The Stability of Matrix Multiplicative Weights Dynamics in Quantum Games

Kyriakos Lotidis*, Panayotis Mertikopoulos♯ and Nicholas Bambos§

Abstract— **In this paper, we study the equilibrium convergence and stability properties of the widely used matrix multiplicative weights (MMW) dynamics for learning in general quantum games. A key difficulty in this endeavor is that the induced quantum state dynamics decompose naturally into (*i*) a classical, commutative component which governs the dynamics of the system's eigenvalues in a way analogous to the evolution of mixed strategies under the classical replicator dynamics; and (*ii*) a non-commutative component for the system's eigenvectors. This non-commutative component has no classical counterpart and, as a result, requires the introduction of novel notions of (asymptotic) stability to account for the nonlinear geometry of the game's quantum space. In this general context, we show that (*i*) only *pure quantum equilibria* can be stable and attracting under the MMW dynamics; and (*ii*) as a partial converse, pure quantum states that satisfy a certain "variational stability" condition are always attracting. This allows us to fully characterize the structure of quantum Nash equilibria that are stable and attracting under the MMW dynamics, a fact with significant implications for predicting the outcome of a multi-agent quantum learning process.**

I. Introduction

The advent of quantum information theory – and, with it, the associated "quantum advantage" [1]–[3] – has had a profound impact on computer science and machine learning, from quantum cryptography and shadow tomography [4], to quantum generative adversarial networks (QGANs) and adversarial learning [5]–[7]. At a high level, the advantages of quantum-based computing are owed to the possibility of preparing superpositions of binary-state quantum systems known as *qubits*: classical bits cannot lie in superposition, so the calculations that can be performed by classical computers are *de facto* limited by their binary alphabet and memory structure. In light of this, quantum computing has the potential to greatly accelerate the development of artificial intelligence algorithms and models, with Google's "Sycamore" 54-qubit processor training an autonomous vehicle model in less than 200 seconds [2].

In a similar manner, when such models are deployed in a multi-agent context – e.g., as in the case of QGANs or autonomous vehicles – the landscape changes drastically relative to classical non-cooperative frameworks. The main reason for this is again the "quantum advantage": due to the intricacies of decoherence and entaglement – two quantum notions that have no classical counterpart – quantum players can have a distinct advantage over "classical" players, achieving higher payoffs at equilibrium than would otherwise be possible [8], [9]. This is again owed to the fact that probabilistic mixing works differently in the quantum and classical worlds: in classical games, a mixed *strategy* is a probabilistic convex combination of the constituent pure strategies; in quantum games, a mixed *state* is a probabilistic mixture of the quantum *projectors* associated to each constituent state. Because of this, a mixed quantum state can return payoffs that lie outside the convex hull of classical mixed strategies, thus providing a tangible advantage to players with access to quantum technologies.

Of course, the extent to which the advantage of quantum players manifests itself is contingent on the players' actually reaching an equilibrium. The recent work of [10] has shown that the problem of computing an approximate Nash equilibrium of a quantum game is included in PPAD, so, by the seminal work of [11], [12], it must be complete for this class (since computing a quantum equilibrium is at least as hard as computing a classical one). Thus, given that the dimensionality of a quantum game is exponential in the number of qubits available to each player, computing a Nash equilibrium of a quantum game quickly becomes an intractable affair, in all but the smallest games. On that account, it seems more reasonable to turn to an online learning paradigm, and instead ask:

- *Are all Nash equilibria equally likely to occur as outcomes of a multi-agent learning procedure?*
- *Is there a class of Nash equilibria with an inherent selection bias – either for or against?*

Our point of departure for these questions is a compact model of learning based on the widely used *matrix multiplicative weights* (MMW) algorithm, originally due to [13]. This learning process can be seen as a quantum extension of the standard Hedge / EXP3 algorithms for single-agent learning in bandits and games [14]–[16], and it has been used extensively for the computation of quantum Nash equilibria in two-player, zero-sum quantum games [17]–[22]. To spotlight the structural properties of the dynamics and to sidestep the technical issues that arise from hyperparameter tuning and the like, we focus throughout on a parameter-free, continuous-time formulation of the dynamics, and we seek to characterize which quantum Nash equilibria are stable and attracting under these dynamics in general quantum games.

*Dept. of Management Science & Engineering, Stanford University, klotidis@stanford.edu.

♯Univ. Grenoble-Alpes, CNRS, Inria, Grenoble INP, LIG, 38000, France panayotis.mertikopoulos@imag.fr.

§Dept. of Electrical Engineering and Management Science & Engineering, Stanford University, bambos@stanford.edu.

P. Mertikopoulos was supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program, and by the French National Research Agency (ANR) in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02), the LabEx PERSYVAL (ANR-11-LABX-0025-01), MIAI@Grenoble Alpes (ANR-19-P3IA-0003). K. Lotidis was supported by the Onassis Foundation (F ZR 033-1/2021-2022). N. Bambos was supported by the Koret Foundation via the Digital Living 2030 project.

Importantly, the quantum state dynamics induced by the MMW learning rule decompose naturally into a "classical" part plus a non-commutative "quantum" component: the eigenvalues of the system follow an equation that is formally analogous to the classical replicator dynamics for finite games, while the system's eigenvectors follow an autonomous flow on the space of unitary matrices that has no classical analogue. Because of this non-commutative structure, deriving the dynamics' equilibrium convergence properties is significantly more difficult because of the nonlinear geometry of the game's state space. In particular, in contrast to finite games (where pure strategies are isolated extreme points), the pure states of a quantum game form a continuous manifold of stationary points (all of them extreme), so the study of stability and convergence questions becomes a highly involved affair.

Nonetheless, despite these topological complications, we show that the MMW dynamics enjoy the following fundamental properties: (i) only *pure quantum equilibria* can be stable and attracting under MMW; and (ii) as a partial converse, pure quantum states that satisfy a certain "variational stability" condition are always stable and attracting. On that account, our results lead to an implicit quantum "purification" principle: under MMW, *mixed states are inherently fragile, and only pure quantum states can be consistently attracting.*[1] This fact has significant implications for predicting the outcome of a multi-agent quantum learning process, and it opens up several research directions with potentially far-reaching implications for the deployment of multi-agent quantum systems

II. Preliminaries

We start by briefly reviewing some basics of quantum game theory and introducing the necessary context for our results.

Notation: Given a (complex) Hilbert space $\mathcal{H}$, we will use Dirac's bra-ket notation to distinguish between an element $|\psi\rangle$ of $\mathcal{H}$ and its adjoint $\langle\psi|$; otherwise, when a basis is implied by the context, we will use the dagger notation "†" to denote the Hermitian transpose $\psi^\dagger$ of ψ. We will also write $\mathbb{H}^d$ for the space of $d\times d$ Hermitian matrices, and $\mathbb{H}^d_+$ for the cone of positive-semidefinite matrices in $\mathbb{H}^d$. Finally, given a real function $f\colon \mathbb{R}\to\mathbb{R}$ and a Hermitian matrix $\mathbf{X}\in\mathbb{H}^d$ with unitary eigen-decomposition $\mathbf{X}=\sum_{\alpha=1}^d x_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha^\dagger$, we will write $f(\mathbf{X})$ for the matrix $f(\mathbf{X})=\sum_{\alpha=1}^d f(x_\alpha)\mathbf{u}_\alpha\mathbf{u}_\alpha^\dagger$.

Quantum games: Following [9], [27], a *quantum game* consists of the following primitives:

1) A finite set of *players* $i\in\mathcal{N}=\{1,\dots,N\}$.

2) Each player $i\in\mathcal{N}$ has access to a complex Hilbert space $\mathcal{H}_i\cong\mathbb{C}^{d_i}$ describing the set of (pure) *quantum states* available to the player (typically a discrete register of qubits). In more detail, a quantum state is an element ψ_i of $\mathcal{H}_i$ with unit norm, so the set of all such states is the unit sphere $\Psi_i \coloneqq \{\psi_i\in\mathcal{H}_i : \|\psi_i\|=1\}$ of $\mathcal{H}_i$. We will write $\Psi\coloneqq\prod_i\Psi_i$ for the space of all ensembles $\psi=(\psi_1,\dots,\psi_N)$ of pure states $\psi_i\in\Psi_i$ that are independently prepared by each $i\in\mathcal{N}$.

3) The players' rewards are determined by their individual *payoff functions* $u_i\colon\Psi\to\mathbb{R}$. These payoff functions are not arbitrary, but are obtained from a joint *positive operator-valued measure* (POVM) quantum measurement process that unfolds as follows [28]: First, we assume given a *finite* set of possible *measurement outcomes* $\omega\in\Omega$ that a referee can observe from the players' quantum states (e.g., measure a player-prepared qubit to be "up" or "down"). Each such outcome $\omega\in\Omega$ is associated to a positive semi-definite operator $\mathbf{P}_\omega\colon\mathcal{H}\to\mathcal{H}$ that acts on the tensor product $\mathcal{H}\coloneqq\bigotimes_i\mathcal{H}_i$ of the players' individual state spaces; we further assume that $\sum_{\omega\in\Omega}\mathbf{P}_\omega=\mathbf{I}$ so the joint probability of observing $\omega\in\Omega$ when the system is at state $\psi\in\Psi$ is

$$P_\omega(\psi)=\langle\psi_1\otimes\dots\otimes\psi_N|\mathbf{P}_\omega|\psi_1\otimes\dots\otimes\psi_N\rangle \tag{1}$$

The payoff to each player $i\in\mathcal{N}$ is given by the outcome of this measurement process via a *payoff observable* $U_i\colon\Omega\to\mathbb{R}$; specifically, in this context, $u_i(\psi)$ denotes the player's expected payoff at state $\psi\in\Psi$, viz.

$$u_i(\psi)\coloneqq\langle U_i\rangle\equiv\sum\nolimits_\omega P_\omega(\psi)\,U_i(\omega). \tag{2}$$

A *quantum game* is then defined as a tuple $\mathcal{Q}\equiv\mathcal{Q}(\mathcal{N},\Psi,u)$ with players, quantum states, and payoff functions as above.

Mixed states: In addition to pure states, each player $i\in\mathcal{N}$ can also prepare probabilistic mixtures thereof, known as *mixed states*. In contrast to mixed strategies in classical, finite games, these mixed states are *not* convex combinations of their pure counterparts; instead, given a family of pure quantum states $\psi_{i\alpha_i}\in\Psi_i$ indexed by $\alpha_i\in\mathcal{A}_i$, a mixed state is described by a *density matrix* of the form

$$\mathbf{X}_i=\sum_{\alpha_i\in\mathcal{A}_i}x_{i\alpha_i}|\psi_{i\alpha_i}\rangle\langle\psi_{i\alpha_i}| \tag{3}$$

where $x_{i\alpha_i}\geq 0$ is the mixing weight of $\psi_{i\alpha_i}$, and we assume that $\operatorname{tr}\mathbf{X}_i=1$ (the states $\psi_{i\alpha_i}$ are not assumed to be orthogonal in this context). By Born's rule, this means that if each player $i\in\mathcal{N}$ prepares a mixed state according to $\mathbf{X}_i$, the probability of observing $\omega\in\Omega$ under $\mathbf{X}=(\mathbf{X}_1,\dots,\mathbf{X}_N)$ will be $P_\omega(\mathbf{X})=\sum_\alpha x_\alpha\langle\psi_\alpha|\mathbf{P}_\omega|\psi_\alpha\rangle$, where, in multi-index notation, $\alpha=(\alpha_1,\dots,\alpha_N)$, $x_\alpha=\prod_i x_{i\alpha_i}$, and $\psi_\alpha=\bigotimes_i\psi_{i\alpha_i}$. Thus, in a slight abuse of notation, the expected payoff to player $i\in\mathcal{N}$ under $\mathbf{X}$ will be

$$u_i(\mathbf{X})=\sum_{\omega\in\Omega}\sum_{\alpha\in\mathcal{A}}x_\alpha P_\omega(\psi_\alpha)\,U_i(\omega)=\sum_{\alpha\in\mathcal{A}}x_\alpha u_i(\psi_\alpha) \tag{4}$$

Equivalently, we can write the above as:

$$\begin{aligned}u_i(\mathbf{X})&=\sum_{\omega\in\Omega}U_i(\omega)\operatorname{tr}[\mathbf{P}_\omega\mathbf{X}_1\otimes\dots\otimes\mathbf{X}_N]\\&=\operatorname{tr}[\mathbf{W}_i\mathbf{X}_1\otimes\dots\otimes\mathbf{X}_N]\end{aligned} \tag{5}$$

where the tensor $\mathbf{W}_i=\sum_{\omega\in\Omega}U_i(\omega)\mathbf{P}_\omega\in\mathcal{H}$ for $i\in\mathcal{N}$ encloses all the relevant payoff information of the game and is the quantum equivalent of the "payoff matrix" of player $i\in\mathcal{N}$. [In this regard, (5) gives a more concrete and concise representation of the payoff structure of $\mathcal{Q}$.]

[1] In classical games, a version of the above results is sometimes referred to as the "folk theorem" of evolutionary game theory [23]–[26].

Contrasting to other classes of games: The expression (4) for a player's expected payoff under a mixed state is reminiscent of mixed extensions of classical finite games, but this association is very tenuous. From a conceptual standpoint, the principal differences are as follows:

1) There is an infinite continuum of pure states $\psi \in \Psi$, not a finite number thereof (as is the case in finite games).
2) The decomposition (3) of a density matrix into pure states is not unique; generically, there may be a continuum of (non-equivalent) families of pure states and mixing weights giving rise to the same density matrix.
3) The convex superposition $\lambda\psi + (1-\lambda)\psi'$ of two pure states ψ and ψ' may give rise to quantum interference terms of the form $|\psi\rangle\langle\psi'|$ and $|\psi'\rangle\langle\psi|$ in the induced payoff; these cross-terms have no analogue in finite games.

Continuous game reformulation: Because of the previous discussion, treating a quantum game as a "tensorial" extension of a finite game can be misleading. Instead, it would be clearer for our purposes to view a quantum game as a *continuous game* where each player $i \in \mathcal{N}$ controls a matrix variable $\mathbf{X}_i$ drawn from the "spectraplex"

$$\boldsymbol{\mathcal{X}}_i = \{\mathbf{X}_i \in \mathbb{H}_+^{d_i} : \operatorname{tr} \mathbf{X}_i = 1\} \tag{6}$$

and the player's payoff function $u_i : \boldsymbol{\mathcal{X}} \equiv \prod_j \boldsymbol{\mathcal{X}}_j \to \mathbb{R}$ is linear in every player's density matrix $\mathbf{X}_j \in \boldsymbol{\mathcal{X}}_j$, $j \in \mathcal{N}$. Since u_i is linear in $\mathbf{X}_i$, letting $\mathbf{V}_i(\mathbf{X}) = \nabla_{\mathbf{X}_i^\top} u_i(\mathbf{X})$ denote the individual payoff gradient of player i, we can write the payoff of player i in terms of the gradient $\mathbf{V}_i$ as:

$$u_i(\mathbf{X}_i; \mathbf{X}_{-i}) = \operatorname{tr}[\mathbf{X}_i \mathbf{V}_i(\mathbf{X})] \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{X}}. \tag{7}$$

Nash equilibrium: In our quantum setting, the classical solution concept of a *Nash equilibrium* (NE) characterizes mixed quantum states $\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$ which discourage unilateral deviations in the sense that

$$u_i(\mathbf{X}^*) \geq u_i(\mathbf{X}_i; \mathbf{X}^*_{-i}) \quad \text{for all } \mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i,\ i \in \mathcal{N} \tag{NE}$$

where we write $(\mathbf{X}_i; \mathbf{X}_{-i}) = (\mathbf{X}_1, \dots, \mathbf{X}_i, \dots, \mathbf{X}_N)$ for the choice of player i relative to all other players. Since $\boldsymbol{\mathcal{X}}_i$ is convex and u_i is linear in $\mathbf{X}_i$, the existence of Nash equilibria follows from the seminal theorem of [29]. Standard arguments [30], [31] show that the Nash equilibria of $\mathcal{Q}$ are precisely the solutions of the variational inequality

$$\operatorname{tr}[\mathbf{V}(\mathbf{X}^*)(\mathbf{X} - \mathbf{X}^*)] \leq 0 \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{X}} \tag{VI}$$

where $\mathbf{V}(\mathbf{X}) = (\mathbf{V}_1(\mathbf{X}), \dots, \mathbf{V}_N(\mathbf{X}))$. We will use this equivalence freely in the sequel.

III. Learning dynamics

The matrix multiplicative weights algorithm: The most widely used algorithm for computing Nash equilibria of quantum games is the so-called *matrix multiplicative weights* (MMW) algorithm which, in our notation, unfolds as

$$\begin{aligned} \mathbf{Y}_i(t+1) &= \mathbf{Y}_i(t) + \eta \mathbf{V}_i(\mathbf{X}(t)) \\ \mathbf{X}_i(t) &= \frac{\exp(\mathbf{Y}_i(t))}{\operatorname{tr}[\exp(\mathbf{Y}_i(t))]} \end{aligned} \tag{MMW}$$

where $\eta > 0$ is a "learning rate" hyperparameter. This algorithm was first introduced in the context of matrix learning by [13], and can be seen as a matrix analogue of the well-known "multiplicative / exponential weights" (EW) algorithm for learning in bandits and games [16], [32]. In the context of quantum games, [17] showed that (MMW) can be used to compute the Nash equilibrium of two-player, zero-sum quantum games by running it for T iterations and taking the time-average $\bar{\mathbf{X}} = (1/T) \sum_{t=1}^{T} \mathbf{X}(t)$ of the generated states; if the learning rate of (MMW) is chosen appropriately – specifically, as $\eta = \mathcal{O}(1/\sqrt{T})$ – the algorithm's output state $\bar{\mathbf{X}}$ is an $\mathcal{O}(1/\sqrt{T})$-equilibrium of the underlying game.

Beyond the two-player, zero-sum case however, the behavior of (MMW) in general quantum games is not well understood. Moreover, an additional limitation from a learning viewpoint is that the guarantees of (MMW) concern the time-averaged state $\bar{\mathbf{X}}$ and not the induced sequence of play $\mathbf{X}(t)$. The behavior of the former can be quite different from the latter, even in min-max quantum games; in particular, as was shown in [19], even in the continuous-time limit where the algorithm's learning rate is taken arbitrarily small, $\mathbf{X}(t)$ may fail to converge altogether.

In view of this, we will take an approach similar to [19], and we will focus on the continuous-time limit of (MMW), namely the dynamics

$$\dot{\mathbf{Y}}_i = \mathbf{V}_i(\mathbf{X}) \qquad \mathbf{X}_i = \mathbf{\Lambda}(\mathbf{Y}_i) \coloneqq \frac{\exp(\mathbf{Y}_i)}{\operatorname{tr}[\exp(\mathbf{Y}_i)]} \tag{MMWD}$$

with $\mathbf{Y}_i \in \boldsymbol{\mathcal{Y}}_i \coloneqq \mathbb{H}^{d_i}$, as per (MMW). The benefit of this reformulation is that it allows us to focus squarely on the dynamics' structural properties without being bogged down by hyperparameter questions and the like. We begin our analysis below by deriving the induced dynamics of the players' mixed quantum states $\mathbf{X}_i \in \boldsymbol{\mathcal{X}}_i$.

The quantum state dynamics of MMW: Under (MMWD), the evolution of the players' mixed states $\mathbf{X}(t)$ is described *implicitly* via that of the auxiliary matrix $\mathbf{Y}(t)$. However, obtaining an *explicit* expression for the dynamics of $\mathbf{X}(t)$ is considerably more difficult because the rules of matrix calculus do not provide an analytic expression for the tensor derivative. Following [33], we will circumvent this difficulty by taking a unitary eigendecomposition of $\mathbf{X}$ of the form $\mathbf{X} = \sum_{\alpha=1}^{d} x_\alpha \mathbf{u}_\alpha \mathbf{u}_\alpha^\dagger$, where $x_\alpha \geq 0$, $\mathbf{u}_\alpha \in \mathcal{H}$ is a unit-norm eigenvector of $\mathbf{X}$ corresponding to x_α. Since, in general, $\mathbf{X}$ and $\mathbf{V}$ do not commute, the eigenvalues and eigenvectors of $\mathbf{X}$ will evolve in a coupled manner determined by (MMWD).

To make this coupling explicit, we will suppress player indices for notational simplicity, and we will differentiate $\mathbf{X}$ in (MMWD) with respect to t. Doing this we obtain:

$$\dot{\mathbf{X}} = \frac{1}{\operatorname{tr}[\exp(\mathbf{Y})]} \frac{d}{dt} \exp(\mathbf{Y}) - \frac{\exp(\mathbf{Y})}{\operatorname{tr}[\exp(\mathbf{Y})]^2} \operatorname{tr}\left[\frac{d}{dt} \exp(\mathbf{Y})\right]$$

and hence, by taking Fréchet derivatives [34], we get:

$$\frac{d}{dt} \exp(\mathbf{Y}) = \int_0^1 e^{(1-s)\mathbf{Y}} \dot{\mathbf{Y}} e^{s\mathbf{Y}}\, ds$$

$$= \mathrm{tr}[\exp(\mathbf{Y})] \int_0^1 \mathbf{X}^{1-s} \mathbf{V}(\mathbf{X}) \mathbf{X}^s \, ds \tag{8}$$

which, after some algebraic manipulations, ultimately yields

$$\begin{aligned} \dot{\mathbf{X}} &= \int_0^1 \mathbf{X}^{1-s} \mathbf{V}(\mathbf{X}) \mathbf{X}^s \, ds - \frac{\exp \mathbf{Y}}{\mathrm{tr}[\exp(\mathbf{Y})]} \int_0^1 \mathrm{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})] \, ds \\ &= \int_0^1 \mathbf{X}^{1-s} \mathbf{V}(\mathbf{X}) \mathbf{X}^s \, ds - \mathrm{tr}[\mathbf{X}\mathbf{V}(\mathbf{X})] \mathbf{X} \end{aligned} \tag{9}$$

Thus, applying $\mathbf{u}_\alpha^\dagger$ to the left and $\mathbf{u}_\beta$ to the right of (9), we get:

$$\mathbf{u}_\alpha^\dagger \dot{\mathbf{X}} \mathbf{u}_\beta = \int_0^1 x_\alpha^{1-s} V_{\alpha\beta} x_\beta^s \, ds - x_\alpha \delta_{\alpha\beta} \sum_\kappa x_\kappa V_{\kappa\kappa} \tag{10}$$

Finally, denoting $\mathbf{u}_\alpha^\dagger \dot{\mathbf{X}} \mathbf{u}_\beta$ by $[\dot{\mathbf{X}}]_{\alpha\beta}$, equation (10) gives the *quantum replicator dynamics* (QRD):

$$[\dot{\mathbf{X}}]_{\alpha\beta} = \begin{cases} x_\alpha [V_{\alpha\alpha} - \sum_\kappa x_\kappa V_{\kappa\kappa}] & \text{for } \alpha = \beta \\ \dfrac{x_\beta - x_\alpha}{\log x_\beta - \log x_\alpha} V_{\alpha\beta} & \text{for } \alpha \neq \beta \end{cases} \tag{QRD}$$

The diagonal part of (QRD) is formally analogous to the replicator dynamics of evolutionary game theory [35]–[37] and captures the evolution of the eigenvalues of $\mathbf{X}(t)$. Thus, (QRD) provides an explicit expression for the evolution of mixed states under (MMWD). Next, we will study in detail how the classical and quantum components of (MMWD) interface to determine the player's long-run behavior.

IV. Convergence and stability

We are now in a position to proceed with our convergence analysis. A key element to keep in mind here is that a quantum game may admit several Nash equilibria, so it is not reasonable to expect a global convergence result that applies to *all* games – this, in fact, would violate the impossibility result of [38]. In view of this, we will focus on the next best thing, that is, to identify those quantum states that are "locally stable and attracting" as formalized below.

Notions of stability and convergence: To state our results, recall first that a *flow* on an abstract metric space $\mathcal{Z}$ is a continuous map $\phi\colon \mathbb{R} \times \mathcal{Z} \to \mathcal{Z}$ such that (*a*) $\phi_0(z) = z$; and (*b*) $\phi_{t+s}(z) = \phi_t(\phi_s(z))$ for all $t, s \in \mathbb{R}$ and all $z \in \mathcal{Z}$. Informally, a flow is usually generated by the *solution orbits* of a system of well-posed ordinary differential equations (ODEs), such as (MMWD): in this interpretation, $\phi_t(z)$ simply denotes the position at time t of the ODE solution that starts at z at time $t = 0$.

With this in mind, given a point $p \in \mathcal{Z}$, we will say that

(i) p is (Lyapunov) *stable* if any orbit of (MMWD) that starts close enough to p remains close enough; formally, for every neighborhood $\mathcal{U}$ of p in $\mathcal{Z}$, we posit that there exists some (smaller) neighborhood $\mathcal{U}'$ of p in $\mathcal{Z}$ such that $\phi_t(\mathcal{U}') \subseteq \mathcal{U}$ for all $t \in \mathbb{R}$.

(ii) p is *attracting* if all nearby orbits converge to p; formally, there exists a neighborhood $\mathcal{U}$ such that $\lim_{t\to\infty} \phi_t(z) = p$ for all $z \in \mathcal{U}$.

In what follows, we will seek to characterize precisely the states that are "stable and attracting" under (MMWD) – or, in more formal language, *asymptotically stable*.

Learning in the spectraplex: Moving forward, a quick look at the quantum replicator dynamics (QRD) reveals the following structural property: an eigenvalue of $\mathbf{X}$ that is initially zero in (QRD) will always remain zero; likewise, an eigenvalue that is initially positive, will always remain positive. Formally, this means that the kernel $\ker(\mathbf{X})$ of $\mathbf{X}$ remains invariant under (QRD); hence, given that the linear span of a Hermitian matrix is the orthocomplement of its kernel, the same holds for $\mathrm{im}(\mathbf{X})$.

The fact that the kernel – or, equivalently, the image – of a density matrix remains invariant under (QRD) is the quantum analogue of the fact that the support of a mixed strategy profile remains invariant under the standard replicator dynamics. In the context of finite games, an immediate consequence of this invariance is that *all* pure strategy profiles are stationary (as zero-dimensional faces of the simplex). This property extends to (QRD) and, in fact, to the entire class of mixed-state dynamics under study: formally, under (QRD), *all pure quantum states are stationary.*

That being said, the major qualitative difference between the quantum and classical regimes is that, in quantum games, there is a *continuum* of pure states, namely the entire manifold Ψ of rank 1 density matrices (a product of spheres). By contrast, in finite games, the pure states are the corners of the simplex Δ spanned by the player's pure strategies, so they are finite in number and *isolated*. As a result, in classical finite games, a pure strategy profile *can* be asymptotically stable; in quantum games, since every pure state is surrounded by other invariant states, *it cannot*.

A second major difference is that, in finite games, strict Nash equilibria are *robust:* a small perturbation of the payoffs of the game does not change the game's strict equilibria. In quantum games, this robustness disappears: indeed, the variational characterization (VI) of Nash equilibria means that $\mathbf{V}(\mathbf{X}^*)$ must be an element of the normal cone to $\boldsymbol{\mathcal{X}}$ at $\mathbf{X}^*$; however, the normal cone to the spectraplex at a matrix of rank 1 has empty topological interior, so the required membership property cannot be robust (for a graphical illustration, see Fig. 1). In particular, any perturbation to the payoff structure of a quantum game may displace the equilibrium in question on the manifold of pure states Ψ.

Consistency and variational stability: In view of the above, we can draw two major conclusions for the quantum setting:

1) Any concept of asymptotic stability must also include a notion of *consistency:* a state cannot be accessed if it is absent from the linear span of the dynamics' initial state.
2) Any concept of robustness must incorporate a notion of *variational stability:* small perturbations to an equilibrium must tend to reinstate it.

We formalize these ideas as follows:

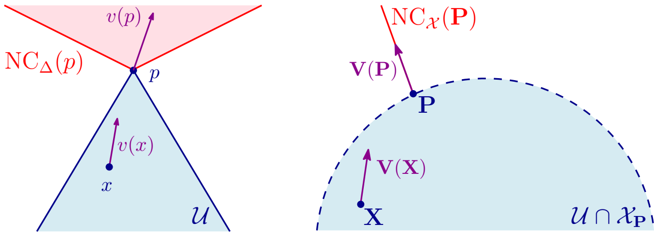

Fig. 1: The geometric discrepancy between the classical and quantum regimes (left and right respectively). In finite games, the normal cone $\mathrm{NC}_{\Delta}(p)$ to the simplex at a pure strategy p has nonempty topological interior, so pure Nash equilibria are generically *robust*: if $v(p)$ is normal to Δ at p, it will remain normal to Δ after a small perturbation. On the other hand, in quantum games, the normal cone $\mathrm{NC}_{\boldsymbol{\mathcal{X}}}(\mathbf{P})$ to the spectraplex at a pure state $\mathbf{P}$ is a ray, so pure Nash equilibria *are not robust*. We also note the different geometry of pure states: in the simplex, pure strategies are isolated extreme points; in the spectraplex, pure states form a continuous manifold.

Definition 1. The *domain of consistency* of a state $\mathbf{P} \in \boldsymbol{\mathcal{X}}$ is the set $\boldsymbol{\mathcal{X}}_{\mathbf{P}} \coloneqq \{\mathbf{X} \in \boldsymbol{\mathcal{X}} : \ker(\mathbf{X}) \leq \ker(\mathbf{P})\}$, i.e., the set of mixed states whose linear span contains that of $\mathbf{P}$. Then, given a flow $\chi\colon \mathbb{R} \times \boldsymbol{\mathcal{X}} \to \boldsymbol{\mathcal{X}}$, we will say that:

1) $\mathbf{P}$ is *consistently attracting* if it attracts all nearby *consistent* initializations, i.e., it admits a neighborhood $\boldsymbol{\mathcal{U}}$ such that $\lim_{t\to\infty} \chi_t(\mathbf{X}) = \mathbf{P}$ for all $\mathbf{X} \in \boldsymbol{\mathcal{U}} \cap \boldsymbol{\mathcal{X}}_{\mathbf{P}}$.
2) $\mathbf{P}$ is *consistently asymptotically stable* if it is stable and consistently attracting.

Finally, to ensure a certain degree of stability to perturbations, we will consider the following notion of variational stability:

Definition 2. We say that $\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$ is *variationally stable* if there exists a neighborhood $\boldsymbol{\mathcal{U}}$ of $\mathbf{X}^*$ in $\boldsymbol{\mathcal{X}}$ such that

$$\operatorname{tr}[\mathbf{V}(\mathbf{X})(\mathbf{X} - \mathbf{X}^*)] < 0 \quad \text{for all } \mathbf{X} \in \boldsymbol{\mathcal{U}} \setminus \{\mathbf{X}^*\}. \tag{VS}$$

Intuitively, Definition 1 captures precisely the accessibility condition that we discussed above, while Definition 2 should be seen as an equilibrium refinement in the spirit of the seminal concept of *evolutionary stability* [39], [40].

With all this in hand, our main result is as follows:

Theorem 1. *Fix some state $\mathbf{X}^* \in \boldsymbol{\mathcal{X}}$ of a quantum game $\mathcal{Q} \equiv \mathcal{Q}(\mathcal{N}, \Psi, u)$. Then:*

a) If $\mathbf{X}^*$ is consistently asymptotically stable, then it is pure.
b) If $\mathbf{X}^*$ satisfies (VS), it is consistently asymptotically stable.

Before presenting the proof of Theorem 1, some remarks are in order. Perhaps the most important one is that, if the standard notion of asymptotic stability is ruled out by the geometry of the game's state space, (MMWD) achieves the next best thing: by definition, states that are consistently asymptotically stable attract all but a measure zero of nearby initial conditions, and Theorem 1 shows that *only* pure states can have this property. This selection result has important implications for quantum games because it shows that the MMW rule essentially "collapses" an initial mixed state to a *specific* pure state – and this, despite the fact that any mixed state can be prepared by an infinity of combinations of pure states.

On the flip side of all this, the implication that variationally stable states are also (consistently) asymptotically stable provides a relevant convergence criterion for (MMWD) and indicates an inherent robustness to variations of player beliefs and predictions. In particular, since (VS) only involves the primitives of the underlying game, the fact that such states are attracting under the MMW dynamics means that they can be seen as *universal attractors* – and since only pure states can have this property, we also infer indirectly that variationally stable states are *a fortiori* pure.

The proof of the *first* part of Theorem 1 consists of three steps. First, we construct a measure μ on $\operatorname{ri} \boldsymbol{\mathcal{X}}$ that is invariant under the flow χ, i.e., $\mu(A) = \mu(\chi_t(A))$ for all measurable sets A of $\operatorname{ri} \boldsymbol{\mathcal{X}}$ and $t \geq 0$, where by $\chi_t(A)$ we denote the image of A after time t, and by $\operatorname{ri} \boldsymbol{\mathcal{X}}$ the relative interior of $\boldsymbol{\mathcal{X}}$. For this, we restrict the dynamics over the dual space $\boldsymbol{\mathcal{Y}}$ to a quotient space $\boldsymbol{\mathcal{Z}}$ that has the same dimension as $\boldsymbol{\mathcal{X}}$, and invoke a volume-conservation argument. Next, we show that if $\mathbf{X}^* \in \operatorname{ri} \boldsymbol{\mathcal{X}}$, it cannot be consistently asymptotically stable. To prove this, we need the following lemma, which, in words, says that if $\mathbf{X}^* \in \operatorname{ri} \boldsymbol{\mathcal{X}}$ is consistently asymptotically stable, then all the orbits starting nearby converge uniformly in time, and whose proof lies at the end of the section.

Lemma 1. *Suppose that $\mathbf{X}^* \in \operatorname{ri} \boldsymbol{\mathcal{X}}$ is consistently asymptotically stable under* (MMWD)*. Then, for every sufficiently small compact neighborhood $\boldsymbol{\mathcal{U}}$ of $\mathbf{X}^*$ in $\operatorname{ri} \boldsymbol{\mathcal{X}}$, we have*

$$\lim_{t\to\infty} \sup_{\mathbf{X} \in \boldsymbol{\mathcal{U}}} \lVert \chi_t(\mathbf{X}) - \mathbf{X}^* \rVert = 0. \tag{11}$$

Finally, we show that if $\mathbf{X}^*$ is consistently asymptotically stable, it is pure, using a reduction argument and considering faces of $\boldsymbol{\mathcal{X}}$ that contain $\mathbf{X}^*$ in their relative interior.

The first two steps preclude convergence to full-rank equilibria, while excluding lower-rank equilibria requires more delicate arguments, where the notion of consistency plays a major role (and has no classical counterpart), and is succeeded in the third step. With the above mind, we are now ready to proceed to the proof of Theorem 1.

Proof of Theorem 1(a). We will present our proof in three steps, as described earlier on.

Step 1: First, we will need to collapse the dual space $\boldsymbol{\mathcal{Y}}$ of $\mathcal{V} \equiv \mathbb{H}^d$ to a subspace $\boldsymbol{\mathcal{Z}}$ with constant trace. The reason for this is that the map $\mathbf{\Lambda}$ is not injective, since for all $\lambda \in \mathbb{R}$, we have $\mathbf{\Lambda}(\mathbf{Y} + \lambda \mathbf{I}) = \mathbf{\Lambda}(\mathbf{Y})$. This means that the inverse image of any point in $\operatorname{im} \mathbf{\Lambda}$ always contains a copy of the real line. This collapse is intended to "quotient out" this redundancy so as to enable volume comparisons later on. To succeed this, consider the transformed matrix variables

$$\mathbf{Z} = \mathbf{Y} - (1/d) \operatorname{tr}[\mathbf{Y}] \mathbf{I} \tag{12}$$

so $\operatorname{tr}\mathbf{Z} = 0$ for all $\mathbf{Y} \in \mathcal{Y}$. Formally, this transformation can be represented via the map $\mathbf{\Pi}\colon \mathcal{Y} \to \mathcal{Z}$, where

$$\mathcal{Z} = \{\mathbf{Z} \in \mathcal{Y} : \operatorname{tr}\mathbf{Z} = 0\} \tag{13}$$

and $\mathbf{\Pi}\colon \mathbf{Y} \mapsto \mathbf{\Pi}(\mathbf{Y}) = \mathbf{Z}$ is defined via (12) above. As such, $\mathbf{Z}$ can be seen as a representative of the equivalence relation $\mathbf{Y} \sim \mathbf{Y} + \lambda\mathbf{I}$, $\lambda \in \mathbb{R}$, so, in turn, $\mathcal{Z}$ can be identified with the quotient $\mathcal{Y}/\sim$. By the definition of $\mathbf{\Lambda}$, we have $\mathbf{\Lambda}(\mathbf{Y}) = \mathbf{\Lambda}(\mathbf{\Pi}(\mathbf{Y})) = \mathbf{\Lambda}(\mathbf{Z})$ for all $\mathbf{Y} \in \mathcal{Y}$, and (MMWD) gives

$$\begin{aligned}\dot{\mathbf{Z}} &= \frac{d}{dt}[\mathbf{Y} - (1/d)\operatorname{tr}[\mathbf{Y}]\mathbf{I}] = \dot{\mathbf{Y}} - (1/d)\operatorname{tr}[\dot{\mathbf{Y}}]\mathbf{I} \\ &= \mathbf{V}(\mathbf{\Lambda}(\mathbf{Y})) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{\Lambda}(\mathbf{Y}))]\mathbf{I} \\ &= \mathbf{V}(\mathbf{\Lambda}(\mathbf{Z})) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{\Lambda}(\mathbf{Z}))]\mathbf{I}\end{aligned} \tag{14}$$

so we can rewrite (MMWD) in terms of $\mathbf{Z}$ as

$$\dot{\mathbf{Z}} = \mathbf{V}(\mathbf{X}) - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{X})]\mathbf{I} \qquad \mathbf{X} = \mathbf{\Lambda}(\mathbf{Z}) \qquad (\text{MMWD}_{\mathcal{Z}})$$

Since $\operatorname{tr}[\dot{\mathbf{Z}}] = \operatorname{tr}[\mathbf{V}(\mathbf{X})] - (1/d)\operatorname{tr}[\mathbf{V}(\mathbf{X})]\operatorname{tr}\mathbf{I} = 0$, it follows that any traceless initial condition of ($\text{MMWD}_{\mathcal{Z}}$) will remain traceless – and hence remain in $\mathcal{Z}$ for all $n \geq 0$. We thus conclude that ($\text{MMWD}_{\mathcal{Z}}$) is a well-posed dynamical system on $\mathcal{Z}$ with induced flow $\zeta\colon \mathbb{R} \times \mathcal{Z} \to \mathcal{Z}$. The following proposition states that the ($\text{MMWD}_{\mathcal{Z}}$) dynamics are volume-preserving, and this key property will be used to construct the invariant measure μ.

Proposition 1. *Let $\mathcal{W} \subseteq \mathcal{Z}$ be an open set of initial conditions of* ($\text{MMWD}_{\mathcal{Z}}$)*, and let $\mathcal{W}_t = \zeta_t(\mathcal{W})$, $t \geq 0$, denote the evolution of $\mathcal{W}$ under the flow $\zeta\colon \mathbb{R}\times\mathcal{Z} \to \mathcal{Z}$ of* ($\text{MMWD}_{\mathcal{Z}}$)*. Then,* $\operatorname{vol}(\mathcal{W}_t) = \operatorname{vol}(\mathcal{W})$ *for all $t \geq 0$, where the form* $\operatorname{vol}(\cdot)$ *stands for the Lebesgue measure on $\mathcal{Z}$.*

The proof of Proposition 1 is based on the fact that each player's payoff function is individually linear in the player's own density matrix, so the individual gradient fields $\mathbf{V}_i$ do not depend on $\mathbf{X}_i$. The full proof lies at the end of the section.

Continuing with the proof of Theorem 1, by descending to the quotient space $\mathcal{Z}$, the map $\mathbf{\Lambda}$ factors through $\mathcal{Z}$ as $\mathbf{\Lambda}_0 : \mathcal{Z} \to \mathcal{X}$ with $\mathbf{\Lambda} = \mathbf{\Lambda}_0 \circ \mathbf{\Pi}$. Now, let λ denote the Lebesgue measure on $\mathcal{Z}$, and let μ denote the pushforward of λ to $\mathcal{X}$ via $\mathbf{\Lambda}_0$, i.e., $\mu(A) = \lambda(\mathbf{\Lambda}_0^{-1}(A))$ for all Borel sets $A \subseteq \mathcal{X}$. Finally, by Proposition 1 and since ζ conjugates χ in the sense that $\chi_t \circ \mathbf{\Lambda}_0 = \mathbf{\Lambda}_0 \circ \zeta_t$ for all $t \geq 0$, we get that $\mu(\chi_t(A)) = \mu(A)$ for all Borel sets A in $\operatorname{ri}\mathcal{X}$ and $t \geq 0$.

Step 2: We will now show that $\mathbf{X}^*$ cannot belong to $\operatorname{ri}\mathcal{X}$. For the sake of contradiction, suppose $\mathbf{X}^* \in \operatorname{ri}\mathcal{X}$ is consistently asymptotically stable, and let $\mathcal{U}$ be a sufficiently small compact neighborhood of $\mathbf{X}^*$, as per Lemma 1. Then, it holds $\lim_{t\to\infty}\sup_{\mathbf{X}\in\mathcal{U}}\|\chi_t(\mathbf{X}) - \mathbf{X}^*\| = 0$, which implies that:

$$\lim_{t\to\infty}\mu(\chi_t(\mathcal{U})) = \mu(\{\mathbf{X}^*\}) = 0 < \mu(\mathcal{U}). \tag{15}$$

However, by *Step 1* of the proof, we have that $\lim_{t\to\infty}\mu(\chi_t(\mathcal{U})) = \mu(\mathcal{U})$, since $\mu(\chi_t(\mathcal{U})) = \mu(\mathcal{U})$ for all $t \geq 0$. Combining it with (15), we arrive at a contradiction.

Step 3: Suppose, finally, that $\mathbf{X}^*$ is consistently asymptotically stable with $\operatorname{rank}(\mathbf{X}^*) > 1$. Note that if $\mathbf{X}^*$ is full-rank, it cannot be consistently asymptotically stable, as shown at the *Step 2* of the proof. Define the set:

$$\mathcal{A}_{\mathbf{X}^*} := \{\mathbf{X} \in \mathcal{X} : \ker(\mathbf{X}) \geq \ker(\mathbf{X}^*)\} \tag{16}$$

which is convex. Hence, considering the restriction of the dynamics on $\mathcal{A}_{\mathbf{X}^*}$, *Step 2* of the proof shows that a point in $\operatorname{ri}\mathcal{A}_{\mathbf{X}^*}$ cannot be consistently asymptotically stable on the induced topology. So, it remains to show that $\mathbf{X}^*$ cannot be consistently asymptotically stable.

For the sake of contradiction, suppose it is. Then, according to Definition 1, there exist $\mathcal{U}$ of $\mathbf{X}^*$ in $\mathcal{X}$ such that $\lim_{t\to\infty}\chi_t(\mathbf{X}) = \mathbf{X}^*$ for all $\mathbf{X} \in \mathcal{U} \cap \mathcal{X}_{\mathbf{X}^*}$. But, since $\mathcal{U}$ is a neighborhood of $\mathbf{X}^*$ in $\mathcal{X}$, we have that

$$(\mathcal{U} \cap \mathcal{X}_{\mathbf{X}^*}) \cap \mathcal{A}_{\mathbf{X}^*} = \mathcal{U} \cap \{\mathbf{X}' \in \mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\} \tag{17}$$

is a neighborhood of $\mathbf{X}^*$, and, thus, the restriction of $\mathcal{U}\cap\{\mathbf{X}' \in \mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\}$ on $\mathcal{A}_{\mathbf{X}^*}$ is a neighborhood of $\mathbf{X}^*$ in $\mathcal{A}_{\mathbf{X}^*}$. Therefore, by our previous argument on the induced dynamics on $\mathcal{A}_{\mathbf{X}^*}$, there exists $\mathbf{X}_0 \in \mathcal{U}\cap\{\mathbf{X}' \in \mathcal{X} : \ker(\mathbf{X}') = \ker(\mathbf{X}^*)\}$ such that $\chi_t(\mathbf{X}_0) \not\to \mathbf{X}^*$, as $t \to \infty$. This contradicts our assumption, so the proof is complete. ■

The second part of Theorem 1 leverages an energy argument in the spirit of Lyapunov's direct method. In the spirit of the analysis of the classical EW algorithm [16], [25], a natural choice of energy function is

$$E(t) = \operatorname{tr}[\mathbf{X}^* \log \mathbf{X}^*] + \log\operatorname{tr}[\exp(\mathbf{Y}(t))] - \operatorname{tr}[\mathbf{X}^*\mathbf{Y}(t)] \tag{18}$$

which is in turn linked to (VS) via Lemma 2 below:

Lemma 2. *Let $\mathbf{X}(t)$, $t \geq 0$, be a trajectory of play under* (MMWD)*, and let $E(t)$ be defined as per* (18)*. Then:*

$$\dot{E}(t) = \operatorname{tr}[\mathbf{V}(\mathbf{X}(t))\,(\mathbf{X}(t) - \mathbf{X}^*)] \tag{19}$$

By a direct calculation, it can be shown that (*i*) $E(t) \geq 0$ with equality if and only if $\mathbf{X}^* = \mathbf{\Lambda}(\mathbf{Y}(t))$; and (*ii*) $E(t) \to 0$ if and only if $\mathbf{\Lambda}(\mathbf{Y}(t)) \to \mathbf{X}^*$ as $t \to \infty$. Thus, putting everything together, we conclude that $\lim_{t\to\infty}E(t)$ exists; then, by a trapping argument, it can be shown that there exists a sequence of times $t_n \to \infty$ such that $\mathbf{X}(t_n) \to \mathbf{X}^*$, which allows us to conclude that $E(t) \to 0$ and ultimately yields our claim. Specifically, we have:

Proof of Theorem 1(b). Suppose that $\mathbf{X}^*$ is variationally stable, i.e., $\operatorname{tr}[\mathbf{V}(\mathbf{X})(\mathbf{X}-\mathbf{X}^*)] < 0$, for all $\mathbf{X} \in \mathcal{U}\backslash\{\mathbf{X}^*\}$, where $\mathcal{U}$ is a neighborhood of $\mathbf{X}^*$. Then, by Lyapunov's direct method for the energy function $E(t)$ defined in (18), we get that $\mathbf{X}^*$ is stable, thus, there exists neighborhood $\mathcal{U}'$ of $\mathbf{X}^*$, such that $\mathbf{X}(t) \in \mathcal{U}$ for all $t \geq 0$, if $\mathbf{X}(0) \in \mathcal{U}'$. This means that if $\mathbf{X}(0) \in \mathcal{U}'$, then $\operatorname{tr}[\mathbf{V}(\mathbf{X}(t))(\mathbf{X}(t) - \mathbf{X}^*)] < 0$ for all $t \geq 0$. In what follows, we will show that $\mathbf{X}(t) \to \mathbf{X}^*$.

Since $\mathbf{X}(t) \in \mathcal{U}$ for all $t \geq 0$, and using Lemma 2 along with (VS), we obtain that $\dot{E}(t) = \operatorname{tr}[\mathbf{V}(\mathbf{X}(t))(\mathbf{X}(t)-\mathbf{X}^*)] < 0$. Hence, $E(t)$ is decreasing if $\mathbf{X}(0) \in \mathcal{U}'$.

Now, we will show that $\mathbf{X}^*$ is an ω-limit of the flow, i.e., there exists a sequence $\{t_n\}_{n\in\mathbb{N}}$ such that $\mathbf{X}(t_n) \to \mathbf{X}^*$ as $n \to \infty$. For the sake of contradiction, suppose such a sequence does not exist. Then, $\|\mathbf{X}(t) - \mathbf{X}^*\|$ is bounded away from zero, which implies that there exists $c > 0$ such that

$$\mathrm{tr}[\mathbf{V}(\mathbf{X}(t))(\mathbf{X}(t) - \mathbf{X}^*)] < -c \tag{20}$$

Integrating over time, we get $E(t) - E(0) < -ct$, which implies that $E(t) \to -\infty$ as $t \to \infty$. This contradicts the non-negativity property of the energy function, thus, we conclude that there exists a sequence $\{t_n\}_{n\in\mathbb{N}}$ such that $\mathbf{X}(t_n) \to \mathbf{X}^*$ as $n \to \infty$. Hence, we get that $E(t_n) \to 0$ as $n \to \infty$. Therefore, since $E(t)$ is decreasing and $E(t_n) \to 0$, we conclude that $E(t) \to 0$ as $t \to \infty$. Thus, we conclude that $\mathbf{X}(t) \to \mathbf{X}^*$, i.e., $\mathbf{X}^*$ is consistently asymptotically stable. ■

Finally, we present the proofs of the intermediate results, required for our analysis.

Proof of Lemma 1. First of all, since $\mathbf{X}^* \in \mathrm{ri}\,\mathcal{X}$, we have that $\ker(\mathbf{X}^*) = \{0\}$, which implies that $\mathcal{X}_{\mathbf{X}^*} = \mathrm{ri}\,\mathcal{X}$. Hence, for any open set O in $\mathrm{ri}\,\mathcal{X}$, it holds that $O \cap \mathcal{X}_{\mathbf{X}^*} = O$. Now, let $\mathcal{U}_0$ be the basin of attraction of $\mathbf{X}^*$, according to the definition of the consistent asymptotic stability in Definition 1, and let $\mathcal{U} \subseteq \mathcal{U}_0$ be a compact neighborhood of $\mathbf{X}^*$. Suppose, for the sake of contradiction, that $\sup_{\mathbf{X}\in\mathcal{U}}\|\chi_t(\mathbf{X}) - \mathbf{X}^*\| \not\to 0$. This implies that there exists $\varepsilon > 0$, a sequence $\{t_n\}_{n\in\mathbb{N}}$ with $t_n \to \infty$ as $n \to \infty$, and $\mathbf{X}_n \in \mathcal{U}$ for $n \in \mathbb{N}$, such that

$$\|\chi_{t_n}(\mathbf{X}_n) - \mathbf{X}^*\| \geq \varepsilon \quad \text{for all } n \in \mathbb{N} \tag{21}$$

Since $\mathcal{U}$ is compact, we may assume (by taking subsequences, if necessary) that $\mathbf{X}_n$ converges to some limit point $\mathbf{X}_\infty \in \mathcal{U}$. Now, since $\mathbf{X}^*$ is Lyapunov stable, there exists a neighborhood $\mathcal{U}'$ of $\mathbf{X}^*$, such that the trajectory $\chi_t(\mathbf{X})$ remains within $\varepsilon/2$-distance of $\mathbf{X}^*$, if $\mathbf{X} \in \mathcal{U}'$, or,

$$\|\chi_t(\mathbf{X}) - \mathbf{X}^*\| < \varepsilon/2 \quad \text{for all } t \geq 0 \text{ and } \mathbf{X} \in \mathcal{U}' \tag{22}$$

Define the hitting time $\tau := \inf\{t \geq 0 : \chi_t(\mathbf{X}_\infty) \in \mathcal{U}'\}$ as the first time that the trajectory enters $\mathcal{U}'$ when starting from $\mathbf{X}_\infty$. It is easy to see that $\tau < \infty$, since $\mathbf{X}_\infty \in \mathcal{U} \subseteq \mathcal{U}_0$, and, thus, $\|\chi_t(\mathbf{X}_\infty) - \mathbf{X}^*\| \to 0$ as $t \to \infty$ by the asymptotic stability of $\mathbf{X}^*$. By continuity of χ, we readily get that there exists a neighborhood $\mathcal{D}$ of $\mathbf{X}_\infty$ such that $\chi_\tau(\mathcal{D}) \subseteq \mathcal{U}''$, where $\mathcal{U}''$ is a neighborhood of $\mathbf{X}^*$ with $\|\chi_t(\mathbf{X}) - \mathbf{X}^*\| < \varepsilon$ for all $t \geq 0$ and $\mathbf{X} \in \mathcal{U}''$. Since $\mathbf{X}_n \to \mathbf{X}_\infty$, we conclude that $\mathbf{X}_n \in \mathcal{D}$ for all sufficiently large n, which, in turn, implies that $\chi_\tau(\mathbf{X}_n) \in \mathcal{U}''$. Moreover, by definition of the sequence $\{t_n\}_{n\in\mathbb{N}}$, we have that $t_n \to \infty$, which gives that $t_n > \tau$ for n sufficiently large, since $\tau < \infty$, as argued. Thus, for n large, we have $\|\chi_{t_n}(\mathbf{X}_n) - \mathbf{X}^*\| < \varepsilon$, which contradicts (21). ■

Proof of Proposition 1. The key observation here is that the dynamics (MMWD$_{\mathcal{Z}}$) are *incompressible* – that is, *divergence-free*. Indeed, since $\mathbf{V}_i$ does not depend on $\mathbf{X}_i$ (it is not possible to suppress player indices here), we readily have $\nabla_{\mathbf{Z}_i^\top}\mathbf{V}_i(\mathbf{\Lambda}(\mathbf{Y})) = 0$ for all $i \in \mathcal{N}$. This immediately implies that the field

$$\mathbf{W}_i(\mathbf{Z}) = \mathbf{V}_i(\mathbf{\Lambda}(\mathbf{Z})) - (1/d_i)\,\mathrm{tr}[\mathbf{V}_i(\mathbf{\Lambda}(\mathbf{Z}))]\mathbf{I} \tag{23}$$

has $\mathrm{div}_{\mathbf{Z}}(\mathbf{W}(\mathbf{Z})) = 0$, so (MMWD$_{\mathcal{Z}}$) is incompressible. Our claim then follows from Liouville's formula [41]. ■

Proof of Lemma 2. Differentiating the energy function $E(t)$ with respect to t, and using (8) we obtain:

$$\begin{aligned}\frac{d}{dt}E(t) &= \frac{1}{\mathrm{tr}[\exp(\mathbf{Y}(t))]}\frac{d}{dt}\mathrm{tr}[\exp(\mathbf{Y}(t))] - \mathrm{tr}[\mathbf{X}^*\dot{\mathbf{Y}}(t)] \\ &= \int_0^1 \mathrm{tr}\left[\mathbf{X}^{1-s}\mathbf{V}(\mathbf{X}(t))\mathbf{X}^s\right] ds - \mathrm{tr}[\mathbf{X}^*\mathbf{V}(\mathbf{X}(t))] \\ &= \mathrm{tr}[\mathbf{V}(\mathbf{X}(t))\,(\mathbf{X}(t) - \mathbf{X}^*)] \qquad (24)\end{aligned}$$

and, the proof is complete. ■

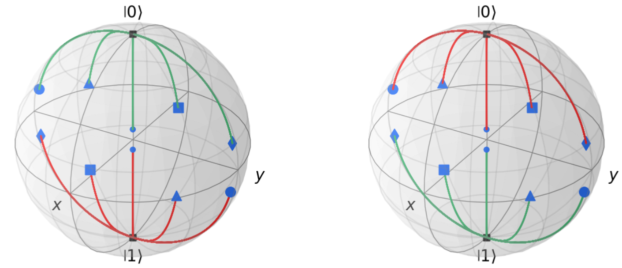

Fig. 2: Convergence of trajectories to variationally stable equilibria in a 2-player quantum anti-coordination game, visualized in Bloch spheres. The green trajectories correspond to player 1, while the red ones to player 2. The blue markers indicate the initial points of the trajectories. In each subfigure separately, the trajectories starting with markers of the same shape correspond to the the trajectories of the two players obtained under the same execution of the dynamics.

V. Simulation Setup

We consider a 2-player symmetric quantum game, which is obtained as the quantum analog of the 2-player finite symmetric game with 2 actions, $\{\alpha_1, \alpha_2\}$ and $\{\beta_1, \beta_2\}$ and payoff matrix $P = \begin{pmatrix}1 & 2\\ 2 & 1\end{pmatrix}$, same for both players. The action profiles (α_1, β_2), (α_2, β_1) are strict Nash equilibria of the finite game. The trajectories are visualized in Bloch spheres [42]. The payoff information of the quantum game is encoded in the Hermitian matrix $W = \mathrm{diag}(1, 2, 2, 1)$, as per (5), which is the same for both players. The green lines correspond the trajectories of player 1, while the red ones to the trajectories of player 2. The blue marker points correspond to the initial points of the trajectories, and in each subfigure of Fig. 2 separately, the trajectories starting with markers of the same shape correspond to the trajectories of the two players obtained under the same execution of the dynamics (10 different initializations of the dynamics in total, 5 in each subfigure). In the left subfigure, we see that the trajectories of the first and second player converge to the density matrices $\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}$ and $\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$, accordingly, i.e., the top and bottom points of the sphere, while in the right subfigure, we see the opposite. This is happening because the initial conditions of the dynamics in the left subfigure lie in the basin of attraction of the one variationally stable equilibrium, while in the right subfigure lie in the basin of the other.

VI. Concluding remarks

When quantum computing models are deployed in a multi-agent context – from autonomous vehicles to quantum GANs – the players' interaction landscape changes dramatically relative to classical interactions. The study of game-theoretic learning in this quantum setting is still in its infancy, so it is not clear at this stage what can be expected by quantum players with bounded rationality. In this regard, the study of the (MMWD) provides the following important insights: the geometric structure of quantum state space leads to an inflation of "learning traps" (stationary states) that have no classical counterpart; nonetheless, the only states that can be stable and attracting under (MMWD) are the game's pure quantum equilibria. Solidifying our understanding of the limits of quantum game-theoretic learning is a particularly fruitful research direction with potentially far-reaching implications for the deployment of multi-agent quantum computing systems.

References

[1] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, August 2018.

[2] F. Arute, K. Arya, R. Babbush, D. Bacon *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, 2019.

[3] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen *et al.*, "Quantum computational advantage using photons," *Science*, vol. 370, no. 6523, pp. 1460–1463, December 2020.

[4] S. Aaronson, "Shadow tomography of quantum states," *SIAM Journal on Computing*, vol. 49, no. 5, January 2020.

[5] P.-L. Dallaire-Demers and N. Killoran, "Quantum generative adversarial networks," *Physical Review A*, vol. 98, no. 1, p. 012324, 2018.

[6] S. Chakrabarti, Y. Huang, T. Li, S. Feizi *et al.*, "Quantum Wasserstein generative adversarial networks," in *NeurIPS '19: Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 2019.

[7] S. Lloyd and C. Weedbrook, "Quantum generative adversarial learning," *Physical Review Letters*, vol. 121, no. 4, July 2018.

[8] D. A. Meyer, "Quantum strategies," *Physical Review Letters*, vol. 82, no. 5, pp. 1052–1055, February 1999.

[9] J. Eisert, M. Wilkens, and M. Lewenstein, "Quantum games and quantum strategies," *Physical Review Letters*, vol. 83, October 1999.

[10] J. Bostanci and J. Watrous, "Quantum game theory and the complexity of approximating quantum Nash equilibria," *Quantum*, vol. 6, 2022.

[11] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou, "The complexity of computing a Nash equilibrium," in *STOC '06: Proceedings of the 38th annual ACM symposium on the Theory of Computing*, 2006.

[12] ——, "The complexity of computing a Nash equilibrium," *Communications of the ACM*, vol. 52, no. 2, pp. 89–97, 2009.

[13] K. Tsuda, G. Rätsch, and M. K. Warmuth, "Matrix exponentiated gradient updates for on-line Bregman projection," *Journal of Machine Learning Research*, vol. 6, pp. 995–1018, 2005.

[14] V. G. Vovk, "Aggregating strategies," in *COLT '90: Proceedings of the 3rd Workshop on Computational Learning Theory*, 1990, pp. 371–383.

[15] N. Littlestone and M. K. Warmuth, "The weighted majority algorithm," *Information and Computation*, vol. 108, no. 2, pp. 212–261, 1994.

[16] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. E. Schapire, "Gambling in a rigged casino: The adversarial multi-armed bandit problem," in *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, 1995.

[17] R. Jain and J. Watrous, "Parallel approximation of non-interactive zero-sum quantum games," in *CCC '09: Proceedings of the 2009 IEEE International Conference on Computational Complexity*, 2009.

[18] S. Aaronson, X. Chen, E. Hazan, S. Kale *et al.*, "Online learning of quantum states," in *NeurIPS '18: Proceedings of the 32nd International Conference of Neural Information Processing Systems*, 2018.

[19] R. Jain, G. Piliouras, and R. Sim, "Matrix multiplicative weights updates in quantum games: Conservation law & recurrence," in *NeurIPS '22: Proceedings of the 36th International Conference on Neural Information Processing Systems*, 2022.

[20] J. van Apeldoorn and A. Gily'en, "Quantum algorithms for zero-sum games," *arXiv: Quantum Physics*, 2019.

[21] T. Li, S. Chakrabarti, and X. Wu, "Sublinear quantum algorithms for training linear and kernel-based classifiers," in *International Conference on Machine Learning*, 2019.

[22] T. Li, C. Wang, S. Chakrabarti, and X. Wu, "Sublinear classical and quantum algorithms for general matrix games," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 10, May 2021.

[23] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. The MIT Press, 2003.

[24] J. Hofbauer and K. Sigmund, "Evolutionary game dynamics," *Bulletin of the American Mathematical Society*, vol. 40, no. 4, July 2003.

[25] P. Mertikopoulos and W. H. Sandholm, "Learning in games via reinforcement and regularization," *Mathematics of Operations Research*, vol. 41, no. 4, pp. 1297–1324, November 2016.

[26] L. Flokas, E. V. Vlatakis-Gkaragkounis, T. Lianeas, P. Mertikopoulos *et al.*, "No-regret learning and mixed Nash equilibria: They do not mix," in *NeurIPS '20: Proceedings of the 34th International Conference on Neural Information Processing Systems*, 2020.

[27] G. Gutoski and J. Watrous, "Toward a general theory of quantum games," in *STOC '07: Proceedings of the 39th annual ACM symposium on the Theory of Computing*, 2007.

[28] I. Chuang and M. Nielsen, *Quantum Computation and Quantum Information*, 2nd ed. Cambridge University Press, 2010.

[29] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the USA*, October 1952.

[30] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, ser. Springer Series in Operations Research. Springer, 2003.

[31] G. Scutari, F. Facchinei, D. P. Palomar, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory in multiuser communication systems," *IEEE Signal Process. Mag.*, May 2010.

[32] P. Mertikopoulos and Z. Zhou, "Learning in games with continuous action sets and unknown payoff functions," *Mathematical Programming*, vol. 173, no. 1-2, pp. 465–507, January 2019.

[33] P. Mertikopoulos and A. L. Moustakas, "Learning in an uncertain world: MIMO covariance matrix optimization with imperfect feedback," *IEEE Trans. Signal Process.*, vol. 64, no. 1, January 2016.

[34] R. M. Wilcox, "Exponential operators and parameter differentiation in quantum physics," *Journal of Mathematical Physics*, vol. 8, 1967.

[35] P. D. Taylor and L. B. Jonker, "Evolutionary stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no. 1-2, 1978.

[36] J. W. Weibull, *Evolutionary Game Theory*. MIT Press, 1995.

[37] W. H. Sandholm, *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press, 2010.

[38] S. Hart and A. Mas-Colell, "Uncoupled dynamics do not lead to Nash equilibrium," *American Economic Review*, vol. 93, no. 5, pp. 1830–1836, 2003.

[39] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, pp. 15–18, November 1973.

[40] J. Maynard Smith, *Evolution and the Theory of Games*. Cambridge: Cambridge University Press, 1982.

[41] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, 2nd ed., ser. Graduate Texts in Mathematics. New York, NY: Springer, 1989.

[42] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.