emerge from interaction over time rather than isolated decisions [Axelrod and Hamilton 1981, Nowak 2006, Sandholm 2010, Sutton and Barto 2018, Leibo et al. 2017].

3.1. Compared Settings

We evaluate three configurations that share the same emotional state representation, the same IPD payoff structure, and the same relational update equations, but differ in how decision-relevant information is represented and transformed into action probabilities.

1. **Classical Fuzzy**: a rule-based fuzzy inference system that directly maps emotional inputs and compact historical features into a cooperation probability;
2. **Quantum-Fuzzy**: a hybrid strategy in which fuzzy outputs are encoded into quantum-fuzzy relational states, from which local, partner, and joint interaction features are extracted and transformed into a decision score;
3. **Quantum-Fuzzy + Classifier**: a downstream predictive policy trained on round-level data generated from the quantum-fuzzy dynamics, using classical, relational, and quantum-derived features as input.

Table 1. Compared decision settings.

Setting	Description
Classical Fuzzy	Fuzzy inference with emotional variables and history signal, directly mapped to cooperation probability.
Quantum-Fuzzy	Fuzzy outputs encoded into a 4-qubit relational state with coupling-sensitive transformations and feature extraction.
Quantum-Fuzzy + Classifier	Supervised model trained on quantum-fuzzy trajectories using emotional, relational, and quantum-derived features.

3.2. Evaluation Axes

The benchmark is organized around five complementary axes, each targeting a different aspect of socially interdependent decision behavior.

1. **Pointwise prediction**: standard predictive evaluation on generated round-level data, used to verify whether the proposed representation remains competitive in local action prediction;
2. **Recovery after betrayal**: controlled scenarios in which a cooperative interaction is disrupted by an induced unilateral betrayal, followed by a recovery window;

3. **Coupling sensitivity**: systematic sweep of inter-agent coupling strength, used to test whether the decision model meaningfully responds to changes in dependence between agents;
4. **Relational robustness**: perturbation-based evaluation under exogenous shocks, designed to assess resilience, recovery speed, and return to cooperation;
5. **Temporal structure**: analysis of regime transitions and temporal clustering, used to investigate how the models organize behavioral trajectories over time.

The first axis preserves the original predictive benchmark and serves as a necessary point of continuity with conventional decision modeling. However, the remaining four axes reflect the main conceptual goal of the framework: to evaluate whether the quantum-fuzzy representation is especially useful in settings where interaction history, dependence, disruption, and adaptation matter more than isolated one-step classification.

In the betrayal-recovery scenario, episodes are divided into three phases: an initial cooperative phase, an induced betrayal event, and a recovery phase. In the reported configuration, 120 episodes of 50 rounds were simulated, with betrayal injected at round 16 and a recovery window of 12 rounds used for the main recovery statistics. In the coupling-sweep scenario, each strategy was evaluated over 90 episodes per condition across four coupling levels: `zero`, `low` (0.2), `medium` (0.45), and `high` (0.75). In the emotional-transition scenario, 100 episodes were generated under progressive emotional drift, with both forward and reverse trajectories in order to probe transition points and hysteresis. In the relational-shocks scenario, 100 episodes were executed under multiple scheduled perturbations, including frustration spikes, trust drops, exogenous defections, and internal noise. Finally, temporal separability was evaluated over 100 episodes using a sliding temporal window of size 5 and stride 3, followed by clustering and low-dimensional projection.

Taken together, these five axes provide a structured evaluation protocol that distinguishes between local predictive quality and system-level relational behavior. This distinction is central to the claims of the paper because the main question is not whether the quantum-fuzzy model is a universally superior classifier, but whether it provides a more expressive representation for relational adaptation in socially interdependent settings.

3.3. Metrics

Predictive performance is evaluated using standard binary classification metrics, including accuracy, precision, recall, F1-score, and ROC-AUC [Powers 2011, Fawcett 2006]. Among these, accuracy and F1-score are adopted as the main reference metrics due to their interpretability and class-balanced assessment. All predictive evaluations are conducted on the `quantum_fuzzy_generated` dataset, ensuring a consistent comparison across strategies.

Relational dynamics are assessed through scenario-specific metrics capturing disruption and recovery. In the betrayal-recovery setting, we measure *recovery time*, *collapse rate*, *post-betrayal payoff*, and *relational stability*. In the coupling sweep, the main outcome is *mutual cooperation rate*, complemented by *system stability* and *mean payoff*, reflecting sensitivity to inter-agent dependence [Axelrod and Hamilton 1981, Nowak 2006, Sandholm 2010].

Robustness under perturbation is evaluated through *resilience score*, *recovery time*, and *return to cooperation*. Emotional-transition behavior is characterized by *transition point*, *smoothness*, *ambiguous zone width*, and *hysteresis*. Temporal organization is assessed via clustering-based metrics, including *silhouette score*, *temporal separability*, and *cluster stability* [Rousseeuw 1987].

This combination of predictive, relational, and temporal metrics allows us to distinguish local predictive performance from interaction-level dynamics, emphasizing the role of relational adaptation in socially interdependent decision processes [Sutton and Barto 2018].

4. Results

4.1. Competitive but Marginal Predictive Gains

The predictive benchmark confirms that the quantum-fuzzy representation remains fully competitive with the classical fuzzy baseline, while yielding only modest gains in pointwise prediction. On the main evaluation dataset, the **Classical Fuzzy** strategy achieved an accuracy of 0.7610 and an F1-score of 0.8039, whereas the **Quantum-Fuzzy** strategy achieved the best overall predictive performance with accuracy 0.7672 and F1-score 0.8153. The **Quantum-Fuzzy + Classifier** condition remained very close, with accuracy 0.7667 and F1-score 0.8151. ROC-AUC values were nearly identical across methods, ranging from 0.7931 to 0.7946.

Table 2. Pointwise predictive performance on the main evaluation dataset. (mean $\pm$ sd, CI95%).

Strategy	Accuracy	Precision	Recall	F1	ROC-AUC
Classical Fuzzy	0.7610 ± 0.0027 [0.7576, 0.7643]	$\mathbf{0.8360 \pm 0.0025}$ **[0.8329, 0.8391]**	0.7742 ± 0.0044 [0.7687, 0.7797]	0.8039 ± 0.0028 [0.8004, 0.8074]	0.7931 ± 0.0018 [0.7909, 0.7953]
Quantum-Fuzzy 4q	0.7672 ± 0.0025 [0.7641, 0.7703]	0.8186 ± 0.0029 [0.8149, 0.8222]	$\mathbf{0.8121 \pm 0.0029}$ **[0.8085, 0.8156]**	$\mathbf{0.8153 \pm 0.0021}$ **[0.8127, 0.8180]**	$\mathbf{0.7946 \pm 0.0018}$ **[0.7923, 0.7969]**
Quantum-Fuzzy + classifier	$\mathbf{0.7667 \pm 0.0022}$ **[0.7639, 0.7694]**	0.8176 ± 0.0038 [0.8129, 0.8224]	0.8126 ± 0.0035 [0.8083, 0.8169]	0.8151 ± 0.0016 **[0.8130, 0.8171]**	0.7937 ± 0.0032 [0.7897, 0.7977]

These results support two conclusions. First, the quantum-fuzzy model does not degrade local predictive behavior; on the contrary, it slightly improves the main predictive metrics relative to the classical baseline. Second, the gains are small in magnitude, which prevents any strong claim of broad predictive superiority. This is an important point for the interpretation of the study. The proposed framework should not be understood primarily as a better classifier in the conventional supervised-learning sense. Rather, the predictive results establish that the quantum-fuzzy representation remains competitive while leaving open the possibility that its main value lies elsewhere.

The classifier-augmented condition is also informative. Because its performance remains nearly tied with the pure quantum-fuzzy model, the results suggest that most of the benefit comes from the relational representation itself rather than from downstream classification complexity. In other words, the main gain appears to be encoded in the structure of the decision state, not merely in the flexibility of an additional learning layer.

4.2. Substantial Gains in Recovery After Betrayal

The strongest improvements in the study emerge in the betrayal-recovery scenario. Here, the quantum-fuzzy strategies exhibit markedly superior relational repair after an induced rupture. The **Classical Fuzzy** baseline required an average recovery time of 5.9068 rounds and exhibited a collapse rate of 0.3247, indicating that a substantial fraction of episodes failed to return to a stable cooperative regime after betrayal. By contrast, the **Quantum-Fuzzy** model reduced recovery time to 2.2833 and collapse rate to 0.0510, while the **Quantum-Fuzzy + Classifier** model further reduced recovery time to 1.4250 and collapse rate to 0.0283.

Table 3. Summary of the strongest relational-dynamics effects.

Scenario	Strategy	Rec. Time	Collapse / Res.	Stability / Return
Betrayal Recovery	Classical Fuzzy	5.9068	0.3247	0.4840
	Quantum-Fuzzy	2.2833	0.0510	0.7554
	Quantum-Fuzzy + Classifier	**1.4250**	**0.0283**	**0.7800**
Coupling Sweep	Classical Fuzzy	–	–	0.1707
	Quantum-Fuzzy	–	–	**0.5500**
	Quantum-Fuzzy + Classifier	–	–	0.5249
Relational Shocks	Classical Fuzzy	7.9596	0.0466	0.1097
	Quantum-Fuzzy	2.7576	**0.2256**	**0.4992**
	Quantum-Fuzzy + Classifier	**2.5455**	0.2077	0.4505

These differences are not minor fluctuations; they indicate a qualitatively different capacity for relational repair. The post-betrayal payoff also improved substantially, rising from 2.1236 in the classical baseline to 2.7735 in the quantum-fuzzy condition and 2.8207 in the classifier-augmented condition. Similarly, relational stability increased from 0.4840 to 0.7554 and 0.7800, respectively. Taken together, these results show that once cooperation is disrupted, the quantum-fuzzy framework is considerably more effective at rebuilding a cooperative interaction trajectory.

This finding is especially relevant because betrayal is one of the most demanding tests for any relational decision model. A strategy that performs well only under stable conditions but fails to recover after a rupture cannot be said to model socially interdependent dynamics in a robust way. The results therefore indicate that the quantum-fuzzy representation is particularly advantageous in settings where disruption, reciprocity, and re-stabilization are central.

4.3. Strong Sensitivity to Coupling

The coupling sweep provides one of the clearest mechanistic signals in the benchmark. In the **Classical Fuzzy** baseline, increasing coupling strength had almost no effect on the main relational outcome. Mutual cooperation remained essentially flat, moving only from 0.1738 under zero coupling to 0.1707 under high coupling, while system stability and mean payoff changed only marginally. This weak response suggests that the classical

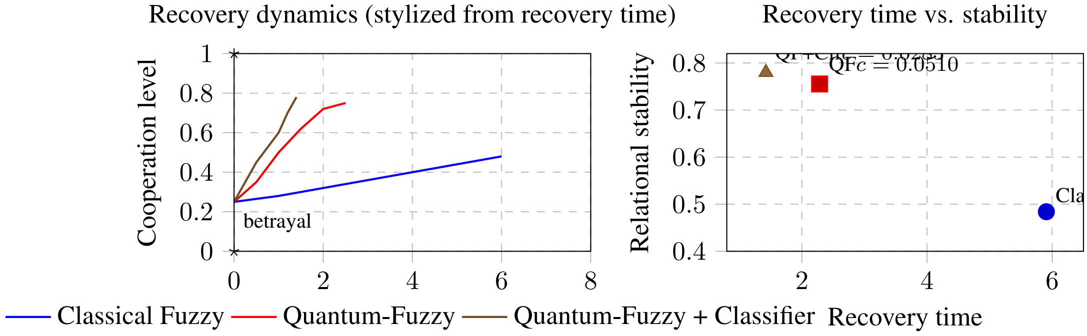

Figure 1. Betrayal-recovery behavior across strategies. Left: stylized recovery profiles derived from observed recovery times, illustrating the relative speed of cooperation restoration. Right: empirical results showing recovery time versus relational stability, with c denoting collapse rate. Quantum-fuzzy models recover substantially faster, exhibit lower collapse rates, and stabilize at higher cooperation levels than the classical baseline.

model, although affected by history and emotional state, does not substantially reorganize its behavior when dependence between agents is externally intensified.

The quantum-fuzzy strategies behave very differently. In the **Quantum-Fuzzy** condition, mutual cooperation rose from 0.1911 under zero coupling to 0.5500 under high coupling, accompanied by an increase in system stability from 0.8236 to 0.8709 and in mean payoff from 1.9342 to 2.6213. The **Quantum-Fuzzy + Classifier** condition showed a similar pattern, increasing mutual cooperation from 0.2427 to 0.5249, while also improving stability and mean payoff. These results indicate that the coupling parameter is not merely a formal component of the model; it has a concrete and substantial behavioral effect.

This pattern provides strong support for the hypothesis that the quantum-fuzzy representation is better aligned with interaction regimes in which inter-agent dependence is an explicit modeling target. The coupling-sensitive controlled operations appear to make the decision state more responsive to relational context, thereby enabling cooperative coordination to emerge as dependence intensifies. This is consistent with the broader interpretation of the framework as a model of relational adaptation rather than a purely local predictor.

4.4. Higher Robustness Under Relational Shocks

A similarly strong advantage appears in the relational-shocks scenario. Under scheduled perturbations, the **Classical Fuzzy** baseline exhibited a resilience score of only 0.0466, an average recovery time of 7.9596, and a return-to-cooperation rate of 0.1097. These values indicate that exogenous shocks frequently destabilize the interaction and that the baseline has limited ability to rebuild cooperative behavior once the relational bond has been perturbed.

The **Quantum-Fuzzy** strategy substantially outperformed the baseline across all main robustness metrics. Its resilience score increased to 0.2256, recovery time fell to 2.7576, and return to cooperation rose to 0.4992. Mean payoff and relational stability also improved considerably, reaching 2.5321 and 0.6506, respectively. The **Quantum-Fuzzy**

+ Classifier condition showed closely related behavior, with resilience score 0.2077, recovery time 2.5455, and return to cooperation 0.4505.

These results reinforce the interpretation that the main strength of the framework lies in how it handles relational perturbation. The quantum-fuzzy representation not only tolerates shocks better, but also reconstructs cooperative structure more effectively afterward. This distinction matters because robustness in socially interdependent environments is not simply a matter of resisting perturbation; it is also a matter of regaining coordination after destabilization. In that sense, the reported gains reflect a deeper improvement in adaptive relational behavior rather than a narrow predictive effect.

4.5. Mixed Temporal Results

The temporal analyses present a nuanced outcome. In the emotional-transition scenario, quantum-fuzzy models sustained cooperation for substantially longer before regime change, with the transition point increasing from 17.18 (**Classical Fuzzy**) to 41.70 (**Quantum-Fuzzy**) and 41.25 (**Quantum-Fuzzy + Classifier**). The ambiguous zone width also decreased sharply (from 0.6204 to 0.2010 and 0.1802), indicating more delayed and more sharply defined transitions under emotional drift.

However, this advantage is not consistent across all temporal metrics. Transition smoothness remained slightly higher in the classical baseline (0.9606 vs. 0.9530 and 0.9559), and temporal separability also favored the classical model (0.6533 vs. 0.6118 and 0.6202), despite higher distinguishability in the quantum-based variants.

Overall, these results do not support a uniform temporal advantage. Instead, they indicate that quantum-fuzzy encoding modifies transition dynamics—delaying collapse and reducing ambiguity—without consistently improving smoothness or separability. This reinforces that the primary strength of the framework lies in relational adaptation rather than global temporal superiority.

5. Discussion

The results consistently show that the primary advantage of the quantum-fuzzy framework lies not in pointwise predictive accuracy, but in its stronger representation of relational dynamics under interdependence, disruption, and recovery. The dynamic scenarios revealed substantially faster post-betrayal recovery, lower collapse rates, strong behavioral responsiveness to coupling variation, and significantly higher resilience under relational shocks. These effects suggest that the framework is especially effective in regimes where decision quality depends on the evolving structure of agent interaction, a phenomenon extensively studied in repeated games [Axelrod and Hamilton 1981, Nowak 2006, Sandholm 2010].

The quantum contribution should be understood as a representational mechanism that captures relational dependence in a compact and behaviorally meaningful way [Busemeyer and Bruza 2012], rather than as a generic classification advantage. This is reinforced by the coupling-sweep experiments, in which the classical baseline remained almost invariant while the quantum-fuzzy strategies showed large increases in mutual cooperation and stability as coupling intensified. The temporal results, however, impose an important qualification: while quantum-fuzzy models delayed regime transitions and reduced ambiguous transition zones, they did not consistently improve smoothness or separability, indicating that the framework's advantage is specific to relational adaptation and recovery.

Regarding limitations: the benchmark is intentionally synthetic, which enables isolation of representational effects but limits behavioral generalization [Sutton and Barto 2018]. Results correspond to a specific implementation—four-qubit encoding, fixed coupling operators, a selected family of scenarios—and broader validation across seeds, parameter sweeps, and encodings is required. The marginal classifier gains suggest the dominant benefit resides in the representation itself, motivating targeted ablation studies. These findings are particularly relevant for reinforcement learning and multi-agent systems, where decision quality emerges from sequential interaction rather than isolated predictions.

6. Conclusion

This paper presented a comparative benchmark between classical fuzzy and quantum-fuzzy decision models in the Iterated Prisoner's Dilemma under emotion-aware relational dynamics. The main contribution of the study is a more precise characterization of where the quantum-fuzzy advantage actually emerges: not as broad predictive dominance, but as a stronger capacity to model rupture, recovery, coupling, and resilience in socially interdependent environments. Across the reported experiments, the quantum-fuzzy strategies remained competitive in pointwise prediction and achieved slight gains in F1-score and accuracy, but their most significant improvements appeared in dynamic relational scenarios, including substantially faster post-betrayal recovery, sharply reduced collapse rates, stronger sensitivity to inter-agent coupling, and higher robustness under relational shocks. At the same time, the temporal analyses showed mixed outcomes, which productively narrows the claim of the paper and strengthens its methodological credibility. Overall, the findings support quantum-fuzzy encoding as a promising direction for modeling adaptive relational behavior in computational intelligence, particularly in problems where interaction history, dependence, and post-disruption reorganization are central. Future work should extend the current benchmark through multi-seed statistical validation, systematic ablation of coupling and encoding components, richer temporal analyses, and broader multi-agent settings that can test whether the observed relational benefits persist beyond the two-agent IPD scenario.

References

Axelrod, R. and Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489):1390–1396.

Busemeyer, J. R. and Bruza, P. D. (2012). *Quantum models of cognition and decision*. Cambridge University Press.

Farias, G. P., Dimuro, G. P., Peter, G. D., and De Manuel Jerez, E. (2013). A bdi-fuzzy agent model for exchanges of non-economic services based on the social exchange theory. In *Anais do VII Workshop-Escola de Sistemas de Agentes, seus Ambientes e Aplicações (WESAAC)*. Sociedade Brasileira de Computação.

Fawcett, T. (2006). An introduction to roc analysis. *Pattern Recognition Letters*, 27(8):861–874.

Kosko, B. (1990). Fuzziness vs. probability. *International Journal of General Systems*, 17(2-3):211–240.

Leibo, J. Z., Zambaldi, V., Lanctot, M., Marecki, J., and Graepel, T. (2017). Multi-agent reinforcement learning in sequential social dilemmas. *Proceedings of AAMAS*.

Nielsen, M. A. and Chuang, I. L. (2010). *Quantum computation and quantum information*. Cambridge University Press.

Nowak, M. A. (2006). *Evolutionary dynamics: exploring the equations of life*. Harvard University Press.

Nowak, M. A. and Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, 355:250–253.

Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):37–63.

Qiskit Development Team (2021). Qiskit: An open-source framework for quantum computing. https://qiskit.org.

Rodrigues, H. D. N., Dimuro, G. P., and Adamatti, D. F. (2017). Um modelo de reputação fuzzy de dimensão variável. In *Anais do XI Workshop-Escola de Sistemas de Agentes, seus Ambientes e Aplicações*, pages 113–124, São Paulo, SP, Brasil. Sociedade Brasileira de Computação.

Rojas, Y. E. L., Dimuro, G. P., and Adamatti, D. F. (2014). Trocas sociais em sistemas multiagentes: Transferência de confiança com base na reputação e na relação de dependência. In *Anais do VIII Workshop-Escola de Sistemas de Agentes, seus Ambientes e Aplicações (WESAAC)*. Sociedade Brasileira de Computação.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65.

Sandholm, W. H. (2010). *Population games and evolutionary dynamics*. MIT Press.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press, 2 edition.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3):338–353.

Zadeh, L. A. (1996). Fuzzy logic = computing with words. *IEEE Transactions on Fuzzy Systems*, 4(2):103–111.