Quantum Games

Junipero Verbeke

University of California, Santa Cruz

(Dated: June 13, 2025)

Abstract

Classical game theory provides a robust mathematical framework for strategic interactions, yet it often leads to dilemmas where individual rationality results in collectively suboptimal outcomes. The introduction of quantum mechanics offers a profound expansion of this framework, promising to resolve these classical impasses by leveraging physical principles like superposition and entanglement. This work establishes a formal bridge between classical and quantum game theory by mapping traditional strategic concepts onto the language of quantum information. We demonstrate how players, by utilizing shared entangled qubits and executing strategies as local unitary operations, can access a vastly larger strategic space than is available classically. This quantum extension of games is shown to fundamentally alter their underlying dynamics. The primary result of this analysis is that quantum strategies can dismantle classical dilemmas, exemplified by the Prisoner's Dilemma, by establishing novel Quantum Nash Equilibria that are Pareto optimal. It is conclusively shown that the expected payoffs for players employing optimal quantum strategies are always greater than or equal to those achievable through any classical means, highlighting a significant and foundational advantage conferred by the quantum domain.

INTRODUCTION

Some people play poker and win a pot, but John von Neumann played poker and invented game theory. The modeling of strategic interactions is one of the most recent branches of mathematics, while also being one of its most applicable forms. Game theory provides a mathematical framework for analyzing situations involving strategic interactions between rational decision-makers, termed players [1]. The fundamental elements include a set of players N, where for each player $i \in N$, there is a set of available strategies S_i, and a payoff function π_i (or E_i) that assigns a numerical value (utility) to player i for each possible combination of strategies chosen by all players [2, 3]. A core assumption is that players act rationally, seeking to maximize their own payoff given their beliefs about the other players' actions. Classical games placing two or more players in situations where both may or may not know the others' decisions already have established solutions in game theory. When players gain access to quantum mechanics, how do their new strategies compare to classical decision-making? Are quantum strategies more profitable than any classical strategy?

Approaching quantum games requires an introduction to the original games presented in Von Neumann's seminal work: Theory of Games and Economic Behavior [4]. This paper begins by studying his examples, expanding the concept of a game from simple two person zero sum, to non-zero sum and eventually static games. Along with the classical cases in each category, their quantum analogues will be developed along with a system to extend any classical game. These analogies lead to a discussion on the quantum equivalents to important game theory concepts. Armed with the mechanics of quantum game theory, we discover advantages over classical strategy in multiplayer games with concrete applications. Specifically, quantum game theory grants new solutions to classically insoluble games. These results are finally quantified by demonstrating that quantum strategies have at least equal payoff (utility) to classical strategies.

I. CLASSICAL TWO-PLAYER GAMES: FOUNDATIONS AND DILEMMAS

This section lays the groundwork for understanding strategic interactions by exploring the foundational concepts of classical two-player games. We begin by distinguishing between two fundamental game types: zero-sum games, where players' interests are in direct opposition, and non-zero-sum games, which allow for outcomes of mutual gain or loss. Through canonical examples like "Matching Pennies" and the "Prisoner's Dilemma," we introduce essential analytical tools, including the Minimax Theorem for resolving zero-sum conflicts, and the concepts of Nash Equilibrium and Pareto Optimality for analyzing more complex scenarios. The discussion culminates in highlighting the central conflict that motivates much of game theory—the tension between individual rational choices and collective well-being, a dilemma illustrated when the Nash Equilibrium of a game is not Pareto Optimal.

A. Zero-Sum Games

In a two-player zero-sum game, the interests of the players are diametrically opposed. Any gain by a player is associated by an equal loss for the other. for For any outcome resulting from the players' chosen strategies $(s_1 \in S_1, s_2 \in S_2)$, the sum of their payoffs is zero: $\pi_1(s_1, s_2) + \pi_2(s_1, s_2) = 0$ [2]. A classic example is Matching Pennies, where each player reveals a coin, and one player wins if the faces match, while the other wins if they differ. The payoff matrix for Matching Pennies is illustrated in Table I. The strategies employed by players can be either pure or mixed. A pure strategy involves choosing a single action deterministically, whereas a mixed strategy involves selecting actions probabilistically according to a specific probability distribution [2, 5]. The central solution concept for zero-sum games is the Minimax Theorem, established by von Neumann. This theorem guarantees the existence of an equilibrium value and optimal strategies (which can be pure or mixed) such that neither player can improve their outcome by unilaterally changing their strategy [5]. This equilibrium corresponds to a saddle point in the payoff matrix.

Player 1 / Player 2	**Heads**	**Tails**
Heads	$(1, -1)$	$(-1, 1)$
Tails	$(-1, 1)$	$(1, -1)$

TABLE I. Payoff matrix for the classical two-player zero-sum game: Matching Pennies.

B. Non-Zero-Sum Games

Unlike zero-sum games, non-zero-sum games allow for outcomes where players can achieve mutual gains or suffer mutual losses, as the sum of their payoffs is not necessarily constant or zero. The Prisoner's Dilemma is the canonical example of a non-zero-sum game and starkly highlights the conflict between individual rationality and collective welfare [3, 6]. In this scenario, two suspects are interrogated separately. If both cooperate (stay silent), they each receive a moderate sentence. If one defects (testifies) while the other cooperates, the defector goes free, and the cooperator receives a harsh sentence. If both defect, they both receive an intermediate sentence. The payoff matrix in Table II illustrates a typical Prisoner's Dilemma.

C. Classical game theory tools

Understanding the dynamics of such games requires key concepts. One is the Nash Equilibrium (NE), which describes a profile of strategies (one for each player) such that no player can improve their expected payoff by unilaterally changing their strategy, assuming other players' strategies remain fixed [1, 5]. In the Prisoner's Dilemma (Table II), mutual defec-

		Player B	
		Cooperate (C)	Defect (D)
Player A	Cooperate (C)	(3, 3)	(0, 5)
	Defect (D)	(5, 0)	(1, 1)

TABLE II. Payoff matrix for a two-player non-zero-sum game: Prisoner's Dilemma. This matrix shows the Pareto optimal outcome (3,3) resulting from mutual cooperation and the Nash equilibrium (1,1) resulting from mutual defection.

tion, yielding payoffs (1,1), constitutes the Nash Equilibrium. Although mutual cooperation would yield a better outcome for both (3,3), defection is the dominant strategy for each individual player regardless of the other's choice. John Nash proved that every finite game has at least one Nash Equilibrium, potentially in mixed strategies [7].

Another important concept is Pareto Optimality (PO). An outcome is defined as Pareto optimal if it is impossible to make one player better off without making at least one other player worse off [2, 8]. In the Prisoner's Dilemma example, the outcome (3,3) achieved through mutual cooperation is Pareto optimal. From this state, any change in strategy by one player to improve their own payoff would necessarily worsen the payoff for the other player (e.g., if Player A defects from (C,C) to gain 5, Player B's payoff drops from 3 to 0).

The core issue, often referred to as "The Dilemma" in game theory, arises when the Nash Equilibrium is not Pareto Optimal. The Prisoner's Dilemma powerfully illustrates this: individually rational choices (defecting) lead to a Nash Equilibrium (1,1) that is collectively suboptimal compared to the Pareto optimal outcome (3,3) that could be achieved if players were able to commit to cooperation [6]. This highlights the fundamental tension between individual incentives and group welfare.

II. INTRODUCING QUANTUM MECHANICS TO GAMES

A. Why Quantize?

Classical game theory assumes strategies are chosen from well-defined sets and combined using classical probability. However, if players have access to quantum resources or interact via quantum systems, the classical framework may be insufficient. Quantum mechanics allows for phenomena like superposition and entanglement, which can fundamentally alter the strategic landscape and potentially offer advantages unobtainable classically [5, 6].

B. Quantum Representation

To incorporate quantum mechanics, classical game elements are mapped onto quantum concepts. The fundamental unit of quantum information employed is the qubit. For a game with two classical strategies, such as Cooperate or Defect, these strategies can be represented by the orthogonal basis states $|0\rangle$ and $|1\rangle$ of a qubit, which is a vector in a 2-dimensional Hilbert space $\mathcal{H}$ [3, 8]. A key distinction from classical bits is the principle of superposition, whereby a qubit can exist in a linear combination of its basis states, formally written as $\alpha|0\rangle + \beta|1\rangle$, where α and β are complex numbers such that the sum of their squared magnitudes, $|\alpha|^2 + |\beta|^2$, equals 1. This property allows players to, in a sense, consider multiple strategies simultaneously [3]. Player actions or strategic choices are then described by unitary operations (U) acting on the qubit states. These mathematical operations are reversible and preserve the normalization of the quantum state, ensuring that $U^\dagger U = I$, where $U^\dagger$ is the Hermitian conjugate of U and I is the identity operator [3, 8]. Examples of such operations include the identity I, which leaves the state unchanged, and the bit-flip X (Pauli-X operator), which transforms $|0\rangle$ to $|1\rangle$ and $|1\rangle$ to $|0\rangle$.

C. Entanglement

When multiple players are involved, their quantum states reside in a composite Hilbert space, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ for two players A and B. Entanglement describes specific states in this composite space that cannot be written as a simple product of individual player states (i.e., they are non-separable) [3, 8]. Entangled states, like the Bell state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, exhibit correlations stronger than any achievable classically. This allows for coordinated strategies and outcomes that have no classical analogue[2, 6]. From a game theory perspective, these stronger correlations are precisely what can alter game outcomes and strategies. In a classical game, player's choices are independent, and any coordination must be explicitly agreed upon or inferred through common knowledge. With entangled states, however, player's measurements become intrinsically linked, even if they are physically separated and cannot communicate. This inherent correlation means that if one player measures their part of the entangled state, the outcome immediately influences the possible outcomes for the other player's measurement without any classical information exchange. This pre-established non-classical link allows players to execute strategies that are perfectly coordinated, often leading to a higher probability of achieving a mutually beneficial outcome, or enabling new equilibrium points that don't exist in the classical version of the game [7]. It's this non-local but perfectly coordinated behavior that give quantum players an advantage, allowing them to "pre-agree" on a joint strategy in a way that classical players cannot. Applying entanglement to the Prisoner's Dilemma, the game begins with both players sharing an entangled pair of qubits, typically prepared in the Bell state $\frac{1}{\sqrt{2}}(|CC\rangle + |DD\rangle)$, where $|C\rangle$ (Cooperate) corresponds to the basis state $|0\rangle$ and $|D\rangle$ (Defect) to $|1\rangle$. A player's strategy is no longer a simple choice between C and D, but a choice of a quantum operation (a unitary matrix) to apply to their individual qubit. A key discovery is the existence of a particular quantum operation which, if applied by both players to their respective qubits, manipulates the entangled state in such a way that it collapses to the $|CC\rangle$ (mutual cooperation) outcome upon measurement. This creates a new, stable Nash Equilibrium. If a player knows their opponent will use this quantum strategy, their best move is to do the same, guaranteeing the Pareto optimal payoff for both and thus resolving the dilemma. Any unilateral deviation from this quantum strategy results in a worse payoff for the deviating player, making mutual cooperation the rational outcome.

D. Measurement

To determine the outcome of a quantum game, a measurement is performed. Quantum measurements are inherently probabilistic. Measuring a state $|\psi\rangle = \alpha |0\rangle + \beta |1\rangle$ in the $\{|0\rangle, |1\rangle\}$ basis yields outcome 0 with probability $|\alpha|^2$ and outcome 1 with probability $|\beta|^2$. The state "collapses" to the measured basis state [3]. For mixed states (probabilistic combinations of pure states), the density matrix formalism ρ is used [3, 8]. The expected payoff for player i is calculated using a corresponding payoff operator $\hat{\pi}_i$, which is constructed from the classical payoffs and measurement projectors: $\bar{\pi}_i = \mathrm{Tr}[\rho_f \hat{\pi}_i]$, where ρ_f is the final state after players' moves [3].

E. Quantization Protocols

Several frameworks exist for structuring quantum games. One of the earliest frameworks, Meyer's Approach, imagines a quantum game as a turn-based competition. In a game like a coin toss, the players use a single quantum object, like a qubit. The game starts with the "coin" in a set initial state (e.g., "Heads"). The first player then performs an action, which isn't a physical flip but the application of a specific quantum operation (a unitary transformation) of their choosing. The quantum coin is then handed to the second player, who applies their own chosen operation to its new state. After the final player has taken their turn, the coin's state is measured. The result of this measurement determines the game's outcome and who wins, demonstrating how one player can use quantum tools to strategically influence an object that is then passed to their opponent [3, 5]. The more common framework, known as the Eisert-Wilkens-Lewenstein (EWL) Protocol, is designed for games where players make their moves simultaneously, like the Prisoner's Dilemma. In this setup, the game begins with a referee or source that prepares two quantum particles, such as a pair of qubits, in a special, entangled state. This means their properties are linked, no matter how far apart they are. One particle is given to Player A and the other to Player B. The players, who can be thought of as being in separate rooms, then make their "move" at the same time by choosing and applying a quantum operation to their own particle. This choice is their strategy. For example, one quantum operation might correspond to the classical move "Cooperate," while another corresponds to "Defect." After they have acted, the particles are brought back together and measured. The specific correlations guaranteed by the initial entanglement mean that the outcome of the measurement for Player A is not independent of Player B's action. The final combined measurement result determines the outcome and thus the payoffs for both players, often allowing for new, more beneficial results that weren't possible in the classical version of the game. [2, 3, 6].

III. QUANTUM GAMES: NEW STRATEGIES AND OUTCOMES

A. Defining Quantum Games

A quantum game extends its classical counterpart by replacing classical strategy spaces with Hilbert spaces, allowing players to prepare initial states (possibly entangled) and employ quantum operations as their strategies. Payoffs are typically determined by measuring the final quantum state that results from the players' actions and then referencing the original classical payoff structure associated with the measurement outcomes [2, 3].

B. Quantum Strategies

The ability to leverage quantum mechanical phenomena significantly expands the set of available strategies beyond classical probabilistic choices. A pure quantum strategy typically corresponds to a player applying a specific unitary operator U_i from an allowed set (such as $SU(2)$ for games based on single qubits) to their part of the shared quantum state [5, 6]. Beyond these, players can employ mixed quantum strategies. These can be conceptualized as classical probability distributions over a set of pure quantum strategies. More generally, a player's strategy or state can be described by a density matrix ρ, which can represent not only a probabilistic mixture of pure states but also potential entanglement with an external environment or other players [1, 7]. In the most comprehensive formulation, any completely positive trace-preserving (CPTP) map acting on a player's subsystem can be considered a representation of a generalized quantum strategy [1].

C. Quantum Equilibria

The concept of equilibrium, central to classical game theory, is adapted and extended within the quantum framework. A key notion is the Quantum Nash Equilibrium (QNE). A quantum state ρ constitutes a QNE if no single player i can improve their expected payoff, $\bar{\pi}_i = \mathrm{Tr}[\rho\hat{\pi}_i]$ (where $\hat{\pi}_i$ is the payoff operator for player i), by unilaterally applying any allowed quantum operation Φ_i to their subsystem [1, 3, 7]. This condition is formally expressed as $\mathrm{Tr}[\rho\hat{\pi}_i] \geq \mathrm{Tr}[(\Phi_i \otimes I_{-i})(\rho)\hat{\pi}_i]$ for all permissible operations Φ_i by player i, where I_{-i} is the identity operation on all other players' subsystems. For a strict QNE, ρ is often required to be a product state ($\rho = \rho_A \otimes \rho_B$), while allowing for entangled states can lead to related concepts like Quantum Correlated Equilibrium (QCE).

Crucially, entanglement plays a pivotal role in many unique phenomena observed in quantum games. The presence of entanglement in the initial state shared by players can lead to the emergence of new Nash equilibria that have no classical counterparts. Conversely, classical Nash Equilibria may become unstable or cease to be equilibria when players are permitted to use quantum strategies, particularly those involving manipulations of entangled states [2, 6]. The degree of entanglement available often dictates whether, and to what extent, quantum advantages can manifest over classical strategies [6].

D. Quantum Advantage in Action

Quantum strategies can demonstrably outperform classical ones in several key examples. As discussed by Meyer [5], in a simple coin-tossing game (PQ Penny Flip), a player utilizing quantum superposition can force a win with certainty against an opponent restricted to classical coin flips, thereby demonstrating a clear quantum advantage in a zero-sum setting [3].

In the Quantum Prisoner's Dilemma, a significant quantum advantage can arise, particularly if players share an initially entangled state. For instance, players might each receive a qubit from a pair prepared in an entangled Bell state, such as $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. This initial entanglement introduces non-classical correlations that are fundamental to how the quantum version of the game unfolds. Instead of choosing classically to "cooperate" or "defect," a quantum player's strategy involves applying a unitary operation to their individual qubit. A common parameterization for such an operation for a single qubit is $\hat{Q} = \begin{pmatrix} \cos(\theta/2) & i\sin(\theta/2) \\ i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$, where the choice of the angle θ (and potentially other phase angles for general SU(2) operations) becomes the strategic move. For example, $\theta = 0$ might map to the classical "Cooperate" action, while $\theta = \pi$ (or $\pi/2$ depending on exact mapping) could represent "Defect," and other values of θ allow for a continuous spectrum of quantum strategies that are superpositions of these classical actions [3, 6]. After both players apply their chosen unitary operations, the composite quantum state of their qubits evolves. Subsequently, a measurement is performed on this final state. The outcome of this measurement is probabilistic, with the probabilities of different classical outcomes (e.g., CC, CD, DC, DD) depending directly on the final quantum state, which in turn depends on the unitary operations chosen by the players and the properties of the initial entangled state. The payoff for each player is then calculated based on these measurement outcomes, typically using the payoff values from the original classical Prisoner's Dilemma matrix. By choosing appropriate quantum strategies in the presence of sufficient entanglement, players can reach outcomes, like mutual cooperation with higher payoffs, that are not stable Nash equilibria in the classical version of the game.

In general, it has been shown that the expected payoff for a player using an optimal quantum strategy is always greater than or equal to the payoff achievable with an optimal classical (mixed) strategy: $\bar{\pi}_{\text{quantum}} \geq \bar{\pi}_{\text{classical}}$ [1, 7]. This is because quantum mechanics expands the available strategy space; classical strategies can be seen as a subset of the broader set of quantum strategies.

CONCLUSION

The introduction of quantum mechanics fundamentally alters the landscape of game theory. This paper has charted a course from the foundational principles of classical game theory, exploring zero-sum interactions like Matching Pennies, non-zero-sum challenges such as the Prisoner's Dilemma, and key concepts including Nash Equilibrium, Pareto Optimality, and the Minimax Theorem, to the enriched strategic domain offered by quantum mechanics. We detailed how quantum phenomena, including the representation of strategies using qubits, the exploitation of superposition, the application of unitary operations, and most critically, the non-local correlations afforded by entanglement, are formally integrated into game structures. Protocols were presented as frameworks for this quantization. By allowing strategies based on these quantum principles, quantum games exhibit new equilibrium structures, such as Quantum Nash Equilibria, and can resolve long-standing classical dilemmas. For instance, the Quantum Prisoner's Dilemma demonstrated that with sufficient entanglement, outcomes aligning with mutual cooperation (and thus Pareto optimality) can become stable equilibria. Quantum strategies thus provide players with tools that can lead to outcomes unattainable through classical means. Our analysis confirmed that these quantum approaches offer payoffs at least as favorable as optimal classical strategies ($\bar{\pi}_{quantum} \geq \bar{\pi}_{classical}$), often enabling significantly improved results and opening new avenues for analyzing strategic interactions.

Looking towards the future of quantum game theory, a critical challenge lies in studying the robustness of quantum advantages against the effects of environmental noise and decoherence, which are inevitable in any realistic physical implementation. Furthermore, this paper has focused on static games, but the application of quantum principles to dynamic and evolutionary games, where strategies can be learned and adapted over time, remains a fertile area of research that intersects with the growing field of quantum machine learning. Finally, translating these theoretical advantages into practical applications, whether in developing novel quantum algorithms, designing robust quantum cryptographic protocols, or modeling complex economic and biological systems, represents the ultimate long-term goal for the field.

---

[1] S. Zhang, Quantum Strategic Game Theory (2011), 64 citations (Semantic Scholar/arXiv) [2025-05-07] 64 citations (Semantic Scholar/arXiv) [2025-04-17] 64 citations (Semantic Scholar/arXiv) [2025-04-17] arXiv:1012.5141 [quant-ph].

[2] A. P. Flitney and D. Abbott, An introduction to quantum game theory (2002), 96 citations (Semantic Scholar/arXiv) [2025-05-07] 96 citations (Semantic Scholar/arXiv) [2025-04-17] arXiv:quant-ph/0208069.

[3] S. Das, Quantumizing Classical Games: An Introduction to Quantum Game Theory (2023), 2 citations (Semantic Scholar/arXiv) [2025-05-07] 2 citations (Semantic Scholar/arXiv) [2025-04-17] 2 citations (Semantic Scholar/arXiv) [2025-04-17] arXiv:2305.00368 [quant-ph].

[4] J. Von Neumann and O. Morgenstern, *Theory of games and economic behavior*, 60th ed., A Princeton classic edition (Princeton University Press, Princeton, NJ, 2007).

[5] D. A. Meyer, Quantum Strategies, Physical Review Letters **82**, 1052 (1999), 592 citations (Crossref/DOI) [2025-05-07] 6 citations (Semantic Scholar/DOI) [2025-05-07] 687 citations (Semantic Scholar/DOI) [2025-04-17] 687 citations (Semantic Scholar/DOI) [2025-04-17] 588 citations (Crossref/DOI) [2025-04-17] 588 citations (Crossref/DOI) [2025-04-17] Publisher: American Physical Society.

[6] J. Eisert, M. Wilkens, and M. Lewenstein, Quantum Games and Quantum Strategies, Physical Review Letters **83**, 3077 (1999), 966 citations (Semantic Scholar/arXiv) [2025-05-07] 763 citations (Crossref/DOI) [2025-05-07] 963 citations (Semantic Scholar/DOI) [2025-04-17] 963 citations (Semantic Scholar/arXiv) [2025-04-17] 760 citations (Crossref/DOI) [2025-04-17] arXiv:quant-ph/9806088.

[7] S. E. Landsburg, Nash Equilibria in Quantum Games (2011), 38 citations (Semantic Scholar/arXiv) [2025-05-07] 7 citations (Semantic Scholar/arXiv) [2025-05-07] 38 citations (Semantic Scholar/arXiv) [2025-04-17] arXiv:1110.1351 [math].

[8] E. G. Hidalgo, Quantum Games and the Relationships between Quantum Mechanics and Game Theory (2016), 8 citations (Semantic Scholar/arXiv) [2025-05-07] 966 citations (Semantic Scholar/DOI) [2025-05-07] 0 citations (Semantic Scholar/DOI) [2025-05-07] 8 citations (Semantic Scholar/arXiv) [2025-04-17] 8 citations (Semantic Scholar/arXiv) [2025-04-17] arXiv:0803.0292 [quant-ph].