

Uncoupled Quantum Dynamics and Nash Equilibria

1 MOTIVATION

Uncoupled learning asks whether equilibrium behavior can emerge when each agent uses only her own payoff information. This restriction is natural in decentralized multi-agent systems: a player may observe or evaluate her own payoff stream, but need not know the payoff functions of the other players. In classical games, this restriction leads to a sharp divide. Hart and Mas-Colell showed that uncoupled dynamics cannot be universally Nash-convergent [HMC03], and in later work they extended this impossibility to stochastic uncoupled dynamics [HMC06]. Subsequent work showed, however, that richer, higher-order rules, whose state includes auxiliary payoff-dependent variables, can locally stabilize certain mixed Nash equilibria [TS23]. All of these works assume that the strategy space is Euclidean (i.e. $\mathbb{R}^d$). This leads to the following question:

How does the classical boundary between convergence and non-convergence for uncoupled learning change when strategies admit a different geometric structure?

This paper studies this question for quantum games, where strategies are density operators rather than probability vectors. The feasible set of density matrices is organized by the subspaces on which those matrices have support: an equilibrium may be full-rank, lying in the interior of the state space, or it may be supported on a proper subspace and hence lie on the boundary—possibly at a degenerate boundary point where unused directions are payoff-tied with the support. Our aim is to determine which classical impossibility results for uncoupled learning still hold in the quantum setting, and which convergence results continue to apply once this matrix geometry is taken into account.

2 PRELIMINARIES

Let player i have finite-dimensional Hilbert space $\mathcal{H}_i$. Its quantum strategy set is

$$D_i = D(\mathcal{H}_i) = \{\rho \in \text{Herm}(\mathcal{H}_i) : \rho \succeq 0, \text{Tr}(\rho) = 1\},$$

the set of density matrices; we call this set the spectraplex. A finite quantum game is specified by Hermitian payoff observables $\Xi_i \in \text{Herm}(\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n)$. For a product profile $\rho = (\rho_1, \dots, \rho_n)$, $u_i(\rho) = \text{Tr}[\Xi_i(\rho_1 \otimes \cdots \otimes \rho_n)]$. A profile ρ^* is a quantum Nash equilibrium if $u_i(\rho_i^*, \rho_{-i}^*) \geq u_i(\rho_i, \rho_{-i}^*)$ for every $\rho_i \in D_i$ and player i . Let $\rho_{\neq i} = \bigotimes_{j \neq i} \rho_j$. Define the payoff-gradient operator as $\Gamma_i(\rho_{-i}) = \text{Tr}_{-i}[\Xi_i(I_i \otimes \rho_{-i})]$, so that the utility function can be written as $u_i(\rho_i, \rho_{-i}) = \text{Tr}[\rho_i \Gamma_i(\rho_{-i})]$, where Tr_{-i} traces out all systems except player i and I_i is the identity on $\mathcal{H}_i$.

Our two obstructions use two feedback models: a deterministic, continuous-time model and a stochastic, repeated-game model with bounded recall. In the first model, a continuous-time learning rule is written as $\dot{\rho}_i = F_i(\rho; \Xi_i)$. It is *uncoupled* if F_i may use player i 's own observable Ξ_i and current state information, but not the observables Ξ_j of other players.

In the second model, if Q_i is the set of pure quantum actions available to player i , and $Q = \prod_{i=1}^n Q_i$ is the set of realized pure quantum action profiles, then a quantum strategy for player i is a sequence $F_t^i : Q^{t-1} \rightarrow D(\mathcal{H}_i)$ such that $F^i = (F_t^i)_{t \geq 1}$, so that, after a history $h_{t-1} = (q(1), \dots, q(t-1)) \in Q^{t-1}$, player i chooses the mixed quantum action $F_t^i(h_{t-1}) \in D(\mathcal{H}_i)$. A quantum strategy mapping is a rule $\mathcal{F}$ assigning to each quantum payoff game $\Xi = (\Xi_1, \dots, \Xi_n)$ a quantum strategy profile $\mathcal{F}(\Xi) = (F^1(\Xi), \dots, F^n(\Xi))$. The strategy mapping can be regarded as a stochastic, discrete-time generalization of continuous-time learning dynamics. The mapping is uncoupled if each player's induced strategy depends only on that player's own payoff observable, namely $F^i(\Xi_1, \dots, \Xi_n) = F^i(\Xi_i)$, or equivalently, for every t and every history h_{t-1} , $F_t^i(h_{t-1}; \Xi_1, \dots, \Xi_n) = F_t^i(h_{t-1}; \Xi_i)$. Finally, a quantum strategy has R -recall if, for every $t > R$, the current choice depends only on the last R realized pure action profiles:

$$F_t^i(h_{t-1}; \Xi_i) = F_t^i(q(t-R), \dots, q(t-1); \Xi_i).$$

If this rule is independent of t , the strategy is stationary, and we write $F^i(q(t-R), \dots, q(t-1); \Xi_i)$. Thus, an uncoupled, stationary, R -recall quantum strategy mapping is a game-to-strategy-profile rule in which each player responds only to their own payoff observable and the most recent R realized pure quantum action profiles.

3 PRIOR WORKS

The equilibrium concept follows Nash's finite-game fixed-point theorem [Nas50]. Learning-in-games asks whether repeated strategic adjustment can reach this equilibrium without centralized information; broad background is given in [FL98, San10]. As discussed above, Hart and Mas-Colell proved that uncoupled dynamics do not universally lead to Nash equilibrium [HMC03]. Foster and Young obtained positive payoff-based convergence-in-behavior results via regret testing [FY06], while Hofbauer and Sandholm established global convergence of stochastic fictitious play for important special classes of games [HS02]. Marden, Young, Arslan, and Shamma developed payoff-based dynamics for multiplayer weakly acyclic games, a setting motivated by distributed control [MYAS09]. Most directly, Toonsi and Shamma introduced higher-order uncoupled gradient-play dynamics and showed that isolated completely mixed equilibria can be locally stabilized under a nonsingularity condition, even though universal guarantees remain impossible [TS23].

Quantum game theory was initiated, in part, by Meyer and by Eisert, Wilkens, and Lewenstein, who showed how allowing quantum strategies can change classical strategic conclusions [Mey99, EWL99]. Our formulation differs from protocol-specific quantizations by taking density matrices as strategies and Hermitian observables as payoffs. Methodologically, the paper also uses projected dynamical systems and variational-inequality ideas [DN93], as well as differentiability properties of metric projections onto convex sets [Sha16]. The main novelty is to combine these strands: uncoupled learning, higher-order stabilization, and the face geometry of quantum state spaces. A recent literature develops no-regret and matrix-multiplicative-weights dynamics for quantum games, establishing regret bounds, recurrence, and time-average convergence [LMB23a, JPS22, LMB23b, LMB23, LPSV24, LPSV25, VVGM⁺25]; our focus is instead on uncoupled *local* Nash stabilization at boundary density-matrix equilibria.

4 OUR CONTRIBUTIONS AND METHODOLOGY

1. A local deterministic obstruction. For the first obstruction, we follow the construction of the three-player game in Section 3 of [HMC03], in which each player controls one qubit. Write a qubit as $\rho(r_i) = \frac{1}{2}(I + r_i \cdot \sigma)$, where $r_i \in \mathbb{R}^3$, $\|r_i\| \leq 1$, and $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the vector of Pauli matrices. For players $i \in \mathbb{Z}_3$ and parameters $\theta_i \in \mathbb{R}^3$, define the utility as $u_{i,\theta}(r_1, r_2, r_3) = r_i \cdot (\theta_i - r_{i+1})$. The corresponding payoff observables Ξ_i^θ for each player i are:

$$\begin{aligned}\Xi_1^\theta &= (\theta_1 \cdot \sigma) \otimes I \otimes I - (\sigma_x \otimes \sigma_x \otimes I + \sigma_y \otimes \sigma_y \otimes I + \sigma_z \otimes \sigma_z \otimes I), \\ \Xi_2^\theta &= I \otimes (\theta_2 \cdot \sigma) \otimes I - (I \otimes \sigma_x \otimes \sigma_x + I \otimes \sigma_y \otimes \sigma_y + I \otimes \sigma_z \otimes \sigma_z), \\ \Xi_3^\theta &= I \otimes I \otimes (\theta_3 \cdot \sigma) - (\sigma_x \otimes I \otimes \sigma_x + \sigma_y \otimes I \otimes \sigma_y + \sigma_z \otimes I \otimes \sigma_z).\end{aligned}$$

Here, $\theta_i \cdot \sigma = \sum_k (\theta_i)_k \sigma_k$. This game has the unique equilibrium $r_i^* = \theta_{i-1}$. However, no continuously differentiable uncoupled dynamic can be Nash-convergent on the whole local family while rendering every equilibrium locally exponentially stable. At $\theta = 0$, changing θ_{i-1} moves player i 's equilibrium but does not change player i 's own observable Ξ_i^θ . Uncoupledness therefore forces the corresponding diagonal block of the Jacobian to vanish for each player. The full Jacobian has trace zero, which is incompatible with all eigenvalues having negative real parts.

2. A stochastic obstruction that is not tied to the computational basis. Let $A = \{\alpha, \beta, \gamma\}$ be the classical action set in the Hart–Mas-Colell three-player example [HMC06]. We encode each action $a \in A$ as the Fourier-basis state $|\tilde{a}\rangle = \text{QFT}_3|a\rangle$. Given a classical payoff $u_i : A^3 \rightarrow \mathbb{R}$, define the quantum observable $\Xi_i = \sum_{a \in A^3} u_i(a) \left(\bigotimes_{m=1}^3 |\tilde{a}_m\rangle\langle\tilde{a}_m| \right)$. Although Ξ_i is generally non-diagonal

in the computational basis, it assigns the same payoff to every encoded pure profile and therefore has the same pure best-reply graph as the classical game; hence the classical impossibility argument of [HMC06] transfers verbatim. This proves that no uncoupled, stationary, one-recall quantum strategy mapping can guarantee almost-sure convergence to a pure quantum Nash equilibrium across all finite quantum games that encode a classical game in an orthonormal (not necessarily computational) product basis.

3. Higher-order quantum gradient dynamics for boundary Nash equilibria. In [TS23], the authors show that for any polymatrix game with an isolated completely mixed Nash equilibrium, there exist (classical) higher-order gradient-play dynamics that converge to that equilibrium. We show that this result can be extended to the quantum setting by reducing the stabilization of a quantum dynamic to a classical linear stabilization problem. The main issue for a boundary Nash equilibrium is that the spectraplex D_i is not locally smooth across different support faces. Hence the stabilization problem depends on whether the equilibrium support is locally fixed or whether nearby projected-gradient directions may activate additional zero-eigenvalue directions.

In the non-degenerate boundary case, the Nash equilibrium $\rho^* = (\rho_1^*, \dots, \rho_n^*)$ is such that each ρ_i^* is full-rank on its support subspace $S_i = \text{supp}(\rho_i^*)$, and all directions outside S_i are strictly suboptimal. Equivalently, the projection locally remains on the fixed face $\mathcal{F}_i = \{\rho_i \in D_i : \text{supp}(\rho_i) \subseteq S_i\}$. Thus the boundary problem reduces to an interior problem on the smaller spectraplex $\mathcal{F}_i$. Let

$$\mathcal{T}(S_i) = \{X = X^\dagger : \text{Tr}(X) = 0, \text{supp}(X) \subseteq S_i\}$$

be the Hermitian, trace-zero tangent space on this face, and let $\mathcal{N}_i$ be a real orthonormal coordinate map for $\mathcal{T}(S_i)$. The corresponding higher-order quantum gradient-play dynamics take the form

$$\begin{aligned}\dot{\rho}_i &= -\rho_i + \Pi_{D_i} \left(\rho_i + \Gamma_i(\rho_{-i}) + \mathcal{N}_i \left[G_i \xi_i + M_i (\mathcal{N}_i^\dagger \Gamma_i(\rho_{-i}) - v_i) \right] \right), \\ \dot{\xi}_i &= E_i \xi_i + K_i (\mathcal{N}_i^\dagger \Gamma_i(\rho_{-i}) - v_i), \\ \dot{v}_i &= \mathcal{N}_i^\dagger \Gamma_i(\rho_{-i}) - v_i.\end{aligned}$$

The signal $\mathcal{N}_i^\dagger \Gamma_i(\rho_{-i}) - v_i$ is a support-restricted washout signal. If the support-restricted payoff Jacobian is nonsingular, then the local closed-loop system reduces to the same real tangent-space stabilization problem as in higher-order classical gradient play. Hence, by choosing suitable uncoupled controller matrices E_i, K_i, G_i, M_i , the given boundary Nash equilibrium can be locally exponentially stabilized.

The degenerate boundary case is more delicate. Here some direction outside the equilibrium support is payoff-tied with the support directions. In classical notation this corresponds to $S_i \subsetneq B_i$, where B_i is the set of best responses at equilibrium; in the quantum setting, the analogous phenomenon is that an orthogonal direction outside $\text{supp}(\rho_i^*)$ has the same first-order payoff as the support. Then the projection is no longer described by a single smooth linearization, but by a critical-face, piecewise-linear system. A simple way to recover the non-degenerate structure is to introduce a support-selective tie-breaking perturbation. Let $\mathcal{P}_i$ denote the orthogonal projection onto the target support. One may replace $\Gamma_i(\rho_{-i})$ in the projected-gradient input by

$$\widehat{\Gamma}_i(\rho_{-i}) = \Gamma_i(\rho_{-i}) - \varepsilon_i(I - \mathcal{P}_i), \quad \varepsilon_i > 0.$$

Since $(I - \mathcal{P}_i)$ lies entirely in the orthogonal complement of the support, we have $\mathcal{N}_i^\dagger (I - \mathcal{P}_i) = 0$ and $\mathcal{N}_i^\dagger \widehat{\Gamma}_i(\rho_{-i}) = \mathcal{N}_i^\dagger \Gamma_i(\rho_{-i})$. This leaves the support-restricted payoff unchanged but strictly penalizes every outside-support direction. Consequently, the degenerate equilibrium becomes effectively non-degenerate relative to the chosen support face. The higher-order support-restricted stabilization argument then applies as before, yielding local convergence to the prescribed boundary Nash equilibrium while avoiding face-switching among tied unused directions.

REFERENCES

- [DN93] Paul Dupuis and Anna Nagurney. Dynamical systems and variational inequalities. *Annals of Operations Research*, 44:7–42, Feb 1993.
- [EWL99] Jens Eisert, Martin Wilkens, and Maciej Lewenstein. Quantum games and quantum strategies. *Phys. Rev. Lett.*, 83:3077–3080, Oct 1999.
- [FL98] Drew Fudenberg and David K. Levine. *The Theory of Learning in Games*. MIT Press, 1998.
- [FY06] Dean P. Foster and H. Peyton Young. Regret testing: Learning to play nash equilibrium without knowing you have an opponent. *Theoretical Economics*, 1(3):341–367, 2006.
- [HMC03] Sergiu Hart and Andreu Mas-Colell. Uncoupled dynamics do not lead to nash equilibrium. *The American Economic Review*, 93(5):1830–1836, 2003.
- [HMC06] Sergiu Hart and Andreu Mas-Colell. Stochastic uncoupled dynamics and nash equilibrium. *Games and Economic Behavior*, 57(2):286–303, 2006.
- [HS02] Josef Hofbauer and William H. Sandholm. On the global convergence of stochastic fictitious play. *Econometrica*, 70(6):2265–2294, 2002.
- [JPS22] Rahul Jain, Georgios Piliouras, and Ryann Sim. Matrix multiplicative weights updates in quantum zero-sum games: Conservation laws and recurrence. In *Advances in Neural Information Processing Systems*, volume 35, pages 4123–4135. Curran Associates, Inc., 2022.
- [LMB23a] Kyriakos Lotidis, Panayotis Mertikopoulos, and Nicholas Bambos. Learning in quantum games. *arXiv preprint arXiv:2302.02333*, 2023.
- [LMB23b] Kyriakos Lotidis, Panayotis Mertikopoulos, and Nicholas Bambos. The stability of matrix multiplicative weights dynamics in quantum games. In *Proceedings of the 62nd IEEE Conference on Decision and Control*, pages 1225–1232. IEEE, 2023.
- [LMBB23] Kyriakos Lotidis, Panayotis Mertikopoulos, Nicholas Bambos, and Jose H. Blanchet. Payoff-based learning with matrix multiplicative weights in quantum games. In *Advances in Neural Information Processing Systems*, volume 36, pages 37942–37974. Curran Associates, Inc., 2023.
- [LPSV24] Wayne Lin, Georgios Piliouras, Ryann Sim, and Antonios Varvitsiotis. No-Regret Learning and Equilibrium Computation in Quantum Games. *Quantum*, 8:1569, December 2024.
- [LPSV25] Wayne Lin, Georgios Piliouras, Ryann Sim, and Antonios Varvitsiotis. Learning in Quantum Common-Interest Games and the Separability Problem. *Quantum*, 9:1689, April 2025.
- [Mey99] David A. Meyer. Quantum strategies. *Phys. Rev. Lett.*, 82:1052–1055, Feb 1999.
- [MYAS09] Jason Marden, H. Young, Grdal Arslan, and Jeff Shamma. Payoff-based dynamics for multiplayer weakly acyclic games. *SIAM Journal on Control and Optimization*, 48:373–396, 02 2009.
- [Nas50] John F. Nash. Equilibrium points in n -person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950.
- [San10] William H. Sandholm. *Population Games and Evolutionary Dynamics*. MIT Press, 2010.
- [Sha16] Alexander Shapiro. Differentiability properties of metric projections onto convex sets. *Journal of Optimization Theory and Applications*, 169:953–964, 2016.
- [TS23] Sarah A. Toonsi and Jeff S. Shamma. Higher-order uncoupled dynamics do not lead to nash equilibrium — except when they do. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS ’23, Red Hook, NY, USA, 2023. Curran Associates Inc.
- [VVG⁺25] Francisca Vasconcelos, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Panayotis Mertikopoulos, Georgios Piliouras, and Michael I. Jordan. A Quadratic Speedup in Finding Nash Equilibria of Quantum Zero-Sum Games. *Quantum*, 9:1737, May 2025.