Matrix Multiplicative Weights Updates in Quantum Zero-Sum Games: Conservation Laws & Recurrence

Anonymous Author(s)
Affiliation
Address
email

Abstract

Recent advances in quantum computing and in particular, the introduction of quantum GANs, have led to increased interest in quantum zero-sum game theory, extending the scope of learning algorithms for classical games into the quantum realm. In this paper, we focus on learning in quantum zero-sum games under *Matrix Multiplicative Weights Update* (a generalization of the multiplicative weights update method) and its continuous analogue, *Quantum Replicator Dynamics*. When each player selects their state according to quantum replicator dynamics, we show that the system exhibits conservation laws in a quantum-information theoretic sense. Moreover, we show that the system exhibits Poincaré recurrence, meaning that almost all orbits return arbitrarily close to their initial conditions infinitely often. Our analysis generalizes previous results in the case of classical games [44, 38].

1 Introduction

The Nash equilibrium has always been a central concept in non-cooperative game theory. While the existence of Nash equilibria is well known in finite games where mixed strategies are allowed [40], it is less clear how to efficiently compute these equilibria themselves. Indeed, much work revolves around finding methods to solve for Nash equilibria in different contexts, as well as analyzing the complexity of solving for the Nash [30, 15].

Despite the dominance of equilibrium computation in classical theory, recent works have begun to move towards attempting to understand the nature of learning in games [20, 47]. The family of regret-minimizing algorithms known as Follow-The-Regularized-Leader (FTRL) have been studied extensively, both for discrete time [7, 13, 41, 45] and continuous time [38, 18, 44, 43] dynamics. These algorithms have found many applications in the domain of machine learning - one such example is that of Generative Adversarial Networks (GANs) [22], which have been successfully modeled using zero-sum game theory. This rich connection has led to a stronger understanding of adversarial learning and improved practical performance of GANs in various settings [18, 29, 33].

On the other side of the coin, quantum computing is a field which has garnered much interest in the machine learning community. Many quantum machine learning algorithms have been proposed, extending standard classical algorithms. One such example is quantum GANs [14, 36, 12]. While such works describe the architecture of quantum GANs, we still do not fully understand the behavior of learning algorithms in the quantum setting. In this paper, we provide an initial analysis of learning in zero-sum quantum games. We do so by studying the learning behaviour of the ubiquitous Multiplicative Weights Update (MWU) update rule [34, 19, 5], a specific instance of the FTRL framework, when applied to such games.

There have been several important results studying MWU and its continuous counterpart, *replicator dynamics*. [7] showed that classical MWU does not converge in day-to-day behaviour to the mixed Nash equilibrium when applied to two player zero-sum games. [38] showed that the orbits of players

Submitted to 36th Conference on Neural Information Processing Systems (NeurIPS 2022). Do not distribute.

in zero-sum games exhibit Poincaré recurrence when using replicator dynamics. However, quantum systems can oftentimes behave in radically different ways to their classical counterparts. Indeed, allowing for quantum strategies can create situations wherein players can gain greater payoffs than if they were playing using only classical strategies (see e.g. the Mermin–Peres magic square game [37, 42]). As such, our extension of the classical results to the quantum realm has to be treated with care.

We focus on a matrix generalization of MWU known as *Matrix Multiplicative Weights Update* (MMWU), which changes the paradigm of standard MWU from cost vectors to cost matrices, and from probability vectors to density matrices. This generalization has been independently discovered and studied by [50] as Matrix Exponentiated Gradient Updates and [5] as the Matrix Multiplicative Weights algorithm. Applications of the MMWU algorithm include solving semi-definite programs (SDPs) [6] and obtaining bounds on the sample complexity for learning problems in quantum computing [27]. From a game theoretic standpoint, [25] analyzed the MMWU algorithm for quantum zero-sum games, proving time-average convergence to an approximate Nash equilibrium. We extend this result by studying the learning behaviour of MMWU in quantum zero-sum games, which could lay the groundwork for developing novel algorithms that achieve convergence to the Nash in day-to-day behaviour in quantum GANs, similar to the case of classical GANs.

Contributions. In this paper, we first study the dynamics of MMWU in quantum zero-sum games, utilizing tools from information theory and classical game theory (Section 3). We provide bounds on the rate of change of the total quantum relative entropy in the system. We then study the continuous counterpart of MMWU, which we call quantum replicator dynamics (Section 4). Analyzing this continuous dynamic allows for various discretizations in order to obtain different discrete-time algorithms for specific purposes. Drawing from multiple recent results in this direction, we show that the total quantum relative entropy in the system is a constant of motion. Furthermore, the dynamics exhibited by both players' trajectories do not converge to equilibrium (in the day-to-day sense) nor oscillate periodically, but rather exhibit a weaker form of periodicity known as *Poincaré recurrence*. Our proof of this result departs from the standard classical method, representing our main technical novelty. Finally, we present several simulations which corroborate our theoretical results (Section 5).

2 Preliminaries and Definitions

2.1 Quantum Theory

Basic concepts. In this paper, we only require a few basic concepts about quantum theory, thus it is not necessary for the reader to have any prior familiarity with the concepts used.

First, we refer to a quantum *register* as a collection of qubits representing a message that is transferred from one party to another. We associate a vector space $\mathcal{H} = \mathbb{C}^n$ with any quantum register. This intuitively represents the maximum number of distinct classical states that can be stored in the register without error. The *state* of a quantum register is represented by a *density matrix*, which is an $n \times n$ positive semi-definite matrix with trace 1. We will use $D(\mathcal{H})$ to denote the set of all density matrices associated with a register that is described by $\mathcal{H}$. One can naturally view such density matrices as linear operators acting on $\mathcal{H}$.

When two registers with associated spaces $\mathcal{A} = \mathbb{C}^n$ and $\mathcal{B} = \mathbb{C}^m$ are considered as a joint register, the associated space is the tensor product $\mathcal{A} \otimes \mathcal{B} = \mathbb{C}^{nm}$. If the two registers are independently prepared in states described by ρ and σ, then the joint state is described by the $nm \times nm$ density matrix $\rho \otimes \sigma$.

Next, for a given vector space $\mathcal{H} = \mathbb{C}^n$, we define $\mathrm{L}(\mathcal{H})$ as the set of all $n \times n$ complex matrices. Furthermore, we denote the subset of $\mathrm{L}(\mathcal{H})$ given by *Hermitian* matrices as $\mathrm{Herm}(\mathcal{H})$. A Hermitian matrix A satisfies the equality $A = A^\dagger$, where $A^\dagger$ denotes the *adjoint* (or conjugate transpose) of matrix A. Subsequently, we define $\mathrm{Pos}(\mathcal{H})$ as the subset of $\mathrm{Herm}(\mathcal{H})$ which consists of all positive semi-definite $n \times n$ matrices.

Finally, the *Hilbert-Schmidt inner product* on $\mathrm{L}(\mathcal{H})$ is defined as $\langle A, B \rangle = \mathrm{Tr}(A^\dagger B)$ for all $A, B \in \mathrm{L}(\mathcal{H})$. Note that $\langle A, B \rangle$ is a real number for any Hermitian matrices A and B, and is also additionally non-negative if A and B are positive semi-definite.

Measurements and Observables. In the context of game theory, we are also interested in the concepts of quantum *measurements* and *observables*. An observable is simply a property of the quantum system which is measurable. The measurement of a register having associated vector space $\mathcal{H} = \mathbb{C}^n$ is a collection of linear operators $\{P_i : 1 \leq i \leq k\} \subset \mathrm{Pos}(\mathcal{H})$ which satisfies $\sum_{i=1}^{k} P_i = \mathbb{1}_{\mathcal{H}}$, where $\mathbb{1}_{\mathcal{H}}$ is the identity matrix on $\mathcal{H}$. If the register corresponding to $\mathcal{H}$ is in a state defined by density matrix ρ and the measurement described by P_i is performed, each outcome i will be observed with probability $\langle P_i, \rho \rangle$. An important note is that two mixed states with the same density matrix are indistinguishable from each other by any measurement.

Additional notation and definitions. A linear mapping of the form $\Phi : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ is called a super-operator. The adjoint super-operator to Φ is $\Phi^{\dagger} : \mathrm{L}(\mathcal{A}) \to \mathrm{L}(\mathcal{B})$ and is uniquely determined by the condition:

$$\langle A, \Phi(B) \rangle = \langle \Phi^{\dagger}(A), B \rangle$$

A super-operator $\Phi : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ is said to be positive if $\Phi(P)$ is positive semi-definite for every choice of positive semi-definite operator $P \in \mathrm{Pos}(\mathcal{B})$. In addition, $\Phi^{\dagger}$ is positive if and only if Φ is positive. There is a one-to-one and linear correspondence between the collection of operators of the form $R \in \mathrm{L}(\mathcal{A} \otimes \mathcal{B})$ and the collection of super-operators $\Phi : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ defined above. Specifically, for each super-operator Φ we can define an operator R as follows:

$$R = \sum_{1 \leq i,j \leq m} \Phi(E_{i,j}) \otimes E_{i,j} \tag{1}$$

where $E_{i,j}$ is the matrix with a 1 in entry (i,j) and 0 elsewhere. Conversely, given an operator $R \in \mathrm{L}(\mathcal{A} \otimes \mathcal{B})$, we can define:

$$\Phi(B) = \mathrm{Tr}_{\mathcal{B}}(R(\mathbb{1}_{\mathcal{A}} \otimes B^{\top})) \tag{2}$$

This correspondence is linear and one can translate back and forth between the two as needed for an given application. R is assumed to be positive semi-definite, so the corresponding super-operator Φ is also positive.

2.2 Quantum Game Theory

A quantum system which can be manipulated by any number of agents, and where the utility of the moves is well defined, quantified and ordered can be conceived as a quantum game. In particular, a 2-player zero-sum quantum game sees players Alice and Bob each sending a quantum state to a referee, who then performs a measurement on these two states to determine their payoffs. Let $\mathcal{A} = \mathbb{C}^n$ and $\mathcal{B} = \mathbb{C}^m$ be the vector spaces that correspond to the state that Alice and Bob send to the referee.

In the classical context, a finite normal-form game consists of a set of players $\mathcal{N} = \{1, \ldots, N\}$, where each player i selects action s_i from a finite set of actions $\mathcal{S}_i = \{1, \ldots, k_i\}$. For the purposes of this paper, we borrow that notation to clarify our setting: we restrict the quantum games we consider to 2-player games where both players have the same action space, meaning that $\mathcal{N} = \{1, 2\}$ and $\mathcal{S} = \{1, \ldots, k\}$. In order to determine the payoffs of the players' actions, the referee performs a joint measurement where Alice's and Bob's states are viewed as a single register. Thus, the referee's measurement can be described by a collection

$$\{R_a : a \in \mathcal{S}\} \subset \mathrm{Pos}(\mathcal{A} \otimes \mathcal{B}) \tag{3}$$

which satisfies the condition $\sum_{a=1}^{k} R_a = \mathbb{1}_{\mathcal{A} \otimes \mathcal{B}}$. We associate each possible measurement outcome a with a payoff for each player. Since we are looking at zero-sum games, if the payoff that Alice receives from the referee is $v(a)$, then Bob's corresponding payoff will be $-v(a)$. Henceforth, we refer to the states sent by Alice and Bob to the referee as ρ and σ respectively. For a given choice of ρ and σ, Alice's expected payoff is:

$$u(\rho, \sigma) = \sum_{a=1}^{k} v(a) \langle R_a, \rho \otimes \sigma \rangle = \langle R, \rho \otimes \sigma \rangle \tag{4}$$

where $R = \sum_{a=1}^{k} v(a) R_a$. Likewise, Bob's corresponding expected payoff is $-u(\rho, \sigma) = -\langle R, \rho \otimes \sigma \rangle$. R is referred to throughout the rest of the paper as a *payoff observable*. A necessary and sufficient condition for matrix R acting on $\mathcal{A} \otimes \mathcal{B}$ to be obtained from some real valued payoff function v is that R is Hermitian. From Equation 2, we can equivalently define the expected payoff of Alice as $u(\rho, \sigma) = \langle \rho, \Phi(\sigma) \rangle$. We will use the latter formulation for the remainder of the paper.

A key notion of equilibrium in classical game theory is the Nash equilibrium. In the quantum setting, we define the pair of quantum states (ρ^*, σ^*) as a Nash equilibrium of R if

$$u(\rho^*, \sigma^*) \geq u(\rho, \sigma^*) \quad \text{and} \quad u(\rho^*, \sigma^*) \geq u(\rho^*, \sigma) \tag{5}$$

for all ρ, σ. That is, neither Alice nor Bob would prefer to unilaterally deviate from playing ρ^* and σ^* respectively.

Finally, we define the notion of a 'fully mixed' Nash equilibrium. An equilibrium (ρ^*, σ^*) is fully mixed if the payoff of a player for deviating from the equilibrium to any other strategy remains exactly equal to their equilibrium payoff. A standard example in classical game theory is Rock-Paper-Scissors, with equilibrium $[1/3, 1/3, 1/3]$ giving payoff 0 to both players.

2.3 Information Theory

We also introduce several information theoretic concepts which will be referenced throughout the paper.

Shannon entropy. The Shannon entropy of a random variable X where each strategy x is obtained with probability $p(x)$ is given by $H(X) = -\sum_x p(x) \log p(x)$. This intuitively is a measure of randomness or uncertainty in the system. A natural generalization of the Shannon entropy to a quantum context is the von Neumann entropy. For a quantum mechanical system defined by density matrix ρ, the von Neumann entropy is given by $S(\rho) = -\text{Tr}(\rho \log \rho)$.

Bregman divergence. We are also interested in the notion of Bregman divergence, which measures the distance between two points. Let $x^* \in \mathcal{X}$ be a Nash equilibrium and let $x \in \mathcal{X}$ be an arbitrary strategy profile. Also, let F be a continuously-differentiable, strictly convex function. The Bregman divergence from x^* to x is given by $D_F(x^* \| x) = F(x^*) - F(x) - \langle \nabla F(x), x^* - x \rangle$.

In the context of quantum games, the Bregman divergence can also be defined using matrix notation. We define the quantum relative entropy between two quantum states using the von Neumann entropy $S(\rho)$. In particular, the quantum relative entropy between two quantum states ρ and σ is defined as $S(\rho \| \sigma) = \text{Tr}\left(\rho(\log \rho - \log \sigma)\right)$.

2.4 Dynamical Systems

Flows. Consider a differential equation $\dot{p} = f(p)$ on a topological space $\mathcal{P}$. The existence and uniqueness theorem for ordinary differential equations guarantees that we can write the unique solution to the differential equation as a continuous map $\phi : \mathcal{P} \times \mathcal{H} \to \mathcal{P}$. This is referred to as the *flow of the differential equation* such that for any point $p \in \mathcal{P}$, $\phi(p, -)$ defines a function of time corresponding to the trajectory of p. Conversely, fixing a time t provides a map $\phi^t \equiv \phi(-, t) : \mathcal{P} \to \mathcal{P}$. In Section 4, we introduce the notion of quantum replicator dynamics, which are Lipschitz continuous differential equations. Hence, a unique flow ϕ of these replicator dynamics exists.

Liouville's theorem. Liouville's theorem can be applied to any system of ordinary differential equations with a continuously differentiable vector field ξ on an open domain $\mathcal{Y} \in \mathbb{R}^d$. The divergence of ξ at $y \in \mathcal{Y}$ is the trace of the Jacobian at y: div $\xi(y) = \sum_{i=1}^d \frac{\partial \xi_i}{\partial y_i}(y)$. Because the divergence is continuous, it is integrable on Lebesgue measurable subsets of $\mathcal{Y}$. Given any such set A, define the image of A under flow ϕ at time t as $A(t) = \{\phi(a, t) : a \in A\}$. $A(t)$ is measurable and of volume vol$[A(t)] = \int_{A(t)} d\mu$. Liouville's formula states that the time derivative of the volume vol$[A(t)]$ exists and links it to the divergence of ξ:

$$\frac{d}{dt}\left[\text{vol } A(t)\right] = \int_{A(t)} \text{div}(\xi) d\mu \tag{6}$$

If div $\xi(y)$ is null at any $y \in \mathcal{Y}$, then the volume is conserved. Since div ξ is continuous, the converse statement is also true - if the volume is conserved on any open set, div $\xi(y)$ has to be null at any point $y \in \mathcal{Y}$.

Algorithm 1: Parallel Approximation of Equilibrium Point

Let $\mu = \epsilon/8$ and let $N = \lceil 64 \log{(nm)}/\epsilon^2 \rceil$.
Initialize: $A_0 = \mathbb{1}_{\mathcal{A}}$, $\rho_0 = A_0/\mathrm{Tr}(A_0)$, $B_0 = \mathbb{1}_{\mathcal{B}}$, and $\sigma_0 = B_0/\mathrm{Tr}(B_0)$.
for $j = 1 \ldots N$ **do**
  $A_j = \exp\left(\mu \sum_{i=0}^{j-1} \Phi(\sigma_i)\right)$
  $\rho_j = A_j/\mathrm{Tr}(A_j)$
  $B_j = \exp\left(-\mu \sum_{i=0}^{j-1} \Phi^*(\rho_i)\right)$
  $\sigma_j = B_j/\mathrm{Tr}(B_j)$
end for

Diffeomorphisms of flows. A function $\mathbf{f}$ between two topological spaces is called a diffeomorphism if i) $\mathbf{f}$ is a bijection, ii) $\mathbf{f}$ is continuously differentiable, iii) $\mathbf{f}$ has a continuously differentiable inverse. Two flows $\Phi^t : A \to A$ and $\Psi^t : B \to B$ are diffeomorphic if there exists a diffeomorphism $\mathbf{f} : A \to B$ such that for each $x \in A$ and $t \in \mathbb{R}$, $\mathbf{f}(\Phi^t(p)) = \Psi^t(\mathbf{f}(p))$. For the purpose of our analysis, the replicator dynamics defined in Equations 17 are translated via a diffeomorphism from the interior of $\mathcal{P}$ to a space $\mathcal{C} = \Pi_{i \in V} \mathbb{R}^{n-1}$, which allows us to show certain desirable properties.

Poincarè recurrence. The concept of Poincarè recurrence arises from Henri Poincarè's celebrated 1890 work regarding the three body problem [46]. He proved that if a dynamical system preserves volume and always remains bounded in its orbits, almost all trajectories return arbitrarily close to their initial position, and do so infinitely often.

Theorem 2.1 (Poincarè recurrence)**.** *If a flow preserves volume and has only bounded orbits then for each open set there exist orbits that intersect the set infinitely often.*

3 MMWU in Quantum Zero-Sum Games

In [25], the MMWU algorithm for zero-sum games is shown to exhibit time-average convergence to an approximate Nash equilibrium in two-player quantum zero-sum games. The MMWU algorithm is shown in Algorithm 1.

Note that here we focus specifically on two-player games, and utilize the expected payoffs defined via super-operators $\Phi : \mathrm{L}(\mathcal{A}) \to \mathrm{L}(\mathcal{B})$ as seen in Equation 2. Moreover, μ is the step-size in the quantum algorithm.

In this section, we examine closely the update steps for each player in the MMWU algorithm (Algorithm 1) and analyze the limiting behaviour of the total quantum relative entropy in the system. First, we introduce two useful facts which will aid in the analysis.

Fact 3.1 (Golden-Thompson inequality [21, 49]). Let A, B be Hermitian matrices. Then

$$\mathrm{Tr}\exp(A + B) \leq \mathrm{Tr}\exp(A)\exp(B) \tag{7}$$

Fact 3.2. Let $0 \leq A \leq \mathbb{1}$ be a PSD matrix and δ be a real number. Then,

$$\exp(\delta A) \leq \mathbb{1} + \delta \exp(\delta) A \tag{8}$$

Next, we put forward two corollaries which will help us prove Theorem 3.5. In particular, these corollaries present important relationships between the values of A and B from one time step to the next and the total quantum relative entropy within the system. We first define $\Delta S(\rho^* \| \rho_j) = S(\rho^* \| \rho_j) - S(\rho^* \| \rho_{j-1})$ and $\Delta S(\sigma^* \| \sigma_j) = S(\sigma^* \| \sigma_j) - S(\sigma^* \| \sigma_{j-1})$.

Corollary 3.3. *The sum of quantum relative entropies in a quantum zero-sum game with fully-mixed Nash equilibrium is given by:*

$$\Delta S(\rho^* \| \rho_j) + \Delta S(\sigma^* \| \sigma_j) = \log \frac{\mathrm{Tr} A_j}{\mathrm{Tr} A_{j-1}} + \log \frac{\mathrm{Tr} B_j}{\mathrm{Tr} B_{j-1}} \tag{9}$$

Corollary 3.4. *The following trace inequalities hold for PSD matrices A and B updated with MMWU:*

$$\Delta S(\rho^* \| \rho_j) + \Delta S(\sigma^* \| \sigma_j) \geq \mu \exp(-\mu) \mathrm{Tr}\left(\rho_j \Phi(\sigma_{j-1})\right) - \mu \exp(\mu) \mathrm{Tr}\left(\rho_{j-1} \Phi(\sigma_j)\right) \tag{10}$$

$$\Delta S(\rho^* \| \rho_j) + \Delta S(\sigma^* \| \sigma_j) \leq \mu \exp(\mu) \mathrm{Tr}\left(\rho_{j-1} \Phi(\sigma_{j-1})\right) - \mu \exp(-\mu) \mathrm{Tr}\left(\rho_{j-1} \Phi(\sigma_{j-1})\right) \tag{11}$$

The proof of Corollary 3.4 relies on Facts 3.1 and 3.2, as well as the equality shown in Corollary 3.3.

In many practical scenarios, one would use decreasing step-sizes when running MMWU. As such, we utilize Corollary 3.4 and take the limit as step-size μ goes to 0 in order to show the following theorem:

Theorem 3.5. *The sum of quantum relative entropies between the player's strategies and a fully mixed Nash equilibrium in a two-player zero-sum quantum game tends to zero when step-size $\mu \to 0$. Specifically,*

$$\lim_{\mu \to 0} \frac{1}{\mu} \left(\Delta S(\rho^* \| \rho_j) + \Delta S(\sigma^* \| \sigma_j)\right) = 0 \tag{12}$$

The proof of Theorem 3.5, along with Corollaries 3.3 and 3.4 can be found in Appendix C.

Theorem 3.5 further motivates an investigation into the continuous time variant of MMWU, which we call quantum/matrix replicator dynamics. In particular, we show that in the continuous case, the sum of quantum relative entropies is invariant. We introduce and study quantum replicator dynamics in detail in Section 4.

4 Replicator Dynamics in Quantum Zero-Sum Games

We have seen that in the case of discrete dynamics (MMWU), as the step-size becomes infinitesimal, the total quantum relative entropy in the system remains invariant. A natural question would then be: does the same result hold in continuous time? In order to explore this question in greater detail, we first need to define the continuous analogue of MMWU, the *quantum replicator dynamics* [47], in the quantum setting.

We start by rewriting the MMWU update steps, but now defined over a continuous time interval $[0, t]$:

$$A(t) = \int_0^t \Phi(\sigma(\tau)) d\tau \tag{13}$$

$$\rho(t) = \exp(A(t))/\mathrm{Tr}(\exp(A(t))) \tag{14}$$

$$B(t) = -\int_0^t \Phi^\dagger(\rho(\tau)) d\tau \tag{15}$$

$$\sigma(t) = \exp(B(t))/\mathrm{Tr}(\exp(B(t))) \tag{16}$$

Note here that we shift the exponential terms from the definition of $A(t)$ and $B(t)$ to the corresponding $\rho(t)$ and $\sigma(t)$ terms. This will help simplify some of the proof techniques later on in the paper. Furthermore, in the rest of the paper we will typically drop from the notation the explicit dependence on t to ease with the notational burden.

It is important to note the following observation, which will be helpful in our later analysis.

Observation 4.1. As $\mu \to 0$, the discrete-time trajectories ρ_j and σ_j defined in Algorithm 1 are equal to the continuous-time trajectories $\rho(t)$ and $\sigma(t)$ defined in Equations 14 and 16.

We now define the *quantum replicator dynamics* as:

$$d\rho/dt = \frac{d}{dt}\left(\frac{\exp(A)}{\mathrm{Tr}(\exp(A))}\right), \quad d\sigma/dt = \frac{d}{dt}\left(\frac{\exp(B)}{\mathrm{Tr}(\exp(B))}\right) \tag{17}$$

It is worth noting that typically, one can write the replicator equations in a form that describes the relative utility that one agent obtains as compared to the average utility overall. However, in the quantum case this is not possible in general, since it relies on the assumption that $\int_0^t \Phi(\sigma(\tau))d\tau$ and $\Phi(\sigma(t))$, and respectively $\int_0^t \Phi^\dagger(\rho(\tau))d\tau$ and $\Phi^\dagger(\rho(t))$ commute.

As a consequence of Observation 4.1, we can also conclude that the dynamical system described by the replicator dynamics defined in Equations 17 is a limit case of the dynamical system described by the MMWU algorithm as $\mu \to 0$.

We are now able to state the main theorem for quantum relative entropy in quantum replicator dynamics.

Theorem 4.2. *When applying matrix/quantum replicator dynamics in a quantum zero-sum game with a fully-mixed Nash equilibrium (ρ^*, σ^*), the sum of quantum relative entropies between the fully-mixed Nash equilibrium and the state of the system $(\rho(t), \sigma(t))$ is invariant on every system trajectory, i.e.:*

$$\frac{d\big(S(\rho^*\|\rho(t)) + S(\sigma^*\|\sigma(t))\big)}{dt} = 0 \tag{18}$$

The proof of this theorem utilizes Observation 4.1 to show that at the limit $\mu \to 0$, the rate of change of the quantum relative entropy is exactly the same whether we are in the discrete or continuous setting. We then apply Theorem 3.5 to complete the proof.

4.1 Poincaré Recurrence in Quantum Zero-Sum Games

Now that we have described analytical results surrounding the day-to-day behaviour of quantum replicator dynamics, we seek to understand the *dynamics* of the trajectories. After all, invariance of quantum relative entropy does not fully describe how the system moves over time. We show that for any two-player zero-sum quantum game updated with replicator dynamics, the system exhibits *Poincaré recurrence*, insofar as the game is zero-sum and has a fully-mixed Nash equilibria. As introduced in Section 2, the notion of Poincaré recurrence is a weaker version of periodicity. To be precise, for almost all initial conditions $\rho_0 \in \mathcal{P}$, the replicator dynamics return arbitrarily close to ρ_0 infinitely often.

Theorem 4.3. *The quantum replicator dynamics given in Equations 17 are Poincaré recurrent in any two player zero-sum game which has a fully-mixed Nash equilibrium.*

The proof of this main theorem involves carefully piecing together several auxiliary results, which we will describe in the rest of the section. Furthermore, we stress that due to the non-commutative nature of quantum systems, the standard (classical) approach of differentiating the discrete-time dynamics in the primal space of probability distributions does not apply directly unless we have the highly unlikely situation where $\int_0^t \Phi(\sigma(\tau))d\tau$ and $\Phi(\sigma(t))$ (resp. $\int_0^t \Phi^\dagger(\rho(\tau))d\tau$ and $\Phi^\dagger(\rho(t))$) commute. This problem with carrying over the standard approach is explicitly discussed in Appendix D.

For the proof in the quantum setting, we first define a *canonical transformation* on the space of the matrices $A(t)$ and $B(t)$, which will be crucial in proving the theorem.

Definition 4.4 (Canonical transformation)**.** We define the canonical transformation of $A'(t)$ and $B'(t)$ to be a mapping of $A(t)$ and $B(t)$ as defined by Equations 13 and 15. In particular, we define

$$\begin{aligned} A'(t) &= A(t) - (v^\dagger A(t) v)\mathbb{1} \\ B'(t) &= B(t) - (v^\dagger B(t) v)\mathbb{1} \end{aligned} \tag{19}$$

where v is a fixed vector defined as $v = [1, 0 \dots 0]^\top$, such that the values of $v^\dagger A(t) v$ and $v^\dagger B(t) v$ are real numbers corresponding to the $(1,1)$-th element of matrices $A(t)$ and $B(t)$ for all t. Notice that this creates matrices $A'(t)$ and $B'(t)$ which have 0 as the $(1,1)$-th entry.

Under the transformation in Definition 4.4, the vector fields $\dot{A}'(t) = F(A')$ and $\dot{B}'(t) = F(B')$ are given by:

$$\begin{aligned} \dot{A}'(t) &= \Phi(\sigma(t)) - (v^\dagger \Phi(\sigma(t)) v)\mathbb{1} \\ \dot{B}'(t) &= -\Phi^\dagger(\rho(t)) + (v^\dagger \Phi^\dagger(\rho(t)) v)\mathbb{1} \end{aligned} \tag{20}$$

where $\frac{d}{dt}\left(v^\dagger A(t) v\right)$ is given by $v^\dagger \frac{dA(t)}{dt} v$.

Moreover, the values of $\rho'(t)$ and $\sigma'(t)$ are defined as:

$$\begin{aligned} \rho'(t) &= \exp(A'(t))/\mathrm{Tr}(\exp(A'(t))) \\ \sigma'(t) &= \exp(B'(t))/\mathrm{Tr}(\exp(B'(t))) \end{aligned} \tag{21}$$

Proposition 4.5. *The dynamics of $\rho(t)$ and $\sigma(t)$ remain the same after undergoing the canonical transformation. Equivalently, $A'(t)$ and $A(t)$ (resp. $B'(t)$ and $B(t)$) admit the same strategy $\rho(t)$ (resp. $\sigma(t)$).*

Proposition 4.6. *The mappings $A'(t)$ and $\rho(t)$ (resp. $B'(t)$ and $\sigma(t)$) are diffeomorphic to one another.*

Proposition 4.6 will be of crucial importance to our proof technique, since we first prove recurrence for the system described by $A'(t)$ and $B'(t)$, then recover recurrence in $\rho(t)$ and $\sigma(t)$.

To show Poincaré recurrence of Equations 17, we first prove two key properties: (i) the flow of $\dot{A}'$ is volume preserving, meaning that the trace of the Jacobian of the respective vector fields $\dot{A}'(t) = F(A')$ and $\dot{B}'(t) = F(B')$ are zero, and (ii) A' and B' have bounded orbits from any interior initial condition. Then, Poincaré recurrence of A' and B' follows from Poincaré's recurrence theorem.

Volume Conservation. We introduce a lemma which shows that in two-player zero-sum quantum replicator dynamics, the canonical transformation produces a dynamical system which preserves volume.

Lemma 4.7. *For two-player zero-sum quantum replicator dynamics, the vector fields that arise as a result of the canonical transformation in Definition 4.4 are volume preserving.*

The proof of Lemma 4.7 follows from considering the flows of $A'(t)$ and $B'(t)$ and showing that the divergences of the vector fields defined in Equations 20 are equal to zero. A straightforward application of Liouville's theorem completes the proof.

Bounded Orbits. We now show that the transformed dynamical system always has bounded orbits when initialized on the interior of the space of probability density matrices.

Lemma 4.8. *For any finite initial points $A(0)$ and $B(0)$, the dynamics mapped to $A'(t)$ and $B'(t)$ via the transformation in Definition 4.4 have bounded orbits.*

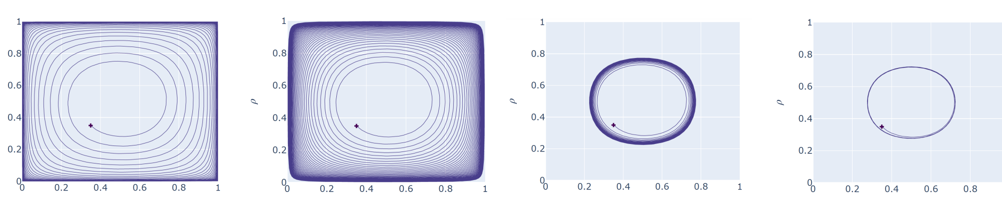
The proof of Lemma 4.8 relies on Theorem 4.2, and leverages the hermicity of the matrices involved. Moreover, we use the fact that the canonically transformed matrices $A'(t)$ and $B'(t)$ have zero as the $(1, 1)$-th element to bound the eigenvalues of $A'(t)$ and $B'(t)$ away from infinity.

Now we are ready to prove Theorem 4.3 using Lemmas 4.7 and 4.8.

Proof of Theorem 4.3. By Lemmas 4.7 and 4.8, as well as the Poincaré recurrence theorem introduced in Section 2, we immediately see that the system of replicator equations given by dA'/dt and dB'/dt are Poincaré recurrent since they are volume preserving and have bounded orbits. Since the flows of $A'(t)$ and $\rho(t)$ are diffeomorphic to one another (likewise for $B'(t)$ and $\sigma(t)$), $d\rho/dt$ and $d\sigma/dt$ are also Poincaré recurrent. This concludes the proof. □

All proofs of the results in this section are provided in Appendix D.

5 Experimental Results



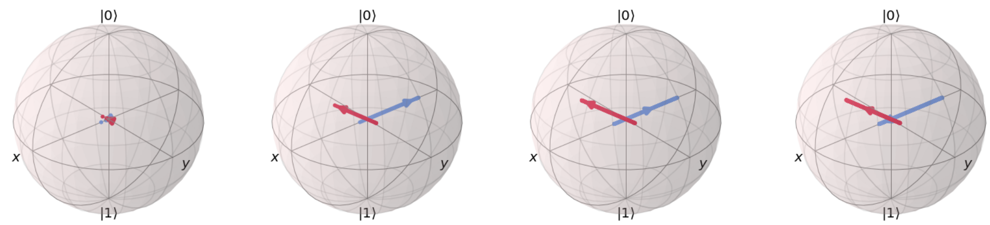
$\mu = \log(1 + \frac{1}{t^{1/4}})$ $\mu = \log(1 + \frac{1}{t^{1/3}})$ $\mu = \log(1 + \frac{1}{t^{1/2}})$ $\mu = \log(1 + \frac{1}{t^{2/3}})$

Figure 1: Eigenvalue trajectories for quantum Matching Pennies game with decreasing μ values.

To corroborate the theoretical results presented in prior sections, we performed relevant simulations of quantum games using both discrete MMWU and replicator dynamics. In the rest of this section, we standardize the use of quantum game matrices obtained via basis transform (described in more detail in Appendix E). This effectively allows us to transform classical games to the matrix setting.

First, we show the trajectories of the first eigenvalue of each player in a quantum Matching Pennies game, obtained using the discrete MMWU algorithm. We see that in accordance to Theorem 3.5, the rate of divergence of the trajectories from the uniform Nash goes to zero for cases with rapidly decreasing learning rate μ.

In the case of replicator dynamics, we present Bloch sphere representations of the trajectories in a quantum Matching Pennies game. The Bloch sphere is a unit 2-sphere representation of a qubit, and we utilize it to visualize the orbits of the replicator dynamics. In particular, the density matrix representing the strategy of each player at each time-step is given as a point within the sphere, and we plot the movement of these orbits over time. According to Theorems 4.2 and 4.3, we expect the trajectories of the replicator dynamics to stay on the interior of the Bloch sphere, since the surface of the sphere corresponds to the pure states of the system. We see from Figure 2 that over time, the system never reaches the boundary of the sphere, which experimentally agrees with our theory.



$t = 1$ $t = 50$ $t = 370$ $t = 1000$

Figure 2: Bloch sphere trajectories for quantum Matching Pennies game between Alice (blue) and Bob (red). The arrowheads represent the current state of each player at each time-step. Notice that over time, the orbits oscillate within the interior of the Bloch sphere.

6 Conclusion

In this paper, we studied the properties of Matrix Multiplicative Weights Update and its continuous analogue, quantum replicator dynamics, in the context of two-player zero-sum quantum games. First, we provide a formulation of quantum replicator dynamics which arises from MMWU. Then, we show that the total quantum relative entropy within such a system is a constant of motion. Finally, we show that in quantum replicator systems with interior Nash equilibria, the dynamics exhibit Poincaré recurrence. This work constitutes an initial step towards analyzing learning behaviour in games with quantum information. In the classical world, showing that conservation laws and recurrence holds has led to a better understanding of game dynamics in increasingly complex settings [48, 39, 43]. Similar work is now potentially possible in the quantum setting. Some other interesting directions for future work include:

- extending our results to multi-agent network generalizations of quantum zero-sum games,
- formalizing the analysis of quantum GANs using our results about learning in quantum zero-sum games, and
- understanding the quantum setting for different classes of games, e.g. potential games.

References

[1] S. Aaronson, X. Chen, E. Hazan, S. Kale, and A. Nayak. Online learning of quantum states. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(12):124019, 2019.

[2] L. Accardi and A. Boukas. Von neumann's minimax theorem for continuous quantum games. *Journal of Stochastic Analysis*, 1(2), Jun 2020.

[3] Z. Allen-Zhu and Y. Li. Follow the compressed leader: Faster online learning of eigenvectors and faster mmwu. In *International Conference on Machine Learning*, pages 116–125. PMLR, 2017.

[4] Z. Allen-Zhu, Z. Liao, and L. Orecchia. Spectral sparsification and regret minimization beyond matrix multiplicative updates. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 237–245, 2015.

[5] S. Arora, E. Hazan, and S. Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

[6] S. Arora and S. Kale. A combinatorial, primal-dual approach to semidefinite programs. *J. ACM*, 63(2), May 2016.

[7] J. P. Bailey and G. Piliouras. Multiplicative weights update in zero-sum games. In *Proceedings of the 2018 ACM Conference on Economics and Computation*, pages 321–338, 2018.

[8] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[9] R. Bondesan and M. Welling. The hintons in your neural network: a quantum field theory view of deep learning, 2021.

[10] J. Bostanci and J. Watrous. Quantum game theory and the complexity of approximating quantum nash equilibria. *arXiv preprint arXiv:2102.00512*, 2021.

[11] F. G. Brandão, A. Kalev, T. Li, C. Y.-Y. Lin, K. M. Svore, and X. Wu. Quantum sdp solvers: Large speed-ups, optimality, and applications to quantum learning. *arXiv preprint arXiv:1710.02581*, 2017.

[12] S. Chakrabarti, H. Yiming, T. Li, S. Feizi, and X. Wu. Quantum wasserstein generative adversarial networks. *Advances in Neural Information Processing Systems*, 32, 2019.

[13] Y. K. Cheung. Multiplicative weights updates with constant step-size in graphical constant-sum games. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[14] P.-L. Dallaire-Demers and N. Killoran. Quantum generative adversarial networks. *Physical Review A*, 98(1):012324, 2018.

[15] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

[16] D. Dong, C. Chen, H. Li, and T.-J. Tarn. Quantum reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 38(5):1207–1220, Oct 2008.

[17] V. Dunjko, J. M. Taylor, and H. J. Briegel. Advances in quantum reinforcement learning. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct 2017.

[18] L. Flokas, E.-V. Vlatakis-Gkaragkounis, and G. Piliouras. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games, 2019.

[19] Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

[20] D. Fudenberg and D. K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[21] S. Golden. Lower bounds for the helmholtz function. *Phys. Rev.*, 137:B1127–B1128, Feb 1965.

[22] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[23] E. Hazan, S. Kale, and S. Shalev-Shwartz. Near-optimal algorithms for online matrix prediction. In *Conference on Learning Theory*, pages 38–1. JMLR Workshop and Conference Proceedings, 2012.

[24] S. B. Hopkins, J. Li, and F. Zhang. Robust and heavy-tailed mean estimation made simple, via regret minimization. *arXiv preprint arXiv:2007.15839*, 2020.

[25] R. Jain and J. Watrous. Parallel approximation of non-interactive zero-sum quantum games. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 243–253. IEEE, 2009.

[26] J. R. Johansson, P. D. Nation, and F. Nori. Qutip: An open-source python framework for the dynamics of open quantum systems. *Computer Physics Communications*, 183(8):1760–1772, 2012.

[27] S. Kale. *Efficient algorithms using the multiplicative weights update method*. Princeton University, 2007.

[28] F. S. Khan, N. Solmeyer, R. Balu, and T. S. Humble. Quantum games: a review of the history, current state, and interpretation. *Quantum Information Processing*, 17(11), 10 2018.

[29] N. Kodali, J. Abernethy, J. Hays, and Z. Kira. On convergence and stability of gans. *arXiv preprint arXiv:1705.07215*, 2017.

[30] C. E. Lemke and J. T. Howson, Jr. Equilibrium points of bimatrix games. *Journal of the Society for Industrial and Applied Mathematics*, 12(2):413–423, 1964.

[31] T. Li, S. Chakrabarti, and X. Wu. Sublinear quantum algorithms for training linear and kernel-based classifiers. In *International Conference on Machine Learning*, pages 3815–3824. PMLR, 2019.

[32] T. Li, C. Wang, S. Chakrabarti, and X. Wu. Sublinear classical and quantum algorithms for general matrix games. *arXiv preprint arXiv:2012.06519*, 2020.

[33] T. Lin, C. Jin, and M. Jordan. On gradient descent ascent for nonconvex-concave minimax problems. In *International Conference on Machine Learning*, pages 6083–6093. PMLR, 2020.

[34] N. Littlestone and M. K. Warmuth. The weighted majority algorithm. *Information and computation*, 108(2):212–261, 1994.

[35] S. Lloyd, M. Mohseni, and P. Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv preprint arXiv:1307.0411*, 2013.

[36] S. Lloyd and C. Weedbrook. Quantum generative adversarial learning. *Physical review letters*, 121(4):040502, 2018.

[37] N. D. Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical review letters*, 65(27):3373, 1990.

[38] P. Mertikopoulos, C. H. Papadimitriou, and G. Piliouras. Cycles in adversarial regularized learning. In *SODA '18 - Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2703–2717, New Orleans, United States, Jan. 2018.

[39] S. G. Nagarajan, D. Balduzzi, and G. Piliouras. From chaos to order: Symmetry and conservation laws in game dynamics. In *International Conference on Machine Learning*, pages 7186–7196. PMLR, 2020.

[40] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[41] G. Palaiopanos, I. Panageas, and G. Piliouras. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos, 2017.

[42] A. Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3-4):107–108, 1990.

[43] J. Perolat, R. Munos, J.-B. Lespiau, S. Omidshafiei, M. Rowland, P. Ortega, N. Burch, T. Anthony, D. Balduzzi, B. D. Vylder, G. Piliouras, M. Lanctot, and K. Tuyls. From poincaré recurrence to convergence in imperfect information games: Finding equilibrium via regularization, 2020.

[44] G. Piliouras and J. Shamma. Optimization despite chaos: Convex relaxations to complex limit sets via poincaré recurrence. In *SODA*, 2014.

[45] G. Piliouras, R. Sim, and S. Skoulakis. Optimal no-regret learning in general games: Bounded regret with unbounded step-sizes via clairvoyant mwu. *arXiv preprint arXiv:2111.14737*, 2021.

[46] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270, 1890.

[47] W. H. Sandholm. *Population Games and Evolutionary Dynamics*. The MIT Press, 2010.

[48] S. Skoulakis, T. Fiez, R. Sim, G. Piliouras, and L. Ratliff. Evolutionary game theory squared: Evolving agents in endogenously evolving zero-sum games. *arXiv preprint arXiv:2012.08382*, 2020.

[49] C. J. Thompson. Inequality with applications in statistical mechanics. *Journal of Mathematical Physics*, 6(11):1812–1813, 1965.

[50] K. Tsuda, G. Rätsch, and M. K. Warmuth. Matrix exponentiated gradient updates for on-line learning and bregman projection. *Journal of Machine Learning Research*, 6(Jun):995–1018, 2005.

[51] J. van Apeldoorn and A. Gilyén. Quantum algorithms for zero-sum games. *arXiv preprint arXiv:1904.03180*, 2019.

[52] A. Youssry, C. Ferrie, and M. Tomamichel. Efficient online quantum state estimation using a matrix-exponentiated gradient method. *New Journal of Physics*, 21(3):033006, 2019.

[53] S. Zhang. Quantum strategic game theory. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 39–59, 2012.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes] The contributions and scope are clearly stated and derived directly from our results.
   (b) Did you describe the limitations of your work? [Yes] The limitations of our work are described in the introduction and in the main text.
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] See Section 6.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes] See Section 6.
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes] All assumptions are stated either in the main body or the appendix.
   (b) Did you include complete proofs of all theoretical results? [Yes] All proofs are provided in the main text or appendix.
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] The code is compiled in several formats and provided in the supplementary material.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [No] Our simulations do not require training.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No] Our simulations are deterministic.
   (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [Yes] Details of the resources and packages used are described in the supplementary material.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] We cite the code packages used.
   (b) Did you mention the license of the assets? [N/A]
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A] Our work does not utilize human-derived data.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A] Our work does not include crowdsourcing or research with human subjects.
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]