Appendix

This supplementary material contains an overview of additional related work in Appendix A, additional examples and definitions in Appendix B, proofs omitted from the paper for space considerations in Appendix C and D, and further experimental results and details in Appendix E.

A Additional Related Work

We now describe a broader class of related work. In particular, we focus on work regarding $a)$ the MMWU algorithm and its applications as well as $b)$ quantum algorithms in game theory.

MMWU Algorithm and Applications. MMWU as an algorithm has been widely studied in the past. For instance, MMWU has been applied to the problem of solving semidefinite programs in [11]. The central idea behind this line of work is the 'quantization' of standard SDP solvers using the matrix multiplicative weights method. Furthermore, MMWU has also been applied to the problem of learning a quantum state, also known as state tomography. [1, 58] have utilized MMWU as a framework for online learning in this setting.

In a broader sense, MMWU can also be considered a member of the class of Matrix Follow-The-Regularized-Leader (FTRL) algorithms. This paradigm has been studied in works such as [26, 4, 27], which admits a wider range of applications such as spectral sparsification of graphs. In addition, some more recent works such as [3] study extensions to FTRL within the regret minimization framework.

Quantum Computing. Quantum information theory has proved to be a useful tool in the fields of reinforcement learning and deep learning [9, 18, 19]. The basic results of classical game theory, such as von Neumann's minimax theorem and more have been shown to hold in the quantum context [31, 2], while recent work has shown analogous complexity and equilibrium results between classical and quantum games [59, 10].

There is also a rich literature towards solving and analyzing games using quantum algorithms. [55] derives a sublinear time quantum algorithm for solving two-player zero-sum games, while [35] derives a similar algorithm for general matrix games. Finally, quantum algorithms have proven to be useful in many machine learning problems such as classification [34] and GANs [39]. For a general overview of quantum algorithms in machine learning see [38] and [8].

B Additional Preliminaries

B.1 An Example: Quantum Matching Pennies

To provide a clearer picture of quantum games, we illustrate the differences between the classical and quantum versions of the famous Matching Pennies game. In classical game theory, Matching Pennies is defined by the payoff matrix $P = \left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$. In the terminology defined above, the payoff observable R would then simply be a diagonal 4×4 matrix with the elements of the classical payoff matrix on the diagonal:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Let us first consider the case where Alice and Bob play randomized mixed strategies in a classical fashion. Alice and Bob will then simultaneously select *vectors* which represent a probability distribution over the strategies *heads* or *tails*. For example, Alice could return the vector $x_A = [1/2, 1/2]$ and Bob could return the vector $x_B = [1/3, 2/3]$. The expected utility for each player is then determined using the payoff matrix:

$$u_A = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}^\top \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix} = 0, \quad u_B = \begin{bmatrix} 1/3 \\ 2/3 \end{bmatrix}^\top \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} = 0$$

For the quantum case, Alice and Bob will instead send density matrices that represent their probabilities of playing a given strategy to a referee. For example, Alice might send the matrix

$x_A = \begin{bmatrix} 1/2 & 1/2i \\ -1/2i & 1/2 \end{bmatrix}$ and Bob could return the matrix $x_B = \begin{bmatrix} 1/3 & 1/2i \\ -1/2i & 2/3 \end{bmatrix}$. The payoffs to each player are then calculated using the payoff observable R as defined in Equation 4.

B.2 Classical MWU Algorithm

In the paper we focus on the matrix version of MWU, with particular focus on the quantum interpretation of the positive semi-definite matrices used in Algorithm 1. However, many of the results we show are analogous to results shown in classical settings.

For two-player zero-sum games (i.e. $A = -B^\top$, where A and B are the payoff matrices for players a and b.), the classical MWU algorithm is shown in Algorithm 2 [22]:

Algorithm 2: MWU for Two-Player Zero-Sum Games

```
Initialize: $x^0_{i,a} = 1/n$ and $x^0_{i,b} = 1/n$ for all $i$ (both players, uniform)
for t=1...T do
    Players $a$ and $b$ choose action $i \in \mathcal{S}$ with probability $x^t_{i,a}$ and $x^t_{i,b}$ respectively.
    for each action $i$ do
        $x^t_{i,a} = \frac{x^{t-1}_{i,a}(1+\epsilon)^{(Ax^{t-1}_b)_i}}{\sum_{j\in\mathcal{S}} x^{t-1}_{j,a}(1+\epsilon)^{(Ax^{t-1}_b)_j}}$
        $x^t_{i,b} = \frac{x^{t-1}_{i,b}(1+\epsilon)^{(Bx^{t-1}_a)_i}}{\sum_{j\in\mathcal{S}} x^{t-1}_{j,b}(1+\epsilon)^{(Bx^{t-1}_a)_j}}$
    end for
end for
```

Furthermore $(Ax^{t-1}_b)_i$ and $(Bx^{t-1}_a)_i$ represent the expected utility obtained by players a and b respectively for playing strategy i.

C Proofs from Section 3

In this section we present proofs of results in Section 3 which were omitted from the main paper due to space considerations.

C.1 Proof of Corollary 3.3

Proof. We use the definition of quantum relative entropy. First, we consider the change in quantum relative entropy from timestep $j-1$ to j.

$$\begin{aligned} S(\rho^*\|\rho_j) - S(\rho^*\|\rho_{j-1}) &= \mathrm{Tr}\left(\rho^*(\log\rho_{j-1} - \log\rho_j)\right) \\ &= \mathrm{Tr}\left(\rho^*\left(\log\frac{A_{j-1}}{\mathrm{Tr}A_{j-1}} - \log\frac{A_j}{\mathrm{Tr}A_j}\right)\right) \\ &= \mathrm{Tr}\left(\rho^*\left(\log A_{j-1} - \log A_j + \log\mathrm{Tr}A_j - \log\mathrm{Tr}A_{j-1}\right)\right) \\ &= \mathrm{Tr}\left(\rho^*\left(-\mu\Phi(\sigma_{j-1}) + \log\mathrm{Tr}A_j - \log\mathrm{Tr}A_{j-1}\right)\right) \end{aligned}$$

Hence,

$$S(\rho^*\|\rho_j) - S(\rho^*\|\rho_{j-1}) = -\mu\mathrm{Tr}(\rho^*\Phi(\sigma^*)) + \mathrm{Tr}(\rho^*(\log\mathrm{Tr}A_j - \log\mathrm{Tr}A_{j-1})) \tag{22}$$

Similarly, we know that

$$S(\sigma^*\|\sigma_j) - S(\sigma^*\|\sigma_{j-1}) = -\mu\mathrm{Tr}(\sigma^*\Phi^\dagger(\rho^*)) + \mathrm{Tr}(\sigma^*(\log\mathrm{Tr}B_j - \log\mathrm{Tr}B_{j-1})) \tag{23}$$

Summing up equations 22 and 23:

$$\begin{aligned} &\Delta S(\rho^*\|\rho_j) + \Delta S(\sigma^*\|\sigma_j) \\ &= \mu\left[-\mathrm{Tr}(\rho^*\Phi(\sigma^*)) + \mathrm{Tr}(\sigma^*\Phi^\dagger(\rho^*))\right] + \log\frac{\mathrm{Tr}A_j}{\mathrm{Tr}A_{j-1}} + \log\frac{\mathrm{Tr}B_j}{\mathrm{Tr}B_{j-1}} \\ &= \log\frac{\mathrm{Tr}A_j}{\mathrm{Tr}A_{j-1}} + \log\frac{\mathrm{Tr}B_j}{\mathrm{Tr}B_{j-1}} \end{aligned}$$

□

C.2 Proof of Corollary 3.4

Proof. We first show the upper bound on the sum $\log\left(\frac{\mathrm{Tr}A_j}{\mathrm{Tr}A_{j-1}}\right)+\log\left(\frac{\mathrm{Tr}B_j}{\mathrm{Tr}B_{j-1}}\right)$. Note that the definition of A_j is:

$$A_j = \exp\left(\mu\sum_{i=0}^{j-1}\Phi(\sigma_i)\right)$$

We can alternatively write:

$$A_j = \exp\left(\log(A_{j-1})+\mu\Phi(\sigma_{j-1})\right) \tag{24}$$

Hence, by taking the trace of A_{j-1} and applying Facts 3.1 and 3.2:

$$\begin{aligned}
\mathrm{Tr}(A_{j-1}) &\le \mathrm{Tr}\left[A_j\exp(-\mu\Phi(\sigma_{j-1}))\right] && \text{(Fact 3.1)}\\
&\le \mathrm{Tr}\left[A_j(\mathbb{1}-\mu\exp(-\mu)\Phi(\sigma_{j-1}))\right] && \text{(Fact 3.2)}\\
&= (\mathrm{Tr}(A_j))(1-\mu\exp(-\mu)\mathrm{Tr}\left(\rho_j\Phi(\sigma_{j-1})\right)\\
&\le (\mathrm{Tr}(A_j))\exp(-\mu\exp(-\mu)\mathrm{Tr}\left(\rho_j\Phi(\sigma_{j-1})\right)). && (1+x\le\exp(x))
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathrm{Tr}(B_{j-1}) &\le \mathrm{Tr}\left[B_j\exp(\mu\Phi^\dagger(\rho_{j-1}))\right]\\
&\le \mathrm{Tr}\left[B_j(\mathbb{1}+\mu\exp(\mu)\Phi^\dagger(\rho_{j-1}))\right]\\
&= (\mathrm{Tr}(B_j))(1+\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_j)\right))\\
&\le (\mathrm{Tr}(B_j))\exp(\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_j)\right)).
\end{aligned}$$

By rearranging and taking matrix logarithm on both sides, we obtain:

$$\log\left(\frac{\mathrm{Tr}A_j}{\mathrm{Tr}A_{j-1}}\right)\ge\mu\exp(-\mu)\mathrm{Tr}\left(\rho_j\Phi(\sigma_{j-1})\right)\qquad\log\left(\frac{\mathrm{Tr}B_j}{\mathrm{Tr}B_{j-1}}\right)\ge-\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_j)\right) \tag{25}$$

The inequality in Equation 10 follows by summing up the two inequalities obtained above and applying Corollary 3.3. Likewise, we can perform similar calculations on $\mathrm{Tr}(A_j)$ and $\mathrm{Tr}(B_j)$ to obtain the statement of Equation 11.

$$\begin{aligned}
\mathrm{Tr}A_j &\le \mathrm{Tr}\left[A_{j-1}\exp(\mu\Phi(\sigma_{j-1}))\right]\\
&\le \mathrm{Tr}\left[A_{j-1}(\mathbb{1}+\mu\exp(\mu)\Phi(\sigma_{j-1}))\right]\\
&= (\mathrm{Tr}A_{j-1})(1+\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right))\\
&\le (\mathrm{Tr}A_{j-1})\exp(\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right))
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathrm{Tr}B_j &\le \mathrm{Tr}\left[B_{j-1}\exp(-\mu\Phi^\dagger(\rho_{j-1}))\right]\\
&\le \mathrm{Tr}\left[B_{j-1}(\mathbb{1}-\mu\exp(-\mu)\Phi^\dagger(\rho_{j-1}))\right]\\
&= (\mathrm{Tr}B_{j-1})(1-\mu\exp(-\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right))\\
&\le (\mathrm{Tr}B_{j-1})\exp(-\mu\exp(-\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right))
\end{aligned}$$

Again rearranging and taking matrix logarithm on both sides:

$$\log\frac{\mathrm{Tr}A_j}{\mathrm{Tr}A_{j-1}}\le\mu\exp(\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right)\quad\log\frac{\mathrm{Tr}B_j}{\mathrm{Tr}B_{j-1}}\le-\mu\exp(-\mu)\mathrm{Tr}\left(\rho_{j-1}\Phi(\sigma_{j-1})\right) \tag{26}$$

Thus, an upper bound on $\Delta S(\rho^*\|\rho_j)+\Delta S(\sigma^*\|\sigma_j)$ is:

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

□

C.3 Proof of Theorem 3.5

In order to prove Theorem 3.5, we utilize Corollary 3.4 and take the limit of the inequalities as $\mu \to 0$.

Proof. Consider the MMWU update from time $j-1$ to j. As $\mu \to 0$, ρ_j and σ_j do not change more than $O(\mu)$. Indeed, since all payoffs in the game are bounded in the MMWU update from time $j-1$ to j, all entries in the numerators increase by at most $\exp(O(\mu)) = 1 + O(\mu)$. Likewise, the denominator is at least as large, but also upper bounded by the previous value of the denominator multiplied by $(1 + O(\mu))$. Hence, every entry in the outputs ρ_j and σ_j are at most $O(\mu)$ from ρ_{j-1} and σ_{j-1}. Moreover, we have that $Tr(\rho_j \Phi(\sigma_{j-1}))$, $Tr(\rho_{j-1}\Phi(\sigma_j))$ and $Tr(\rho_{j-1}\Phi(\sigma_{j-1}))$ are all within $O(\mu)$ of each other. Using Taylor expansion, both the upper bounds and lower bounds given in Equations 10,11 are of the order of $O(\mu^2)$ and the theorem follows.

□

D Proofs and Additional Results from Section 4

In this section, we focus on the key results shown in Section 4. In particular, we first present some connections between the quantum and classical replicator dynamics. Then, we show the comprehensive proof of Theorem 4.3, which relies on Lemmas 4.7 and 4.8. In addition, the proofs of all other results can also be found in this section.

D.1 Proof of Observation 4.1

Proof. It suffices to show that at the limit, the integrals $A(t)$ and $B(t)$ defined in 13 and 15 can be written in the discrete summation form as seen in Algorithm 1. First we write the Riemann sum for $A(t)$ by taking j infinitesimal intervals in time interval $[0, t]$, each of width μ:

$$A(t) = \int_0^t \Phi(\sigma(\tau))d\tau = \sum_{i=0}^{j-1} \left(\frac{t}{j}\right) \Phi(\sigma_i)$$

However $t = \mu j$, so the above can be written as $\sum_{i=0}^{j-1} \mu\Phi(\sigma_i)$. Taking limit as $\mu \to 0$, it is clear to see that the matrix exponent of the continuous-time trajectory $A(t)$ is equal to its discrete-time counterpart A_j. Hence, the trajectories $\rho(t)$ and ρ_j are equivalent at the limit. A similar argument holds for $\sigma(t)$ and σ_j at the limit. □

D.2 Connections between Quantum and Classical Replicator Dynamics

In this section we show that in the special case where commutativity holds, the quantum replicator dynamics presented in Section 4 are equivalent to that of the classical replicator equations. Recall that we used the integral form of the replicator dynamics in the main text. However, typically in the classical case one can write the replicator dynamics in a form that represents the utilities obtained by each player:

$$\dot{x} = x\left(Ay - \left(x^\top Ay\right)\vec{\mathbb{1}}\right)$$

where x and y are n-dimensional probability vectors representing the strategies of each player, A is the payoff matrix and $\vec{\mathbb{1}}$ is the n-dimensional all-one vector. In the quantum setting, we require a few additional results and assumptions in order to write the Equations 17 in the classical form. The following lemma arrives directly as a result of the series definition of matrix exponential.

Lemma D.1. *If matrices P and $\frac{dP}{dt}$ commute then*

$$\frac{d(\exp(P))}{dt} = \frac{dP}{dt} \cdot \exp(P) = \exp(P) \cdot \frac{dP}{dt}$$

Proof.

$$\begin{aligned}\frac{d}{dt}(\exp(P)) &= \frac{d}{dt}\left[\mathbb{1} + P + \frac{1}{2}P^2 + \frac{1}{3!}P^3 + \cdots\right] \\ &= \frac{dP}{dt}\left[\mathbb{1} + P + \frac{1}{2}P^2 + \cdots\right] \\ &= \frac{dP}{dt}\cdot\exp(P) = \exp(P)\cdot\frac{dP}{dt} \qquad \text{(by commutativity)}\end{aligned}$$

□

By leveraging Lemma D.1 and assuming that commutativity holds, we can state the following lemma:

Lemma D.2. *If $\int_0^t \Phi(\sigma(\tau))d\tau$ & $\Phi(\sigma(t))$, and $\int_0^t \Phi^\dagger(\rho(\tau))d\tau$ & $\Phi^\dagger(\rho(t))$ commute, then the replicator system when applied to a two-player zero-sum quantum game is equivalent to:*

$$d\rho/dt = \rho[\Phi(\sigma) - \mathrm{Tr}(\rho\Phi(\sigma))\mathbb{1}] \tag{27}$$

$$d\sigma/dt = \sigma[-\Phi^\dagger(\rho) + \mathrm{Tr}(\sigma\Phi^\dagger(\rho))\mathbb{1}] \tag{28}$$

Proof. By definition we have that $\rho = \exp(A)/\mathrm{Tr}(\exp(A))$ where $A = \int_0^t \Phi(\sigma(\tau))d\tau$. By applying Lemma D.1 we have that if $\int_0^t \Phi(\sigma(\tau))d\tau$ and $\Phi(\sigma(t))$ commute then $d\exp(A)/dt = \exp(A)\Phi(\sigma) = \Phi(\sigma)\exp(A)$.

$$\begin{aligned}d\rho/dt &= \frac{(d\exp(A)/dt)\mathrm{Tr}(\exp(A)) - \exp(A)d(\mathrm{Tr}(\exp(A)))/dt}{(\mathrm{Tr}\exp(A))^2} \\ &= \frac{(d\exp(A)/dt)\mathrm{Tr}(\exp(A)) - \exp(A)\mathrm{Tr}(d\exp(A)/dt)}{(\mathrm{Tr}\exp(A))^2} \\ &= \frac{\exp(A)\Phi(\sigma)\mathrm{Tr}(\exp(A)) - \exp(A)\mathrm{Tr}(\exp(A)\Phi(\sigma))}{(\mathrm{Tr}\exp(A))^2} \\ &= \rho[\Phi(\sigma) - \mathrm{Tr}(\rho\Phi(\sigma))\mathbb{1}]\end{aligned}$$

The proof in the case of $d\sigma/dt$ is similar. □

We can see that Equations 27 and 28 take a familiar form - indeed, the quantum replicator dynamics model the difference in utility obtained by each player compared to the average utility in much the same way as the classical case. In general, commutativity holds only in the case when all matrices $\Phi^\dagger(\rho(t))$ and $\Phi(\sigma(t))$ are diagonal, which is precisely the classical setting. The assumption that commutativity holds in the quantum setting is a very strong one, and in general commutativity does not hold once we enter the realm of quantum information.

Reminiscent of a similar result in classical evolutionary game theory, one can write a proof for the invariance of total entropy directly using the formulation of replicator dynamics. However given the commutativity assumption, this 'proof' only holds in the case where we are writing classical probability vectors and game payoff matrices in the quantum notation.

Proof of Theorem 4.2 (Classical case). We will prove that $d\mathrm{Tr}\big(\rho^*(\log\rho^* - \log\rho(t)) + \sigma^*(\log\sigma^* - \log\sigma(t))\big)/dt = 0$. To do that we will focus on the terms related to the first agent, i.e., $d\mathrm{Tr}\big(\rho^*(\log\rho^* - \log\rho(t))\big)/dt = -d\mathrm{Tr}\big(\rho^*\log\rho(t))\big)/dt$.

$$\begin{aligned}
d\mathrm{Tr}\big(\rho^* \log \rho(t))\big)/dt &= \mathrm{Tr}\big(\rho^* d(\log \rho(t))/dt\big) \\
&= \mathrm{Tr}\big(\rho^* \rho(t)^{-1} d\rho(t)/dt\big) \quad \text{(by Lemma } D.1^1\text{)} \\
&= \mathrm{Tr}\big(\rho^* \rho^{-1} \rho[\Phi(\sigma) - \mathrm{Tr}(\rho\Phi(\sigma))I]\big) \quad \text{(by Lemma D.2)} \\
&= \mathrm{Tr}\big(\rho^* [\Phi(\sigma) - \mathrm{Tr}(\rho\Phi(\sigma))I]\big) \\
&= \mathrm{Tr}\big(\rho^* \Phi(\sigma)\big) - \mathrm{Tr}\big(\mathrm{Tr}(\rho\Phi(\sigma))\rho^*]\big) \\
&= \mathrm{Tr}\big(\rho^* \Phi(\sigma)\big) - \mathrm{Tr}(\rho\Phi(\sigma)) \quad \text{(Since } \mathrm{Tr}\big(\rho^*\big) = 1\text{)} \\
&= \mathrm{Tr}\big(\rho^* \Phi(\sigma^*)\big) - \mathrm{Tr}(\rho\Phi(\sigma))
\end{aligned}$$

The last line comes from the assumption that the equilibrium (ρ^*, σ^*) is 'fully mixed', i.e. the payoff of an agent when deviating from the Nash equilibrium to any other strategy remains exactly equal to her equilibrium payoff (e.g., Rock-Paper-Scissors). A similar analysis for the second agent results in:

$$d\mathrm{Tr}\big(\sigma^* \log \sigma(t))\big)/dt \quad = \quad -\mathrm{Tr}\big(\sigma^* \Phi^\dagger(\rho^*)\big) + \mathrm{Tr}(\sigma \Phi^\dagger(\rho))$$

These terms cancel out and the theorem follows. □

However in general we cannot use the above formulation, but rather need an argument tailored to the quantum setting using Theorem 3.5 and Observation 4.1. The general proof of Theorem 4.2 is presented here:

D.3 Proof of Theorem 4.2

In the proof of this theorem, we leverage Observation 4.1 and Theorem 3.5.

Proof. First we note that by expanding the limit definition of the derivative we obtain:

$$\begin{aligned}
&\frac{d\big(S(\rho^* \| \rho(t)) + S(\sigma^* \| \sigma(t))\big)}{dt} \\
&= \lim_{h \to 0} \frac{\big(S(\rho^* \| \rho(t+h)) + S(\sigma^* \| \sigma(t+h))\big) - \big(S(\rho^* \| \rho(t)) + S(\sigma^* \| \sigma(t))\big)}{h}
\end{aligned}$$

Here, we apply Observation 4.1 in the following manner: If we perform the substitution $j \to t + h$, $j - 1 \to t$ and $\mu \to h$, we obtain a reformulation of the same limit in the language of stepsizes of duration μ. As $\mu \to 0$, the difference between the continuous time flow solution and its discrete-time Euler approximation is of order $O(\mu^2)$. Thus, we can compute this limit evaluated along the points of the discrete-time MMWU trajectory. This limit by Theorem 3.5 is equal to zero.

□

D.4 Proof of Proposition 4.5

Proof. We first consider the definition of $\rho'(t)$.

$$\begin{aligned}
\rho'(t) &= \frac{\exp(A'(t))}{\mathrm{Tr}(\exp(A'(t)))} \\
&= \frac{\exp\left(\int_0^t \Phi(\sigma(\tau))d\tau - (v^\dagger A(t) v)\mathbb{1}\right)}{\mathrm{Tr}\left(\exp\left(\int_0^t \Phi(\sigma(\tau))d\tau - (v^\dagger A(t) v)\mathbb{1}\right)\right)} \\
&= \frac{\exp\left(\int_0^t \Phi(\sigma(\tau))d\tau\right) \cdot \exp\left(-(v^\dagger A(t) v)\mathbb{1}\right)}{\mathrm{Tr}\left(\exp\left(\int_0^t \Phi(\sigma(\tau))d\tau\right) \cdot \exp\left(-(v^\dagger A(t) v)\mathbb{1}\right)\right)}
\end{aligned}$$

---

[1] Note here that $\rho(t)^{-1}$ exists since the exponential of a matrix is always an invertible matrix.

Note that the denominator can be written as the trace of a matrix where each diagonal entry is the corresponding value of A multiplied by $\exp(-v^\dagger A(t)v)$. Thus, the above can be rewritten as

$$\begin{aligned}\rho'(t) &= \frac{\exp(A(t)) \cdot \exp\left(-(v^\dagger A(t)v)\mathbb{1}\right)}{\exp(-v^\dagger A(t)v) \cdot \mathrm{Tr}\left(\exp(A(t))\right)} \\ &= \frac{\exp(A(t))}{\mathrm{Tr}(\exp(A(t)))} \cdot \mathbb{1} \\ &= \rho(t)\end{aligned}$$

We can perform similar computations to show the same holds true for $\sigma'(t)$ and $\sigma(t)$. □

D.5 Proof of Proposition 4.6

Proof. Consider the map f between A' and ρ. We have shown in Proposition 4.5 that the map in one direction is $\rho = f(A') = \frac{\exp(A')}{\mathrm{Tr}(\exp(A'))}$. This is continuously differentiable. Now let us consider the inverse map f^{-1}. In particular, we show that the inverse mapping exists and is equal to $A' = f^{-1}(\rho) = \log(\rho) - (v^\dagger \log(\rho)v)\mathbb{1}$.

Indeed,

$$\begin{aligned}f(f^{-1}(\rho)) &= \frac{\exp(\log(\rho) - (v^\dagger \log(\rho)v)\mathbb{1})}{\mathrm{Tr}(\exp(\log(\rho) - (v^\dagger \log(\rho)v)\mathbb{1}))} \\ &= \frac{\exp\left(\log(\rho)\right) \cdot \exp\left(-(v^\dagger \log(\rho)v)\mathbb{1}\right)}{\mathrm{Tr}\left(\rho \cdot \exp\left(-(v^\dagger \log(\rho)v)\mathbb{1}\right)\right)} \\ &= \frac{\rho \cdot \exp\left(-(v^\dagger \log(\rho)v)\mathbb{1}\right)}{\exp\left(-(v^\dagger \log(\rho)v)\right)} \\ &= \rho \cdot \mathbb{1} = \rho\end{aligned}$$

where we have used the fact that $\mathrm{Tr}(\rho) = 1$. Furthermore, note that the inverse mapping $f^{-1}(\rho) = \log(\rho) - (v^\dagger \log(\rho)v)\mathbb{1}$ is smooth.

Hence, since f is a bijection and the inverse map is differentiable, $A'(t)$ and $\rho(t)$ are diffeomorphic to each other. Similarly, $B'(t)$ and $\sigma(t)$ are diffeomorphic to each other. □

Lemmas 4.7 and 4.8 are crucial to proving our recurrence result. Typically, volume conservation is not difficult to prove, since the canonical transformation guarantees that the dynamical system becomes separable, which then allows for a simple application of Liouville's theorem. However, proving that the orbits of the dynamical system are bounded is more challenging. We carefully consider the definitions of $A'(t)$ and $B'(t)$ and utilize properties regarding their eigenvalues to show this property.

D.6 Proof of Lemma 4.7

Proof. Let $\Psi_{A'}$ denote the flow of $A'(t)$ and $\Psi_{B'}$ denote the flow of $B'(t)$. First let us consider $A'(t)$.

$$\begin{aligned}\dot{A}' = \frac{dA'}{dt} &= \Phi(\sigma(t)) - (v^\dagger \Phi(\sigma(t))v)\mathbb{1} \\ &= \Phi\left(\frac{\exp(B(t))}{\mathrm{Tr}(\exp(B(t)))}\right) - \left(v^\dagger \Phi\left(\frac{\exp(B(t))}{\mathrm{Tr}(\exp(B(t)))}\right)v\right)\mathbb{1}\end{aligned}$$

Now, we derive the first order partial-derivatives of $F(A')$ with respect to $A'(t)$:

$$\begin{aligned}\frac{\partial F(A')}{\partial A'} &= \frac{\partial}{\partial A'}\left(\Phi\left(\frac{\exp(B(t))}{\mathrm{Tr}(\exp(B(t)))}\right) - \left(v^\dagger \Phi\left(\frac{\exp(B(t))}{\mathrm{Tr}(\exp(B(t)))}\right)v\right)\mathbb{1}\right) \\ &= \frac{\partial}{\partial A'}\left(\Phi\left(\frac{\exp(B'(t))}{\mathrm{Tr}(\exp(B'(t)))}\right) - \left(v^\dagger \Phi\left(\frac{\exp(B'(t))}{\mathrm{Tr}(\exp(B'(t)))}\right)v\right)\mathbb{1}\right) \\ &= 0\end{aligned}$$

Clearly, the partial derivative $\frac{\partial F(A')}{\partial A'}$ only depends on the value of $B'(t)$ and not $A'(t)$. This implies that the vector field is separable, and so the first order partial derivative with respect to $A'(t)$ is zero. By a similar argument, the partial derivative $\frac{\partial F(B')}{\partial B'}$ is also zero. By definition, the diagonal of the Jacobian matrix describing the vector fields is zero, and hence the divergence (which is the trace of the Jacobian) is zero as well. We can then directly apply Liouville's theorem to conclude that the flows $\Psi_{A'}$ and $\Psi_{B'}$ are volume preserving. □

D.7 Proof of Lemma 4.8

Proof. First, $\rho(0) = \frac{\exp(A(0))}{\mathrm{Tr}(\exp(A(0)))}$ (resp. $\sigma(0)$) is bounded away from the boundary, since we assumed that since $A(0)$ and $B(0)$ are finite matrices. Moreover, $A'(0)$ and $B'(0)$ are also finite. The sum of quantum relative entropies at time 0 is thus:

$$S(\rho^*\|\rho(0)) + S(\sigma^*\|\sigma(0)) = \mathrm{Tr}(\rho^*(\log\rho^* - \log\rho(0))) + \mathrm{Tr}(\sigma^*(\log\sigma^* - \log\sigma(0)))$$

This sum is finite since $\rho(0), \sigma(0)$ have full support. Indeed, the support of ρ^* and σ^* are contained in the support of $\rho(0)$ and $\sigma(0)$ respectively. By Theorem 4.2, the sum of quantum relative entropies $S(\rho^*\|\rho(t)) + S(\sigma^*\|\sigma(t))$ is bounded above by a fixed value C for all time t. Now, let $\lambda_{\min}(\rho^*)$ denote the minimum eigenvalue of ρ^*. We have that $\rho^* \geq \lambda_{\min}(\rho^*)\mathbb{1}$. Note the following:

$$\begin{aligned}\mathrm{Tr}(\rho^*(\log\rho^*)) - \mathrm{Tr}(\rho^*(\log\rho(t))) &< C\\ -\mathrm{Tr}(\rho^*(\log\rho(t))) &< D\\ -\mathrm{Tr}(\lambda_{\min}(\rho^*)\mathbb{1}(\log\rho(t))) &< D\end{aligned}$$

where C and D are positive and finite real numbers. Using the fact that the trace of a matrix is basis independent, we use a basis where $\rho(t)$ is diagonal. Here, the minimum eigenvalue of $\rho(t)$ cannot go to zero, otherwise $-\mathrm{Tr}(\lambda_{\min}(\rho^*)\mathbb{1}(\log\rho(t)))$ goes to $+\infty$, a clear contradiction to the inequality above. A similar argument holds for $\sigma(t)$. This is equivalent to saying that all the eigenvalues of $\rho(t)$ and $\sigma(t)$ lie in the interval $[\epsilon, 1-\epsilon]$ for some small $\epsilon > 0$.

Note that the value of $A'_{1,1}(t)$ is always zero by design. We also know by construction that $A'(t)$ is Hermitian, since $\sigma(t)$ is Hermitian, Φ constitutes a hermicity preserving map and the integral of Hermitian matrices remains Hermitian. Hence, the smallest eigenvalue of $A'(t)$, $\lambda_{\min}(A'(t))$ is upper bounded by the smallest element on the diagonal of $A'(t)$, namely $\lambda_{\min}(A'(t)) \leq 0$, since $A'_{1,1}(t) = 0$. If the largest eigenvalue $\lambda_{\max}(A'(t))$ goes to $+\infty$, then the corresponding smallest eigenvalue of $\rho(t) = \frac{\exp(A'(t))}{\mathrm{Tr}(\exp(A'(t)))}$ goes to 0. Likewise if the smallest eigenvalue $\lambda_{\min}(A'(t))$ goes to $-\infty$, the smallest eigenvalue of $\rho(t)$ goes to 0 as well. We have previously shown that this cannot occur, since the eigenvalues of $\rho(t)$ are all bounded away from zero. Hence, it follows that all eigenvalues of $A'(t)$ are bounded away from infinity.

Now, consider the scalar $v^\dagger A'(t)w$ for arbitrary bounded vectors v and w. $A'(t)$ is Hermitian, so by the spectral theorem it has an orthonormal eigenbasis. Now we express v and w in terms of this n-dimensional eigenbasis (which we denote by $\mathbb{e}$):

$$v = \sum_{i=1}^{n} a_i \mathbb{e}_i \quad \text{and} \quad w = \sum_{i=1}^{n} b_i \mathbb{e}_i$$

where a_i and b_i are scalars. Thus,

$$\begin{aligned}v^\dagger A'(t)w &= \left(\sum_{i=1}^{n} a_i \mathbb{e}_i\right)^\dagger A'(t)\left(\sum_{i=1}^{n} b_i \mathbb{e}_i\right)\\ &= \sum_{i=1}^{n} a_i b_i \lambda_i\end{aligned}$$

where λ_i are the eigenvalues of $A'(t)$ corresponding to $\mathbb{e}_i$. Here a_i, b_i and λ_i are all bounded, so it follows that all elements of matrix $A'(t)$ are bounded away from infinity as well. Likewise, the elements of matrix $B'(t)$ remain bounded from infinity and thus the dynamical system described by $A'(t)$ and $B'(t)$ have bounded orbits.

□

E Experiments

In order to formalize a method for simulating and understanding quantum games, we propose a method to generate complex R matrices with the same eigenvalues as the values of the classical payoff matrix. This means that for any standard eigenbasis of a classical game, we create a new, complex eigenbasis where the eigenvalues are the same, but the eigenvectors are complex. Intuitively, this can be viewed as mapping a classical game matrix to the Hilbert space, effectively allowing generalizations of classically studied games to the semi-definite context. We introduce the following lemma, which describes a method to obtain a basis transformation to Hilbert space.

Lemma E.1. *For any $n \times n$ two player zero-sum game with classical payoff matrix P, let $\lambda_{i,j} = P_{i,j}$ and let V and W be unitary matrices of size $n \times n$. Then, the matrix R defined as*

$$R = \sum_{i,j} \lambda_{i,j} (V_i \otimes W_j)(V_i^\dagger \otimes W_j^\dagger) \tag{29}$$

produces a transformation from the classical payoff space $\mathcal{P}$ to the density matrix space $\mathcal{C}$. In particular, R is a complex $n^2 \times n^2$ matrix which satisfies the following properties:

- *The matrix R is Hermitian.*
- *The eigenvalues of R correspond to the elements of the classical payoff matrix P.*

Proof. First note that Equation 29 is equivalent to taking the following basis transformation:

$$\sum_{i,j} \lambda_{i,j} (V_i \otimes W_j)(V_i^\dagger \otimes W_j^\dagger) = (V \otimes W)\Lambda(V^\dagger \otimes W^\dagger)$$

where $\Lambda = \begin{bmatrix} \lambda_{1,1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_{n,n} \end{bmatrix}$ is the $n^2 \times n^2$ diagonal matrix formed by placing the elements of classical payoff matrix P on the diagonal. Moreover, V and W are both unitary, so the tensor products $V \otimes W$ and $V^\dagger \otimes W^\dagger$ are also unitary. This holds because:

$$\begin{aligned} (V \otimes W)(V \otimes W)^\dagger &= (V \otimes W)(V^\dagger \otimes W^\dagger) \\ &= VV^\dagger \otimes WW^\dagger \\ &= \mathbb{1} \otimes \mathbb{1} = \mathbb{1} \end{aligned}$$

This implies that R can be written as a unitary diagonalization of the form $PDP^\dagger$, where $P = (V \otimes W)$ and $D = \Lambda$. Hence R is Hermitian and has real eigenvalues by the spectral theorem. In addition, by construction the eigenvalues of R are the elements of the classical payoff matrix P. $\square$

Moreover, the initial conditions of a simulation of a two-player quantum zero-sum game can be similarly obtained by using the same orthonormal bases as in Lemma E.1 and applying a change of basis to the initial conditions in the standard basis.

E.1 Additional Experimental Details (Matching Pennies)

In the main text, we showed several simulations that illustrate the theoretical results in Sections 3 and 4. Here we present detailed descriptions of the simulations, as well as additional experiments which serve to corroborate auxiliary results in the paper.

To generate Figure 1, we used a modified version of Algorithm 1, where instead of constant step-size μ we define $\mu = \log(1 + \frac{1}{t^a})$, where a is an exponent which determines the rate of decrease of the step-size, and t is the time-step of the simulation. Intuitively, this means that for higher values of a, the step-sizes go to zero at a faster rate. Empirically, we observe that if the exponent is greater than $\frac{1}{2}$, then we can clearly see the phenomenon described in Theorem 3.5, since the quantum relative entropy in the system increases when the trajectories move away from the fully mixed Nash equilibrium.

In the case of replicator dynamics, we show experimentally (Figure 3) that the sum of quantum relative entropies for both players is a constant of motion as shown in Theorem 4.2. We first use

Lemma E.1 to obtain the R matrix of a quantum game given some complex basis and classical payoff matrix (Matching Pennies, for instance). Then, we run the discrete Algorithm 1 with decreasing step-size to compute the interior ϵ-approximate Nash of the quantum game. The replicator equations are then solved given a set of initial conditions for each player. Thereafter, we compute the quantum relative entropy between each player's strategy at each discretized time step and the ϵ-approximate Nash equilibrium.

Notice that although the quantum relative entropies of the first player (pink) and second player (purple) are oscillating over time, their sum remains approximately constant.

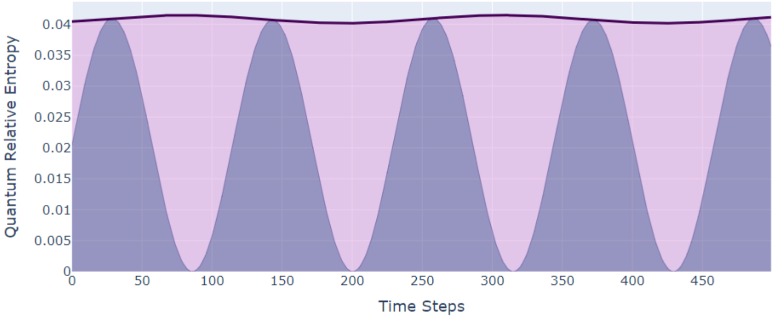


Figure 3: Approximately constant sum of quantum relative entropy values for two-player quantum Matching Pennies game.

Finally, in order to generate Figure 2, we perform a similar simulation to the above. However, we also transform the trajectories of each player using the Pauli matrices, which are given as:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{30}$$

The expectations of the density matrices under this transformation (commonly referred to as the Bloch vector) produce the coordinates on the x, y and z axes within the Bloch sphere. We plot these coordinates for each time step and hence obtain the trajectories for each player in the Bloch sphere.

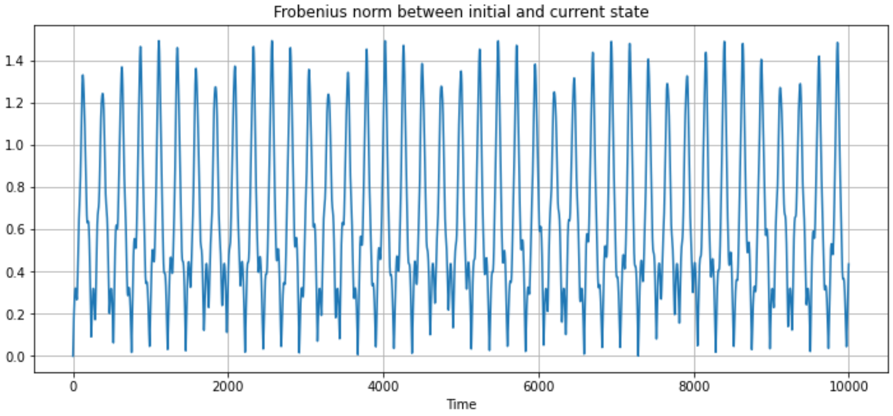
E.2 Larger-Scale Experiments

In the main text, we study a single-qubit system in the form of a Matching Pennies game. However, our results extend to games with larger strategy sets, some of which can be interpreted as multi-qubit systems. We study a 16×16 game with a payoff matrix in $[-1, 1]$ which can be played via 2 qubits. Moreoever, this game is selected because it has an interior (and uniform) Nash equilibrium. Similarly, we generate an 64×64 generalization of the 16×16 game to represent an example of a 3-qubit game with interior Nash. In Figure 4, we plot the Frobenius norm between the initial condition and the system state when updated using quantum replicator dynamics (Equations 17). Notice that despite the relatively erratic behaviour of the strategies, they eventually return to the initial condition, implying that recurrence holds (Theorem 4.3).

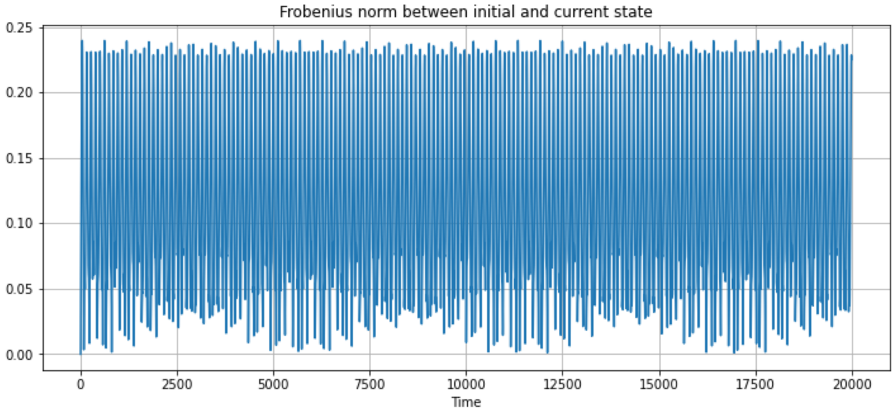
Furthermore, to confirm the volume conservation property shown in Theorem 4.2 we perform the same simulation as in Figure 3 for the same 2 and 3 qubit systems as above, obtaining Figure 5. However, a minor note is that we perform these simulations without the basis transform of Lemma E.1 to avoid potential numerical errors since the matrices are of larger dimension. The initial conditions are random density matrices, as before. Note that the sum of quantum relative entropies between each player and the interior Nash equilibrium is constant as expected from our theoretical results.

E.3 Reproducibility Details

All experiments performed for this work were done using Python 3.7 and have been compiled into a Jupyter notebook for ease of viewing. Running the code requires basic scientific computing packages such as NumPy, SciPy and Cython, as well as data visualization packages such as Matplotlib and Plotly. In addition, we use QuTip [29], a quantum simulation package. Most of our code has been edited such that it can be easily executed on a standard computer in a matter of minutes.

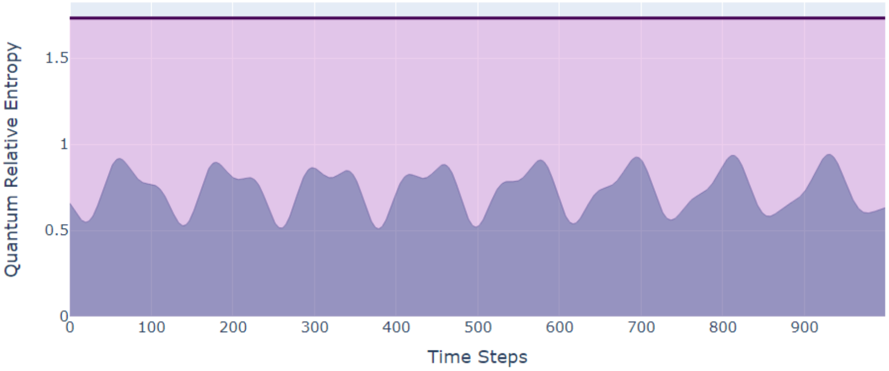


2-qubit system

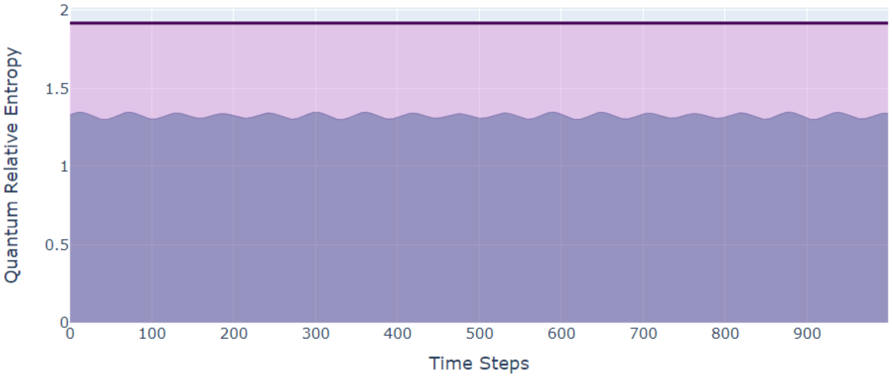


3-qubit system

Figure 4: Frobenius norm between the initial condition and current system state for 2 and 3 qubit systems with interior Nash equilibrium. Note that at approx. 7200 iterations the 2-qubit system returns arbitrarily close to 0, and the same occurs at approx. 12000 iterations for the 3-qubit system.



2-qubit system



3-qubit system

Figure 5: Sum of quantum relative entropies of each player to the Nash equilibrium for 2 and 3 qubit systems with interior Nash.