

Hamiltonian of polymatrix zero-sum games

Toshihiro Ota^{1,2,*} and Yuma Fujimoto^{1,3,†}

¹*AI Lab, CyberAgent, Shibuya, Tokyo 150-0002, Japan*

²*THEMS, RIKEN, Wako, Saitama 351-0198, Japan*

³*RCIES, Soken University, Hayama, Kanagawa 240-0193, Japan*



(Received 22 October 2025; accepted 3 December 2025; published 26 December 2025)

The understanding of a dynamical system's properties can be significantly advanced by establishing it as a Hamiltonian system and then systematically exploring its inherent symmetries. By formulating agents' strategies and cumulative payoffs as canonically conjugate variables, we identify the Hamiltonian function that generates the dynamics of polymatrix zero-sum games. We reveal the symmetries of our Hamiltonian and derive the associated conserved quantities, showing how the conservation of probability and the invariance of the Fenchel coupling are intrinsically encoded within the system. Furthermore, we propose the *dissipation Follow-the-Regularized-Leader* (DFTRL) dynamics by introducing a perturbation that dissipates the Fenchel coupling, proving convergence to the Nash equilibrium and linking DFTRL to last-iterate convergent algorithms. Our results highlight the potential of Hamiltonian dynamics in uncovering the structural properties of learning dynamics in games and pave the way for broader applications of Hamiltonian dynamics in game theory and machine learning.

DOI: [10.1103/3gfh-lg1q](https://doi.org/10.1103/3gfh-lg1q)

I. INTRODUCTION

Dynamical system analysis plays a fundamental role in game theory. In the context of learning in games [1], multiple agents dynamically update their strategies over time in an attempt to maximize their own payoffs. The utility function of each agent depends not only on their own strategy but also on those of others, and simultaneous optimal strategies across all agents are characterized by the Nash equilibrium [2]. However, when the utilities of agents are in direct conflict (zero-sum games), the resulting learning dynamics often exhibit recurrent behaviors and convergence to a Nash equilibrium may fail to occur [3].

The learning algorithm known as Follow the Regularized Leader (FTRL) provides a systematic formulation of dynamics in online learning [3,4]. In the FTRL algorithm, each agent's strategy evolves according to their cumulative payoff, and it has been shown that in zero-sum games the Fenchel coupling is conserved throughout the dynamics. This conservation law implies that, under the FTRL dynamics, the distance between the strategies and the Nash equilibrium remains essentially invariant, thereby explaining the failure of convergence to the equilibrium. Although heuristic modifications to the FTRL algorithm have been proposed to

mitigate this issue, a comprehensive understanding remains elusive.

Conserved quantities generally provide crucial insights into the system under consideration, as a sufficient number of conserved quantities can even render the system solvable. Hamiltonian dynamics offers a sophisticated mathematical framework for systematically analyzing *symmetries* and associated conservation laws. Attempts to examine learning dynamics in games using a Hamiltonian perspective date back several decades [5], yet most studies have rarely extended beyond two-agent scenarios. Bailey and Piliouras [6] made a notable observation that, in polymatrix zero-sum games [7–9], it is promising that agents' strategies and their cumulative payoffs are thought of as canonical conjugates for formulating the FTRL dynamics as a Hamiltonian system. Nevertheless, their approach resulted only in the reconstruction of existing conserved quantities such as the Fenchel coupling, without identifying a genuine Hamiltonian function for polymatrix zero-sum games [10]. In any case, previous studies have not fully exploited the paradigm of “symmetry and conservation law” inherent in Hamiltonian dynamics.

In this paper, we address the aforementioned issues by leveraging an analysis of symmetry, conservation laws, and their breaking within the framework of Hamiltonian dynamics. Specifically, our contributions are as follows:

(1) We establish the Hamiltonian function H_{FTRL} of the learning dynamics in polymatrix zero-sum games, with agents' strategies and cumulative payoffs serving as canonically conjugate variables, and demonstrate that the Hamiltonian system generated by H_{FTRL} produces the FTRL dynamics.

(2) We elucidate the symmetries of the Hamiltonian H_{FTRL} and derive the associated conserved quantities, revealing

*Contact author: ota_toshihiro@cyberagent.co.jp

†Contact author: fujimoto.yuma1991@gmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

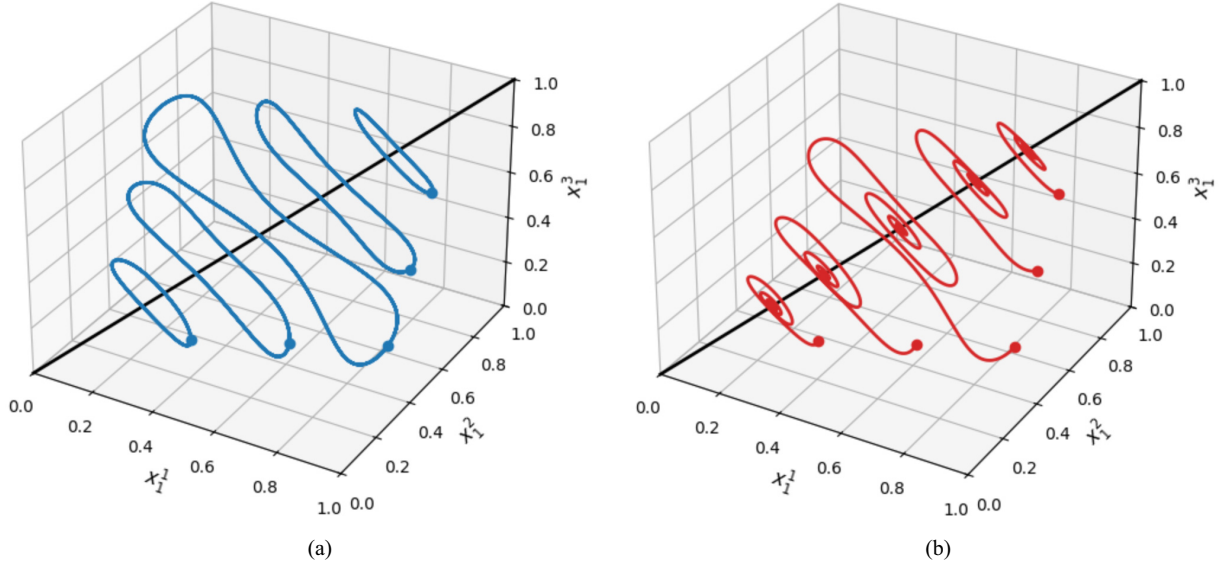


FIG. 1. Three-player Matching Pennies. The diagonal straight line from $(0, 0, 0)$ to $(1, 1, 1)$ represents the Nash equilibria. The solution trajectories show (a) cyclic (FTRL) and (b) convergent (DFTRL) behavior.

explicitly how the conservation of probability in strategies and the temporal invariance of the Fenchel coupling are intrinsically encoded within H_{FTRL} . We also examine angular-momentum-like conservation laws, which have not been studied in the literature.

(3) We propose the *dissipation FTRL* (DFTRL) dynamics by introducing a perturbation designed to dissipate (monotonically decrease) the Fenchel coupling along the trajectory. We prove that the DFTRL dynamics converges to the Nash equilibrium (as numerically illustrated in Fig. 1) and coincides with the continuous optimistic (CO) and extra-gradient FTRL algorithms in the limit of infinitesimal perturbations.

Related work

The analysis of learning dynamics in zero-sum games from the perspective of Hamiltonian dynamics has a long history [5, 11–15]. These studies typically adopt a formulation in which the canonical variables are given by a pair of strategies, thereby restricting their analysis to two-agent settings. Moreover, the Hamiltonians proposed in these works are tailored to specific game dynamics, such as the replicator, projection, or best-response dynamics, and do not possess any symmetry giving rise to known conserved quantities. Another line of research has focused on the Poisson structure of evolutionary games [16] and the conserved quantities of the game within that framework [17]. While their Hamiltonian can also be defined for non-zero-sum games, it needs an equilibrium in its definition and is restricted to replicator dynamics.

To address the nonconvergent behavior observed in FTRL dynamics, several modified versions of FTRL have been proposed. Representative examples include the optimistic [18–22], extra-gradient [18, 23–25], and negative-momentum [26–30] FTRL algorithms. While all these variants of FTRL are known to converge to the Nash equilibrium, their similarities and differences remain an active subject of discussion. For instance, it has been reported that the extra-gradient FTRL

converges even in time-varying games, where the payoff matrix evolves over time, whereas the other two do not [31, 32]. In Sec. III C, we will demonstrate that the proposed DFTRL algorithm offers a unified perspective for classifying these existing variants in the continuum limit.

II. PRELIMINARIES

To fix the notation, in this section we describe a brief overview of polymatrix games and the FTRL learning dynamics. We basically follow the discussions in Refs. [3, 6].

We denote a graphical *polymatrix game* Γ by the following three tuple:

$$\Gamma = (\mathcal{N}, \mathcal{A}, \mathcal{U}), \quad (1)$$

where $\mathcal{N} = \{1, \dots, n\}$ represents the set of agents (players), $\mathcal{A} = \{A_1, \dots, A_n\}$ are finite sets of actions that an agent i takes,

$$A_i = \{a_1^i, \dots, a_{N_i+1}^i\}, \quad N_i \geq 0, \quad (2)$$

and $\mathcal{U} = \{U^{(ij)}\}_{i,j \in \mathcal{N}}$ is a set of payoff matrices:

$$U^{(ij)} \in \mathbb{R}^{|A_i| \times |A_j|}, \quad U^{(ii)} \equiv 0. \quad (3)$$

Each point of the N_i -simplex,

$$\mathcal{X}^{N_i} := \left\{ z \in \mathbb{R}_{\geq 0}^{|A_i|} \mid \sum_{a \in A_i} z_a = 1 \right\}, \quad (4)$$

is called a (*mixed*) *strategy*, whose entries correspond to the probabilities of the actions that an agent may select. In particular, a strategy is called *fully mixed* when it lies in the relative interior of the simplex. We collectively denote a set of n strategies as

$$x = (x^1, \dots, x^n) \in \mathcal{X}^{N_1} \times \dots \times \mathcal{X}^{N_n} =: \mathcal{X}. \quad (5)$$

Given the mixed strategies of agents i and j , $x^i \in \mathcal{X}^{N_i}$ and $x^j \in \mathcal{X}^{N_j}$, agent i receives *utility* from agent j : $(x^i)^\top U^{(ij)} x^j$. The total utility of agent i is [33]

$$u_i(x^i; x^{j(\neq i)}) = \sum_{j(\neq i)} (x^i)^\top U^{(ij)} x^j, \quad (6)$$

which results in n optimization problems:

$$\max_{x^i \in \mathcal{X}^{N_i}} u_i(x^i; x^{j(\neq i)}) = \max_{x^i \in \mathcal{X}^{N_i}} (x^i)^\top \sum_{j(\neq i)} U^{(ij)} x^j, \quad (7)$$

for $i = 1, \dots, n$. In Γ , $x^* = (x^{1*}, \dots, x^{n*}) \in \mathcal{X}$ is called a *Nash equilibrium* if

$$\begin{aligned} \forall i \in \mathcal{N}, \quad \forall x^i \in \mathcal{X}^{N_i}, \\ u_i(x^{i*}; (x^{j(\neq i)})^*) \geq u_i(x^i; (x^{j(\neq i)})^*). \end{aligned} \quad (8)$$

That is, the Nash equilibrium is an n -simultaneous solution to the above optimization problem.

A game Γ is called *zero sum (coordination)* if

$$\forall i, j \in \mathcal{N}, \quad U^{(ji)} = \sigma(U^{(ij)})^\top, \quad (9)$$

where $\sigma = -1$ (+1). We denote zero-sum or coordination polymatrix games by $\Gamma_\sigma := (\mathcal{N}, \mathcal{A}, \mathcal{U}, \sigma)$. A significant nature of Γ_{-1} is that the total utility of all the agents always vanishes due to the zero-sum property,

$$\sum_{i \in \mathcal{N}} u_i(x^i; x^{j(\neq i)}) = \sum_{i, j \in \mathcal{N}} (x^i)^\top U^{(ij)} x^j = 0, \quad (10)$$

which implies that the total payoffs circulate among agents and are conserved within the game.

In the subsequent sections, we may refer to curves such as

$$\phi : I \rightarrow \mathcal{X}^{N_i} \text{ or } \mathbb{R}^{|A_i|} \quad (11)$$

as the *learning dynamics*, where I is an interval, and write their images in the codomains by $x^i(t)$ or $y^i(t)$ for convenience.

Follow the regularized leader

Let $C_{SC}^\infty(M, \mathbb{R}) (\subset C^\infty(M, \mathbb{R}))$ be the set of strongly convex and L -smooth functions on a convex subset M of Euclidean space. Given initial payoff vectors $y^i(0) \in \mathbb{R}^{|A_i|}$, the agents in Γ_σ update their strategies by [34]

$$y^i(t) = y^i(0) + \int_0^t \sum_{j(\neq i)} U^{(ij)} x^j(s) ds, \quad (12)$$

$$x^i(t) = \arg \max_{z \in \mathcal{X}^{N_i}} (\langle z, y^i(t) \rangle - h_i(z)), \quad (13)$$

where $\langle \cdot, \cdot \rangle$ is the Euclidean inner product and $h_i \in C_{SC}^\infty(\mathcal{X}^{N_i}, \mathbb{R})$ are regularizer functions. This learning rule is known as the Follow-the-Regularized-Leader dynamics [3]. For the function $h_i^* : \mathbb{R}^{|A_i|} \rightarrow \mathbb{R}$, which is the convex conjugate of h_i ,

$$h_i^*(y^i) = \max_{z \in \mathcal{X}^{N_i}} (\langle z, y^i \rangle - h_i(z)), \quad (14)$$

it is known [35] that the right-hand side of Eq. (13) can be expressed as the gradient of h_i^* :

$$x^i(t) = \nabla h_i^*(y^i) |_{y^i=y^i(t)}. \quad (15)$$

To summarize, the FTRL dynamics is a dynamical system consisting of the following three components:

- (1) Setup of game: $\Gamma_\sigma = (\mathcal{N}, \mathcal{A}, \mathcal{U}, \sigma)$;
- (2) Regularizers: a set of $h_i \in C_{SC}^\infty(\mathcal{X}^{N_i}, \mathbb{R})$ or their dual h_i^* ; and
- (3) Learning rule: Equation (13) or its dual (15), with Eq. (12).

In particular, it is noted that this dynamical system is characterized by the choice of payoff matrices $\{U^{(ij)}\}_{i,j \in \mathcal{N}}$ (the “rule” of a game) and the choice of “potential functions” h_i (or h_i^*).

III. THE HAMILTONIAN OF POLYMATRIX ZERO-SUM GAMES

We focus on polymatrix zero-sum games Γ_{-1} and show that the FTRL dynamics in Γ_{-1} is understood as a Hamiltonian system. We investigate the system through the lens of *symmetry* and the conservation laws. For readers unfamiliar with Hamiltonian dynamics, see Appendix B.

A. FTRL dynamics as Hamiltonian systems

As noted in Eq. (10), the total utility in zero-sum games is a constant over time and thus could serve as a Hamiltonian function for the learning dynamics. With this viewpoint, let us first observe that, by taking time derivatives, the FTRL dynamics in Eqs. (12) and (15) can be rewritten as follows:

$$\begin{aligned} \frac{dy^i(t)}{dt} &= \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) \\ &= \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j) |_{y^j=y^j(t)}, \end{aligned} \quad (16)$$

$$\begin{aligned} \frac{dx^i(t)}{dt} &= \nabla \nabla h_i^*(y^i) |_{y^i=y^i(t)} \frac{dy^i(t)}{dt} \\ &= \text{Hess}(h_i^*(y^i(t))) \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t), \end{aligned} \quad (17)$$

where we have used Eq. (15) and Hess is the Hessian. A point here is that the right-hand sides of these equations are Lipschitz continuous, owing to the strong convexity of the regularizer functions. From the standard existence and uniqueness theorem of initial value problems, Eqs. (16) and (17) on I admit a unique solution that is indeed the FTRL dynamics [Eqs. (12) and (15)]. Hence, these differential equations are equivalent to the FTRL dynamics.

Now, we define the phase space of Γ_{-1} by $M_{\Gamma_{-1}} = T^*\mathbb{R}_{\geq 0}^{\sum_i |A_i|} \simeq \mathbb{R}_{\geq 0}^{\sum_i |A_i|} \times \mathbb{R}^{\sum_i |A_i|}$ equipped with the symplectic structure $\omega \simeq J$ as described in Appendix B. For a given choice of $h_i^* \in C_{SC}^\infty(\mathbb{R}^{|A_i|}, \mathbb{R})$, we define the *FTRL Hamiltonian* $H_{\text{FTRL}} : M_{\Gamma_{-1}} \rightarrow \mathbb{R}$ by

$$H_{\text{FTRL}}(x, y) = \sum_{i,j \in \mathcal{N}} \langle \nabla h_j^*(y^j), U^{(ji)} x^i \rangle. \quad (18)$$

Theorem 1. The FTRL dynamics in the form of Eqs. (16) and (17) is a Hamiltonian system defined by Eq. (18).

Proof. What we need to show is that the flow equation on $M_{\Gamma_{-1}}$ generated by Eq. (18) accompanied with the symplectic structure reduces to the FTRL dynamics [(16) and (17)]. First,

the flow equation for $x(t)$ (“position”) induced by Eq. (18) is

$$\begin{aligned} \frac{dx^k(t)}{dt} &= \nabla_{y^k} H_{\text{FTRL}}(x, y) \big|_{x=x(t), y=y(t)} \\ &= \sum_{i,j \in \mathcal{N}} \delta^{kj} (\nabla \nabla h_j^*(y^j(t)))^\top U^{(ji)} x^i(t) \\ &= \text{Hess}(h_k^*(y^k(t))) \sum_{i \in \mathcal{N}} U^{(ki)} x^i(t), \end{aligned} \quad (19)$$

which is equivalent to Eq. (17). Next, by utilizing the symmetry of the inner product, $\langle \alpha, \beta \rangle = \langle \beta, \alpha \rangle$, one finds the flow equation for $y(t)$ (“momentum”),

$$\begin{aligned} \frac{dy^k(t)}{dt} &= -\nabla_{x^k} H_{\text{FTRL}}(x, y) \big|_{x=x(t), y=y(t)} \\ &= -\sum_{i,j \in \mathcal{N}} \delta^{ki} (U^{(ji)})^\top \nabla h_j^*(y^j(t)) \\ &= \sum_{j \in \mathcal{N}} U^{(kj)} \nabla h_j^*(y^j(t)), \end{aligned} \quad (20)$$

where the last equality follows from the assumption that our game is zero sum: $(U^{(jk)})^\top = -U^{(kj)}$ for all $j, k \in \mathcal{N}$. ■

The aforementioned Lipschitz continuity ensures that the Hamiltonian system of H_{FTRL} describes the FTRL dynamics in polymatrix zero-sum games. The energy conservation law in this Hamiltonian system leads to the conservation law of the total utility of the game Γ_{-1} :

$$\begin{aligned} E_{\text{FTRL}}(t) &= H_{\text{FTRL}}(x(t), y(t)) \\ &= \sum_{i,j \in \mathcal{N}} \langle x^i(t), U^{(ij)} x^j(t) \rangle \equiv 0, \end{aligned} \quad (21)$$

which is the equivalent of Eq. (10), as expected. Now that the learning dynamics in Γ_{-1} is characterized by our proposed Hamiltonian (18), we investigate the system in terms of symmetry and the conservation laws in the following subsections.

B. Symmetry and conservation law from the FTRL Hamiltonian

The Hamiltonian (18) is defined on the symplectic manifold $M_{\Gamma_{-1}} \simeq \mathbb{R}_{\geq 0}^{\sum_i |A_i|} \times \mathbb{R}_{\leq 0}^{\sum_i |A_i|}$. The base manifold (configuration space) $\mathbb{R}_{\geq 0}^{\sum_i |A_i|}$ seems too large to be the space of strategies, since the strategies x^i in games should be thought of as probabilities of actions and have to sum to 1: $\sum_a x_a^i = 1$. Interestingly, this condition is implemented as the *symmetry* of our Hamiltonian and its conservation law, as the following proposition demonstrates.

Proposition 1. If $\nabla h_i^*(y^i)$ are translationally invariant for all $i \in \mathcal{N}$, then the solutions $x(t)$ to the Hamiltonian system of H_{FTRL} are constrained onto the probability simplices $\mathcal{X}$.

The proof follows straightforwardly from the following lemma:

Lemma 1. If a Hamiltonian function on $M_{\Gamma_{-1}}$ has the translation symmetry in $y^i \in \mathbb{R}^{|A_i|}$, then the following sum of the elements of the corresponding conjugate variable is a conserved quantity:

$$G_S^i(x, y) = \sum_{a \in A_i} x_a^i. \quad (22)$$

This lemma is derived from Noether’s theorem (see Appendixes B and D 1).

Proof. Given translationally invariant $\nabla h_i^*(y^i)$ for all $i \in \mathcal{N}$, the FTRL Hamiltonian (18) has the translation symmetry in $y^i \in \mathbb{R}^{|A_i|}$. From Lemma 1, in this Hamiltonian system we have the conserved quantities

$$G_S^i(x(t), y(t)) = \sum_{a \in A_i} x_a^i(t) = \text{const.}, \quad (23)$$

for all i . Suppose that the initial conditions $\sum_a x_a^i(0) = 1$ are imposed, then the solution trajectory of the Hamiltonian system generated by Eq. (18) remains restricted to $x(t) \in \mathcal{X}$, completing the proof. ■

We find that the Hamiltonians for both the entropic regularizer (the replicator dynamics [36–38]) and the Euclidean regularizer (the projection dynamics [39–43]) have translation symmetry in all $y^i \in \mathbb{R}^{|A_i|}$, implying that they yield reasonable Hamiltonian dynamics in games:

Corollary 1. In the Hamiltonian systems of H_{FTRL} for the entropic regularizer and the Euclidean regularizer, the solution trajectories of $x(t)$ are constrained onto the probability simplices $\mathcal{X}$.

Proof. From Proposition 1, it suffices to show the translation invariance of the corresponding gradients of the dual regularizers, ∇h_i^* . For the entropic regularizer, we have

$$\nabla h_i^*(y^i)_a = \frac{\exp(y_a^i)}{\sum_{a'} \exp(y_{a'}^i)}, \quad (24)$$

and for the Euclidean regularizer,

$$\nabla h_i^*(y^i)_a = y_a^i - \frac{1}{|A_i|} \left(\sum_{a'} y_{a'}^i - 1 \right). \quad (25)$$

These explicit formulas and other details can be found in Appendix A. One finds that the right-hand sides of these gradients are invariant under the constant shift in y^i , $y_a^i = y_a^i + \epsilon$ for all $a \in A_i$, which proves the statement. ■

In addition, for the entropic and Euclidean regularizers, we know that $\sum_a \nabla h_i^*(y^i)_a = 1$, which leads us to another conserved quantity.

Proposition 2. Suppose we have a game Γ_{-1} with the Hamiltonian (18). If a fully mixed Nash equilibrium exists and the elements of $\nabla h_j^*(y^j)$ sum to a constant independent of j , then the following function is a conserved quantity of the Hamiltonian system:

$$G_F(x, y) = \sum_{i \in \mathcal{N}} (h_i(x^{i*}) + h_i^*(y^i) - \langle y^i, x^{i*} \rangle). \quad (26)$$

Proof. The proof is derived by applying Noether’s theorem to the Hamiltonian (18) along with the symmetry under the canonical transformation,

$$x'^i = x^i + \nabla h_i^*(y^i) - x^{i*}, \quad y'^i = y^i, \quad (27)$$

for all $i \in \mathcal{N}$. Using the assumptions, one can confirm that the FTRL Hamiltonian H_{FTRL} is invariant under this transformation:

$$H_{\text{FTRL}}(x', y') = H_{\text{FTRL}}(x, y). \quad (28)$$

It turns out that the generator of this canonical transformation takes the form

$$g_F(x, y) = \sum_{i \in \mathcal{N}} (h_i^*(y^i) - \langle y^i, x^{i*} \rangle) + \text{const.} \quad (29)$$

Since $\sum_i h_i(x^{i*})$ is a constant, from Noether's theorem it follows that the function G_F is the conserved quantity in the Hamiltonian system of H_{FTRL} . ■

The complete proof is provided in Appendix D2. In the literature, the following function,

$$F(p, y) = \sum_{i \in \mathcal{N}} (h_i(p^i) + h_i^*(y^i) - \langle y^i, p^i \rangle), \quad (30)$$

is referred to as the Fenchel coupling [43,44]. The Fenchel coupling at $p = x^*$, which measures the distance to the Nash equilibrium, is a well-known conserved quantity of the FTRL dynamics in zero-sum games [3]. Proposition 2 shows that our Hamiltonian (18) involves the Fenchel coupling as a conserved quantity associated with the corresponding symmetry. Due to this, we also refer to G_F as the Fenchel coupling.

If we interpret the symmetry transformation (27), it can be viewed as a coordinate shift under which the total utility of the system remains invariant when each agent's "current strategy" x^i is adjusted by the difference between its "regularized best response" $\nabla h_i^*(y^i)$ and the equilibrium x^{i*} . If an agent is at equilibrium ($\nabla h_i^*(y^i) = x^{i*}$), no shift is applied; otherwise, the strategy x is adjusted by the magnitude of this error. [Due to the zero-sum property, the Hamiltonian (18) is insensitive to shifts in x corresponding to the difference from the best response.] This shift coincides exactly with the gradient of the Fenchel coupling. Then, by Noether's theorem, the Fenchel coupling, which serves as the generating function for this shift, must remain constant. This is analogous to the conservation of momentum in physics, where the symmetry is invariance of a Hamiltonian under spatial translations (shifts).

Finally, we propose two "angular momentum" conservation laws for the Hamiltonian systems of H_{FTRL} . The key points are the commutativity of rotation matrices and the payoff matrices, and the equivariance of ∇h_i^* under the rotation.

Proposition 3. Suppose we have a game Γ_{-1} with the Hamiltonian (18).

(1) Let $R_i \in \text{SO}(|A_i|)$, $i \in \mathcal{N}$, be rotation matrices. If the followings hold for all $i, j \in \mathcal{N}$,

$$U^{(ij)} R_j = R_i U^{(ij)}, \quad (31)$$

$$\nabla h_i^*(R_i y^i) = R_i \nabla h_i^*(y^i), \quad (32)$$

then the following function is a conserved quantity of the Hamiltonian system:

$$G_K(x, y) = \sum_{i \in \mathcal{N}} \langle x^i, K_i y^i \rangle, \quad (33)$$

where $R_i = e^{\epsilon K_i}$ and $K_i \in \mathfrak{so}(|A_i|)$.

(2) Let $\mathcal{R} \in \text{SO}(n)$ be a rotation matrix. Suppose that all the action sets and the regularizers are identical for all the agents: $A_i \simeq A$ and $h_i^* = h^*$ for all $i \in \mathcal{N}$. If the followings hold for all $i, j \in \mathcal{N}$,

$$U \mathcal{R} = \mathcal{R} U, \quad (34)$$

$$\nabla h^*(\mathcal{R} y) = \mathcal{R} \nabla h^*(y), \quad (35)$$

where U is the block payoff matrix with the entries $U_{ij} = U^{(ij)}$, then the following function is a conserved quantity of the Hamiltonian system:

$$G_\Omega(x, y) = \sum_{i, j \in \mathcal{N}} \langle x^i, \Omega_{ij} y^j \rangle =: \langle x, \Omega y \rangle, \quad (36)$$

where $\mathcal{R} = e^{\epsilon \Omega}$ and $\Omega \in \mathfrak{so}(n)$.

The proof is provided in Appendix D3. This proposition indicates that there may be game-theoretic analogs of the conservation laws for angular momentum in physics. The first one shows the conservation of the total "angular momentum" of each agent, and the second corresponds to the conservation of the system's total angular momentum. We currently do not have a clear intuitive interpretation of these conserved quantities in game theory and are not aware of any existing literature on this topic. This result suggests that the Hamiltonian (18) may generate conserved quantities that are not recognized in the standard FTRL dynamics context. These results remain interesting from a mathematical standpoint and may provide a deeper understanding of learning dynamics in games.

We here provide a useful expression of Proposition 3 as a corollary.

Corollary 2. Suppose we have a game Γ_{-1} with the Hamiltonian (18) and assume that the equivariance condition holds for below cases.

(1) Suppose that all the action sets are identical, $A_i \simeq A$ for all $i \in \mathcal{N}$. If the payoff matrices are given by a single skew-symmetric matrix $B \in \mathfrak{so}(|A|)$ as

$$U^{(ij)} = cB, \quad B^\top = -B, \quad (37)$$

where $c \in \mathbb{R}$ is a constant, then the rotation matrices $R_i = e^{\epsilon B} \in \text{SO}(|A|)$ commute with the payoff matrices and the following quantity is conserved:

$$G_K(x, y) = \sum_{i \in \mathcal{N}} \langle x^i, B y^i \rangle. \quad (38)$$

(2) If the payoff matrices are given by the tensor product of a skew-symmetric matrix $\Sigma = (\sigma_{ij}) \in \mathfrak{so}(n)$ and a symmetric matrix S as

$$U^{(ij)} = \sigma_{ij} S, \quad S^\top = S, \quad (39)$$

then the rotation matrix $\mathcal{R} = e^{\epsilon \Sigma} \in \text{SO}(n)$ commutes with the payoff matrix U and the following quantity is conserved:

$$G_\Omega(x, y) = \langle x, \Sigma y \rangle. \quad (40)$$

One should note that the payoff matrices in these cases, $U^{(ij)} = cB$ and $U^{(ij)} = \sigma_{ij} S$, satisfy the zero-sum condition, $(U^{(ij)})^\top = -U^{(ji)}$, due to the skew-symmetry of B and the symmetry of S . The payoff matrices in the first case define a symmetric game, $U^{(ij)} = U^{(ji)}$, while those in the second case define a skew-symmetric game, $U^{(ij)} = -U^{(ji)}$. Thus, roughly speaking, a class of symmetric zero-sum games has the rotation symmetry in the strategy space with the corresponding angular momentum(like) conservation law, while a class of skew-symmetric zero-sum games has the rotation symmetry among agents with another angular momentum conservation law. In Appendix D3, we demonstrate the explicit examples

of these angular momentum conservation laws: two-player Rock-Paper-Scissors as a symmetric game and two-player Matching Pennies as a skew-symmetric game.

C. Dissipation FTRL dynamics

We now have the conserved quantity G_F of the Hamiltonian system, which measures the distance from a fully mixed Nash equilibrium. However, it implies that the learning dynamics does not converge to the Nash equilibrium, but rather shows a cyclic behavior around it. To resolve this issue, we introduce a perturbation to the FTRL dynamics that breaks the conservation law for G_F , so that the learning dynamics converges to the Nash equilibrium, while the Hamiltonian itself (the total utility) is kept as in Eq. (21) to sustain the global zero-sum property of the game.

We consider a modified FTRL dynamics by adding perturbation terms to Eqs. (16) and (17), with $\alpha \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned}\frac{dx^i(t)}{dt} &= \text{Hess}(h_i^*(y^i(t))) \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) + \alpha f^i(x(t), y(t)), \\ \frac{dy^i(t)}{dt} &= \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j(t)) + \alpha g^i(x(t), y(t)).\end{aligned}\quad (41)$$

We observe that the energy conservation law [Eq. (21)] holds as long as the learning dynamics satisfies the relations $x^i(t) = \nabla h_i^*(y^i(t))$. From this observation, one can confirm that the FTRL Hamiltonian remains conserved along the solutions to the modified FTRL dynamics [Eq. (41)] under the perturbation relations,

$$f^i(x, y) = \text{Hess}(h_i^*(y^i)) g^i(x, y). \quad (42)$$

For generic perturbations f and g that do not satisfy Eq. (42), the solution is not guaranteed to take the form $x^i(t) = \nabla h_i^*(y^i(t))$, and consequently the Hamiltonian is not necessarily conserved. From the viewpoint of Hamiltonian conservation, there is freedom in the choice of g (or f). Leveraging this freedom, we can deform the dynamics while preserving the Hamiltonian, i.e., the zero-sum nature of the game. This property is a significant consequence of the FTRL Hamiltonian (18); in general Hamiltonian systems, adding a perturbation such as Eq. (41) would typically break conservation of the Hamiltonian.

Theorem 2. Suppose we have a game Γ_{-1} . If a fully mixed Nash equilibrium exists and the elements of $\nabla h_j^*(y^j)$ sum to a constant independent of j , then the modified FTRL dynamics [Eq. (41)] together with the following perturbations for all $i \in \mathcal{N}$ under the relations [Eq. (42)] lead to convergence to the Nash equilibrium:

$$g^i(x, y) = \sum_{j \in \mathcal{N}} U^{(ij)} \text{Hess}(h_j^*(y^j)) \sum_{k \in \mathcal{N}} U^{(jk)} x^k. \quad (43)$$

Proof. We show that the Fenchel coupling G_F is monotonically decreasing under the modified FTRL dynamics [Eq. (41)] with Eqs. (42) and (43). Using the assumptions, the time

evolution of the Fenchel coupling turns out to be

$$\begin{aligned}\frac{d}{dt} G_F(x(t), y(t)) &= \alpha \sum_{i \in \mathcal{N}} \langle g^i(x(t), y(t)), x^i(t) - x^{i*} \rangle \\ &= -\alpha \sum_j \left\langle \text{Hess}(h_j^*(y^j(t))) \right. \\ &\quad \times \sum_k U^{(jk)} (x^k(t) - x^{k*}), \\ &\quad \left. \times \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle.\end{aligned}\quad (44)$$

Note that $x(t)$ and $y(t)$ here are the solutions to Eq. (41). Since $\alpha \geq 0$ and the Hessian of the dual regularizers h_j^* is positive definite, we obtain

$$\frac{d}{dt} G_F(x(t), y(t)) \leq 0. \quad (45)$$

The equality is reached by $\alpha = 0$, which reduces to the ordinary FTRL dynamics. Thus, in the modified FTRL dynamics for $\alpha > 0$, we have the convergent solutions to the Nash equilibrium. ■

We illustrate the complete proof in Appendix D 4. It is noted that, from the relations (42), the learning dynamics turns out to satisfy $x^i(t) = \nabla h_i^*(y^i(t))$. Due to this dissipative property of the Fenchel coupling G_F , we refer to the proposed modified FTRL dynamics [Eq. (41)] with the relations [(42) and (43)] as the *dissipation FTRL* dynamics. The DFTRL dynamics generically converges to the Nash equilibrium under the perturbations [Eq. (43)], which can be understood as an effective tuning of the payoff matrices of the game:

$$\begin{aligned}x^i(t) &= \nabla h_i^*(y^i(t)), \\ \frac{dy^i(t)}{dt} &= \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j(t)) + \alpha g^i(x(t), y(t)) \\ &= \sum_{j \in \mathcal{N}} (U^{(ij)} + \alpha \sum_{k \in \mathcal{N}} U^{(ik)} \text{Hess}(h_k^*(y^k(t))) U^{(kj)}) x^j(t).\end{aligned}\quad (46)$$

The algorithm of DFTRL is interpreted as follows. The agents play a game Γ_{-1} with the payoff matrices satisfying the *local* zero-sum property, $(U^{(ij)})^\top = -U^{(ji)}$. Meanwhile, a “game master” secretly tunes the actual payoffs by Eq. (43) to converge the learning dynamics while maintaining the *global* zero-sum property [Eq. (21)], namely, the conservation of the total utility. We will demonstrate in numerical simulations in the next section that the DFTRL dynamics really converges to the Nash equilibrium. To close this section, we would like to mention the relation between our DFTRL algorithm and the known last-iterate convergent FTRL algorithms.

Proposition 4. For small α , the DFTRL dynamics is equivalent to the first-order expansion of the continuous optimistic [18–22] and the continuous extra-gradient (CEG) [18,23–25] FTRL dynamics.

The proof is provided in Appendix E. In the literature, the optimistic and extra-gradient FTRL dynamics are known to converge to the Nash equilibrium. Since the perturbation coefficient is set to be small enough empirically, this proposition

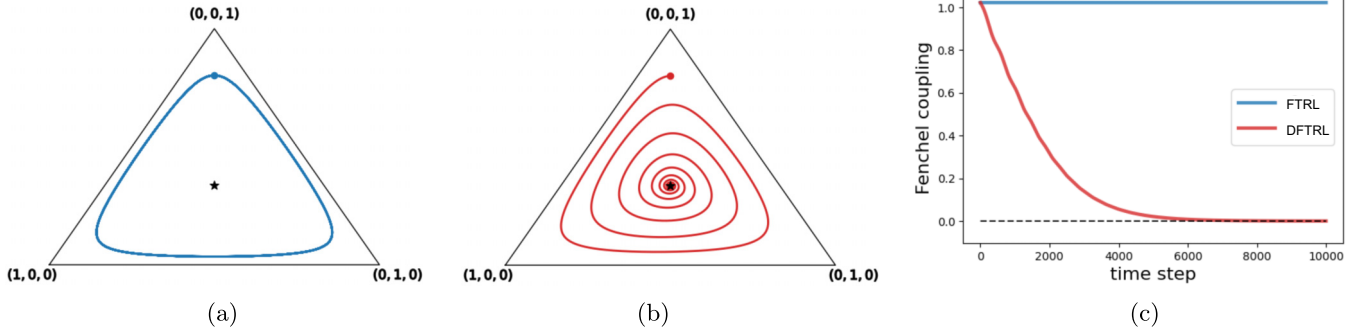


FIG. 2. Two-player Rock-Paper-Scissors. The stars at the center of the simplices represent the Nash equilibrium. (a) FTRL ($\alpha = 0$), (b) DFTRL ($\alpha = 0.15$) and (c) Fenchel coupling G_F .

suggests that our DFTRL algorithm provides a unified and in-depth understanding of the mechanism behind the two. We would like to note that the continuous negative-momentum [26–30] FTRL dynamics reduces to the ordinary FTRL dynamics, rather than the DFTRL, which we also verify in Appendix E.

IV. EXPERIMENTS

We present two experimental simulations that demonstrate our theoretical results [45]. We illustrate typical examples in this section. Further details and additional simulations are provided in Appendix G.

The first example is the two-player Rock-Paper-Scissors game. We pick the entropic regularizer for $i = 1$ and 2:

$$h_i(x^i) = \sum_a x_a^i \ln x_a^i, \quad h_i^*(y^i) = \text{lse}(y^i). \quad (47)$$

We solve the dynamical system equation (41) under Eqs. (42) and (43), with the initial condition $x^1(0) = x^2(0) = (0.1, 0.1, 0.8)$. Figure 2 shows the result. Figures 2(a) and 2(b) represent the solution trajectories of $x^1(t)$ for $\alpha = 0$ and 0.15, respectively. The stars at the center of the simplices represent the Nash equilibrium: $x^{i*} = (1/3, 1/3, 1/3)$. As mentioned in the previous section, $\alpha = 0$ reduces the system to the ordinary FTRL dynamics, and we see the cyclic solution in Fig. 2(a), as expected. Figure 2(c) shows that the Fenchel coupling G_F for DFTRL monotonically decreases, while it is conserved for FTRL, as proven in Proposition 2 and Theorem 2.

The second case is the three-player Matching Pennies game. We solve the dynamical system as in the previous example with the entropic regularizer, and the result is shown in Fig. 1. These figures depict the time evolution of the tuple $(x_1^1(t), x_1^2(t), x_1^3(t))$ in the unit cube. This game has the continuum of Nash equilibria,

$$(x_1^{1*}, x_1^{2*}, x_1^{3*}) = (p, p, p), \quad 0 \leq p \leq 1, \quad (48)$$

which the diagonal straight lines represent in Fig. 1. The solution trajectories in Figs. 1(a) and 1(b) correspond to the cases of $\alpha = 0$ and 0.1, respectively. Again, we observe that the solutions for $\alpha = 0$ exhibit cyclic behavior, and for $\alpha = 0.1$ monotonically converge to the Nash equilibrium.

V. DISCUSSION

This paper discussed a unique formulation and algorithms for the learning dynamics in games. We established the Hamiltonian of polymatrix zero-sum games and showed that the FTRL dynamics is indeed the Hamiltonian system of the proposed H_{FTRL} [Eq. (18)] in Theorem 1. We examined the Hamiltonian system through the lens of symmetry and found that the conserved quantities in the FTRL dynamics are intrinsically encoded in H_{FTRL} (Propositions 1–3). The Fenchel coupling G_F , one such conserved quantity, measures the distance to the Nash equilibrium, the conservation of which leads to the nonconvergent learning dynamics. Considering the perturbation that monotonically decreases G_F but preserves the conservation law of H_{FTRL} , we proposed the DFTRL algorithm in Theorem 2, which yields convergent dynamics to the Nash equilibrium. We also clarified the relation between DFTRL and the known last-iterate convergent algorithms in Proposition 4.

A limitation of this study is that we considered only fully mixed strategies. This is simply to ensure compatibility with the conservation laws. Including a proper treatment on the boundary of the strategy space $\mathcal{X}$ would provide a more complete description of the formulation. Beyond this, our work opens many avenues for future research. At first, different types of perturbations may exist that yield convergent dynamics, or there might be a uniqueness theorem for our perturbation. In fact, under certain conditions, the known last-iterate convergent FTRL algorithms reduce to the DFTRL as demonstrated in Proposition 4. Such considerations would further deepen our understanding of convergence in learning in games. Next, since our Hamiltonian H_{FTRL} reproduces well-known conserved quantities and even generates unique ones (see also Appendix F 1), other properties characterizing polymatrix zero-sum games may also be derived from H_{FTRL} . The symmetries (of agents or actions) have been of interest in game theory for a long time [46,47]. In connection with these symmetries, our Hamiltonian could further derive the corresponding conservation laws and potentially provide insights into solving games. Third, the Hamiltonian H_{FTRL} could serve as a promising starting point for the quantum version of the FTRL dynamics in polymatrix zero-sum games. As we have the canonically conjugate variables satisfying $\{x_a^i, y_b^j\}_{\text{P}} = \delta^{ij} \delta_{ab}$, we can perform the canonical quantization by replacing the variables with the operators

satisfying the canonical commutation relations, $[\hat{x}_a^i, \hat{y}_b^j] = i\hbar \delta^{ij} \delta_{ab}$. In analogy with the standard quantum mechanics context, we can naturally define the polymatrix zero-sum *quantum* game through $H_{\text{FTRL}}(\hat{x}, \hat{y})$ and compare this system with existing works [48–52]. Finally, the conditions of polymatrix and zero sum could be further relaxed by broadening the scope of our research. In the study of such situations, the key would be to establish an appropriate Hamiltonian and to explore the symmetry thereof. We believe that this paper not only reorganizes the FTRL dynamics, but also provides a compass for establishing emerging research fields.

All the experiments in this paper are conducted on an Intel(R) Xeon(R) CPU @ 2.20 GHz. The operating system is Ubuntu 22.04.4 LTS. We use the `torchdiffeq` framework [53,54] for solving the differential equations. Running the code for each simulation takes at most a few minutes.

DATA AVAILABILITY

The data that support the findings of this article are openly available [45].

APPENDIX A: PROPERTIES AND EXAMPLES OF FTRL DYNAMICS

We here summarize some known properties and examples of learning in games relevant to our discussions in the main text. To begin with, we recap the Karush-Kuhn-Tucker (KKT) condition, also known as the generalized method of Lagrange multipliers. The notations overlap somewhat with those in the main text, but we believe there is no confusion.

Let M be an n -dimensional smooth manifold and

$$f : M \rightarrow \mathbb{R}, \quad g : M \rightarrow \mathbb{R}^m, \quad h : M \rightarrow \mathbb{R}^l \quad (\text{A1})$$

are assumed to be in C^1 . Let us consider a nonlinear optimization problem,

$$\min_{p \in M} f(p), \quad (\text{A2})$$

with the constraints

$$g(p) \leq 0, \quad h(p) = 0. \quad (\text{A3})$$

The Karush-Kuhn-Tucker condition [55,56]. Suppose that $p^* \in M$ is a local minimizer of the above constrained optimization problem, and that the Guignard constraint qualification holds at p^* . Then, there exist (unique) Lagrange multipliers

$$\mu \in \mathbb{R}^m, \quad \lambda \in \mathbb{R}^l, \quad (\text{A4})$$

such that at p^* the following conditions hold:

$$df + \mu^\top dg + \lambda^\top dh = 0, \quad (\text{A5})$$

$$h(p^*) = 0, \quad (\text{A6})$$

$$\mu \geq 0, \quad g(p^*) \leq 0, \quad \mu^\top g(p^*) = 0. \quad (\text{A7})$$

1. Properties of Nash equilibrium

As introduced in the main text, the Nash equilibrium is defined as follows:

$$x^* = (x^{1*}, \dots, x^{n*}) \in \mathcal{X} \quad (\text{A8})$$

is a Nash equilibrium if it satisfies

$$\forall i \in \mathcal{N}, \quad x^{i*} \in \arg \max_{x^i \in \mathcal{X}^{N_i}} u_i(x^i; (x^{j(\neq i)})^*). \quad (\text{A9})$$

This means that the Nash equilibrium is a solution to the optimization problems of

$$f(x^i) = -u_i(x^i; (x^{j(\neq i)})^*) = -(x^i)^\top \sum_{j \in \mathcal{N}} U^{(ij)} x^{j*}, \quad (\text{A10})$$

over $M_i = \mathbb{R}^{N_i+1}$ with the constraints

$$g(x^i) = -(x_1^i, \dots, x_{N_i+1}^i) \leq 0, \quad (\text{A11})$$

$$h(x^i) = \sum_{a \in A_i} x_a^i - 1 = 0. \quad (\text{A12})$$

Thus, from the KKT condition, at the Nash equilibrium x^{i*} there exist Lagrange multipliers $\mu^i \in \mathbb{R}^{N_i+1}$ and $\lambda^i \in \mathbb{R}$, such that

$$\sum_a \left(- \left(\sum_j U^{(ij)} x^{j*} \right)_a - \mu_a^i + \lambda^i \right) dx_a^i = 0, \quad (\text{A13})$$

$$\sum_a x_a^{i*} = 1, \quad (\text{A14})$$

$$\mu_a^i \geq 0, \quad x_a^{i*} \geq 0, \quad \sum_a \mu_a^i x_a^{i*} = 0. \quad (\text{A15})$$

From the last condition, we find that for each $a \in A_i$, either $\mu_a^i = 0$ or $x_a^{i*} = 0$ holds. If $x_a^{i*} > 0$ for all $a \in A_i$, then $\mu^i = 0$ and the KKT condition reduces to the ordinary method of Lagrange multipliers:

$$\sum_j U^{(ij)} x^{j*} = \lambda^i \mathbf{1}, \quad (\text{A16})$$

with the constraint $\sum_a x_a^{i*} = 1$. $\mathbf{1} := (1, \dots, 1)$ is the all-one vector. Otherwise, if $x_{a'}^{i*} = 0$ for some a' , then $\mu_{a'}^i \geq 0$ and the condition becomes

$$\left(\sum_j U^{(ij)} x^{j*} \right)_{a'} = \lambda^i - \mu_{a'}^i \leq \lambda^i, \quad (\text{A17})$$

with the constraint $\sum_a x_a^{i*} = 1$.

For a zero-sum game, the total utility for all the agents always vanishes due to the antisymmetry of the payoff matrices:

$$\begin{aligned} \sum_{i \in \mathcal{N}} u_i(x^i; x^{j(\neq i)}) &= \sum_{i,j} (x^i)^\top U^{(ij)} x^j \\ &= - \sum_{i,j} (x^j)^\top U^{(ji)} x^i \\ &= 0. \end{aligned} \quad (\text{A18})$$

This fact implies that the constants λ^i in the aforementioned property for the Nash equilibrium satisfy the corresponding

condition,

$$\begin{aligned}
 0 &= \sum_{i \in \mathcal{N}} u_i(x^{i*}, (x^{j(\neq i)})^*) \\
 &= \sum_{i \in \mathcal{N}} (x^{i*})^\top \sum_j U^{(ij)} x^{j*} \\
 &= \sum_{i \in \mathcal{N}} (x^{i*})^\top \lambda^i \mathbf{1} \\
 &= \sum_{i \in \mathcal{N}} \lambda^i \sum_a x_a^{i*} \\
 &= \sum_{i \in \mathcal{N}} \lambda^i, \tag{A19}
 \end{aligned}$$

where we assumed x^* is fully mixed. Overall, the fully mixed Nash equilibrium for a zero-sum game has the property that for $i = 1, \dots, n$, there exists a set of constants $\lambda^i \in \mathbb{R}$ that satisfies

$$\sum_{j \in \mathcal{N}} U^{(ij)} x^{j*} = \lambda^i \mathbf{1}, \quad \sum_{i \in \mathcal{N}} \lambda^i = 0. \tag{A20}$$

Next, we provide two examples of the FTRL dynamics. To describe them, we solve the right-hand side of Eq. (13) or (14), a nonlinear optimization problem. By assuming that strategies are always fully mixed, as in the Nash equilibrium properties above, this optimization problem is solved by the reduced KKT condition, namely, the ordinary method of Lagrange multipliers. Suppose ζ^i solve the optimization problems of

$$f(x^i) = -(\langle x^i, y^i \rangle - h_i(x^i)), \tag{A21}$$

over $M_i = \mathbb{R}^{N_i+1}$ with the constraint $h(x^i) = \sum_a x_a^i - 1 = 0$, for $i = 1, \dots, n$. Then, there exist Lagrange multipliers $v^i \in \mathbb{R}$ that satisfy

$$\sum_a (-y_a^i + \nabla h_i(\zeta^i)_a + v^i) dx_a^i = 0, \tag{A22}$$

$$\sum_a \zeta_a^i = 1. \tag{A23}$$

By definition, $\zeta^i = \arg \max_{z \in \mathcal{X}^{N_i}} (\langle z, y^i \rangle - h_i(z))$. In many cases, the constants v^i are determined by the second condition on ζ^i .

2. Entropic regularizer and the replicator dynamics

One of the most widely known examples of learning dynamics is so-called the *replicator dynamics* [36–38], which is given by the following entropic regularizer function:

$$h_i(x^i) = \sum_a x_a^i \ln x_a^i. \tag{A24}$$

For this regularizer function, Eq. (A22) becomes

$$\zeta_a^i = \exp(y_a^i - v^i - 1), \tag{A25}$$

for all $a \in A_i$. By summing over a , the constraint for ζ^i gives

$$e^{v^i+1} = \sum_a \exp(y_a^i). \tag{A26}$$

Thus, we find

$$\zeta_a^i = \frac{\exp(y_a^i)}{\sum_{a'} \exp(y_{a'}^i)}. \tag{A27}$$

By substituting this expression to the right-hand side of Eq. (14), we obtain

$$h_i^*(y^i) = \text{lse}(y^i) := \ln \sum_a \exp(y_a^i) \tag{A28}$$

and

$$\nabla h_i^*(y^i)_a = \frac{\exp(y_a^i)}{\sum_{a'} \exp(y_{a'}^i)} =: \text{softmax}(y^i)_a, \tag{A29}$$

$$\text{Hess}(h_i^*(y^i))_{ab} = \delta_{ab} \text{softmax}(y^i)_a - \text{softmax}(y^i)_a \text{softmax}(y^i)_b. \tag{A30}$$

With these at hand, we find the learning rule,

$$\begin{aligned}
 \frac{dx^i(t)}{dt} &= \text{Hess}(h_i^*(y^i(t))) \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) \\
 &= x^i(t) \odot \left(\sum_j U^{(ij)} x^j(t) - \left\langle x^i(t), \sum_j U^{(ij)} x^j(t) \right\rangle \mathbf{1} \right), \tag{A31}
 \end{aligned}$$

where $\odot$ is the elementwise product, also known as the Hadamard product, and we used the relation $x^i(t) = \nabla h_i^*(y^i(t)) = \text{softmax}(y^i(t))$.

3. Euclidean regularizer and the projection dynamics

Another widely used example of learning dynamics is given by the Euclidean quadratic regularizer:

$$h_i(x^i) = \frac{1}{2} \|x^i\|_2^2. \tag{A32}$$

Similarly to the entropic regularizer case, the optimization problem is solved by Eq. (A22) and the constraint as

$$\zeta_a^i = y_a^i - \frac{1}{|A_i|} \left(\sum_{a'} y_{a'}^i - 1 \right). \tag{A33}$$

We now have

$$h_i^*(y^i) = \frac{1}{2} \|y^i\|_2^2 - \frac{1}{2} \|c^i\|_2^2, \tag{A34}$$

where $c^i \in \mathbb{R}^{|A_i|}$ are identical-entry vectors:

$$c^i = \frac{1}{|A_i|} \left(\sum_{a'} y_{a'}^i - 1 \right) \mathbf{1}. \tag{A35}$$

The gradients and the Hessians are

$$\nabla h_i^*(y^i)_a = y_a^i - \frac{1}{|A_i|} \left(\sum_{a'} y_{a'}^i - 1 \right), \quad (\text{A36})$$

$$\text{Hess}(h_i^*(y^i))_{ab} = \delta_{ab} - \frac{1}{|A_i|}, \quad (\text{A37})$$

which lead to the projected reinforcement learning process. The agents' mixed strategies are then known to follow the *projection dynamics* [39–43]:

$$\begin{aligned} \frac{dx^i(t)}{dt} &= \text{Hess}(h_i^*(y^i(t))) \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) \\ &= \sum_j U^{(ij)} x^j(t) - \frac{1}{|A_i|} \sum_a \left(\sum_j U^{(ij)} x^j(t) \right)_a \mathbf{1}. \end{aligned} \quad (\text{A38})$$

APPENDIX B: LIGHTNING REVIEW OF HAMILTONIAN DYNAMICS

For accessibility to broader readers, we describe minimal facts on Hamiltonian dynamics relevant to our main discussions. Here again, some notations overlap with those in the main text, but there should be no confusion. For more details on Hamiltonian dynamics, we refer the reader to Refs. [57,58].

Let (M, ω) be a real $2n$ -dimensional symplectic manifold equipped with the symplectic form ω . A typical example is the space of “position” and “momentum,” $M = T^*\mathbb{R}^n \simeq \mathbb{R}^n \times \mathbb{R}^n \ni (q, p)$, on which the symplectic form takes the form

$$\omega = \sum_{a=1}^n dp_a \wedge dq_a \simeq \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix} =: J, \quad (\text{B1})$$

where 1_n is the $n \times n$ identity matrix. Even if the global structure of M is subtle, the symplectic form can be expressed as in the above at least locally, owing to Darboux's theorem. Due to this fact, in this Appendix we mostly consider this typical example since it suffices for our purpose.

Given a real smooth function f on M , $f \in C^\infty(M)$, the *Hamiltonian system* induced by f is defined by the following flow equation for the curve $\phi : I \rightarrow M$, accompanied by the symplectic form:

$$\frac{d\phi(t)}{dt} = J \nabla f(z) |_{z=\phi(t)}, \quad z \in M. \quad (\text{B2})$$

Specifically, using the local coordinate $z = (q, p)$ and the abuse of notation $\phi(t) = (q(t), p(t))$, the flow equation becomes

$$\begin{aligned} \frac{dq(t)}{dt} &= \nabla_p f(q, p) |_{q=q(t), p=p(t)}, \\ \frac{dp(t)}{dt} &= -\nabla_q f(q, p) |_{q=q(t), p=p(t)}. \end{aligned} \quad (\text{B3})$$

Such a flow-generating function f of a Hamiltonian system is itself called the *Hamiltonian* and is often written as H . In Hamiltonian dynamics, [59] the symplectic manifold M is referred to as the *phase space* or *state space*.

1. Energy conservation law

In Hamiltonian dynamics, the value of the flow-generating function $f(z)$ on the trajectory of the solution to the flow equation (B2) is called *energy*:

$$E : I \rightarrow \mathbb{R}, \quad E(t) := f(\phi(t)). \quad (\text{B4})$$

A well-known fact is that this energy is conserved along the time interval I , as long as the function f has no explicit dependence on I (time). In other words, the value of the flow-generating function $f(z)$ is invariant along the trajectory of the solution to the flow equation: Using again the local coordinate $z = (q, p)$,

$$\begin{aligned} \frac{dE(t)}{dt} &= \frac{d}{dt} f(q(t), p(t)) \\ &= \left\langle \nabla_q f(q, p) |_{q=q(t), p=p(t)}, \frac{dq(t)}{dt} \right\rangle \\ &\quad + \left\langle \nabla_p f(q, p) |_{q=q(t), p=p(t)}, \frac{dp(t)}{dt} \right\rangle \\ &= \left\langle -\frac{dp(t)}{dt}, \frac{dq(t)}{dt} \right\rangle + \left\langle \frac{dq(t)}{dt}, \frac{dp(t)}{dt} \right\rangle \\ &= 0, \end{aligned} \quad (\text{B5})$$

where we have used the fact that $\phi(t) = (q(t), p(t))$ satisfies the flow equation (B3) and the symmetricity of the inner product.

2. Noether's theorem in Hamiltonian dynamics

Conservation law in Hamiltonian dynamics is generalized by the so-called Noether's theorem. In a Hamiltonian system, besides the energy, we can identify conserved quantities according to the symmetry of the Hamiltonian.

Theorem B1. If a Hamiltonian function f on a symplectic manifold (M, ω) admits the one-parameter group of canonical transformations generated by a function $G \in C^\infty(M)$, then G is conserved on the Hamiltonian flow of f .

For the details, see Ref. [58]. In a word, this theorem states that if a Hamiltonian is invariant under a certain (canonical) transformation, the Hamiltonian system has an associated conserved quantity. In the physics context, we refer to such a transformation (or equivalently a group action) as the *symmetry* of the Hamiltonian system. We will illustrate two examples below for intuition. Instead of providing the proof or more details of this theorem, here we only describe a rough procedure for the construction of the conserved quantity associated with the corresponding symmetry. Let us consider a Hamiltonian system generated by $f \in C^\infty(M)$ and the following infinitesimal canonical transformation on the local coordinate (q, p) of M :

$$q'_a = q_a + \delta q_a, \quad p'_a = p_a + \delta p_a, \quad (\text{B6})$$

where

$$\delta q_a = \epsilon \{q_a, G\}_P, \quad \delta p_a = \epsilon \{p_a, G\}_P, \quad (\text{B7})$$

for $G \in C^\infty(M)$ and $\epsilon \in \mathbb{R}_{\geq 0}$. $\{\cdot, \cdot\}_P$ denotes the Poisson bracket. What Noether's theorem tells us is that if the Hamiltonian is invariant under this canonical transformation for a

certain G ,

$$f(q', p') = f(q, p), \quad (\text{B8})$$

which leads to $\{G, f\}_P = 0$, then the function G is a conserved quantity in the Hamiltonian system: Along the trajectory of the solution to the flow equation (B3),

$$\begin{aligned} \frac{d}{dt}G(q(t), p(t)) &= \left\langle \nabla_q G(q, p) \Big|_{q=q(t), p=p(t)}, \frac{dq(t)}{dt} \right\rangle \\ &\quad + \left\langle \nabla_p G(q, p) \Big|_{q=q(t), p=p(t)}, \frac{dp(t)}{dt} \right\rangle \\ &= (\langle \nabla_q G(q, p), \nabla_p f(q, p) \rangle \\ &\quad - \langle \nabla_p G(q, p), \nabla_q f(q, p) \rangle) \Big|_{q=q(t), p=p(t)} \\ &= \{G, f\}_P \\ &= 0. \end{aligned} \quad (\text{B9})$$

To provide intuition, we present two well-known examples from the physics literature. Let $M = T^*\mathbb{R}^3 \simeq \mathbb{R}^3 \times \mathbb{R}^3$ be the phase space of a single particle of mass m moving in the three-dimensional space. The Hamiltonian $H : M \rightarrow \mathbb{R}$, describing generic classical mechanical systems, takes the form

$$H(q, p) = \frac{1}{2m} \|p\|_2^2 + V(q), \quad (\text{B10})$$

where $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ is referred to as the potential function, which characterizes the system.

(1) *Momentum conservation law.* If the Hamiltonian H has a translation symmetry,

$$H(q', p') = H(q, p), \quad (\text{B11})$$

$$q' = q + e, \quad p' = p, \quad (\text{B12})$$

where e is a constant vector, then the associated conserved quantity is

$$G^P(q, p; e) = \langle e, p \rangle. \quad (\text{B13})$$

In the physical sense, this implies that if the potential function is invariant under a certain constant shift of the position coordinate, the corresponding momentum of the particle is conserved in the Hamiltonian system.

(2) *Angular momentum conservation law.* If the Hamiltonian H has a rotation symmetry,

$$H(q', p') = H(q, p), \quad (\text{B14})$$

$$q' = Rq, \quad p' = Rp, \quad (\text{B15})$$

where $R \in \text{SO}(3)$, then the associated conserved quantity is the angular momentum along R . If the invariance holds for any rotation, $\forall R \in \text{SO}(3)$, the whole angular momentum vector of the particle is conserved in the Hamiltonian system,

$$G_a^L(q, p) = (q \times p)_a, \quad a = 1, \dots, 3, \quad (\text{B16})$$

where $\times$ is the three-dimensional exterior product.

As these examples demonstrate, the examination of symmetry of the Hamiltonian has a direct connection to the conserved quantity of the system. Conserved quantities provide us an enormous amount of information of the system under consideration, as a sufficient number of conserved

quantities even lead to the solvability of the system. Due to this, in the study of a dynamical system, it is essential in understanding the system to establish an appropriate Hamiltonian function and to elucidate the symmetry thereof.

APPENDIX C: COMMENTS ON THE PAPER BY BAILEY AND PILIOURAS

In this Appendix, we recap the main arguments in the paper by Bailey and Piliouras [6], emphasizing that their use of the term ‘‘Hamiltonian’’ deviates from the standard convention, explained in Appendix B. It turns out that their central claim essentially amounts to the tautological statement that ‘‘the solution to the FTRL dynamics satisfies the FTRL dynamics’’. As a result, their analysis ultimately reduces to a study of the FTRL dynamics and previously known conserved quantities.

With the notations in Sec. II, suppose that a game Γ_σ is given with a set of regularizer functions $h_i \in C_{\text{SC}}^\infty(\mathcal{X}^{N_i}, \mathbb{R})$. One defines the ‘‘cumulative strategies’’ by

$$X^i(t) := \int_0^t x^i(s) ds. \quad (\text{C1})$$

Then, for the choice of h_i , the FTRL dynamics describes the time evolution of the learning dynamics $(X(t), y(t))$ in Γ_σ by

$$\begin{aligned} \frac{dX^i(t)}{dt} &= \nabla h_i^*(y^i) \Big|_{y^i=y^i(t)}, \\ \frac{dy^i(t)}{dt} &= \sum_{j \neq i} U^{(ij)} x^j(t), \end{aligned} \quad (\text{C2})$$

with the relations $\dot{X}^i(t) = \dot{X}^i(t) = \nabla h_i^*(y^i(t))$. One introduces the ‘‘free Hamiltonian’’ for the agents by

$$H_0(X, y) = \sum_{i \in \mathcal{N}} h_i^*(y^i), \quad (\text{C3})$$

and the ‘‘interaction Hamiltonian,’’

$$H_{\text{int}}^\sigma(X, y) = -\sigma \sum_{i \in \mathcal{N}} h_i^* \left(\beta^i + \sum_{k \neq i} U^{(ik)} X^k \right), \quad (\text{C4})$$

where $X^i, y^i \in \mathbb{R}^{|A_i|}$, and $\beta^i \in \mathbb{R}^{|A_i|}$ are some constants. The total ‘‘Hamiltonian’’ of the system is a function on $\mathbb{R}^{\sum_i |A_i|} \times \mathbb{R}^{\sum_i |A_i|}$ defined by their sum,

$$\begin{aligned} H^\sigma(X, y) &= H_0(X, y) + H_{\text{int}}^\sigma(X, y) \\ &= \sum_{i \in \mathcal{N}} \left[h_i^*(y^i) - \sigma h_i^* \left(\beta^i + \sum_{k \neq i} U^{(ik)} X^k \right) \right] \\ &= \sum_{i \in \mathcal{N}} h_i^*(y^i) - \sigma \sum_{j \in \mathcal{N}} h_j^* \left(\beta^j + \sum_{i \neq j} U^{(ji)} X^i \right). \end{aligned} \quad (\text{C5})$$

Theorem C1. The function H^σ [Eq. (C5)], with $\sigma = -1$ and $\beta^i = y^i(0)$ for all $i \in \mathcal{N}$, is invariant along the solution trajectory of the FTRL dynamics [Eq. (C2)].

Using Eq. (C2) and the assumption, the invariance of $H^{\sigma=-1}$ along the solution trajectory is readily proven as

$$\begin{aligned} \frac{d}{dt} H^{\sigma=-1}(X(t), y(t)) &= \sum_{i \in \mathcal{N}} \left[\left\langle \nabla h_i^*(y^i) \mid_{y^i=y^i(t)}, \frac{dy^i(t)}{dt} \right\rangle \right. \\ &\quad \left. + \left\langle \nabla h_i^*(z^i) \mid_{z^i=y^i(0)+\sum_k U^{(ik)} X^k(t)}, \sum_j U^{(ij)} \frac{dX^j(t)}{dt} \right\rangle \right] \\ &= 2 \sum_{i \in \mathcal{N}} \left\langle x^i(t), \sum_j U^{(ij)} x^j(t) \right\rangle \\ &= 0, \end{aligned} \quad (\text{C6})$$

where at the last line we have used the zero-sum property of the game [Eq. (A18)]. From this theorem, we recognize that the function $H^{\sigma=-1}$ is a conserved quantity for polymatrix zero-sum games along the solution trajectories of the FTRL dynamics. However, the function $H^{\sigma=-1}$ is naively not regarded as a Hamiltonian function for the FTRL dynamics. Let $M' = T^* \mathbb{R}^{\sum_i |A_i|} \simeq \mathbb{R}^{\sum_i |A_i|} \times \mathbb{R}^{\sum_i |A_i|}$ be a phase space and (X, y) be the (global) canonical coordinate of M' . Suppose we have a function $H^{\sigma=-1} : M' \rightarrow \mathbb{R}$ given as in Eq. (C5). Then, the Hamiltonian system induced by $H^{\sigma=-1}$ is described by the flow equations (B3) as

$$\begin{aligned} \frac{dX^k(t)}{dt} &= \nabla_{y^k} H^{\sigma=-1}(X, y) \mid_{X=X(t), y=y(t)} \\ &= \nabla h_k^*(y^k) \mid_{y^k=y^k(t)}, \end{aligned} \quad (\text{C7})$$

$$\begin{aligned} \frac{dy^k(t)}{dt} &= -\nabla_{X^k} H^{\sigma=-1}(X, y) \mid_{X=X(t), y=y(t)} \\ &= \sum_j U^{(kj)} \nabla h_j^*(y^j) \mid_{y^j=y^j(t) + \sum_l U^{(jl)} X^l(t)}. \end{aligned} \quad (\text{C8})$$

As can be seen from the right-hand side of Eq. (C8), this Hamiltonian system is generally not equivalent to the FTRL dynamics (C2). If one imposes a constraint by hand on the variables X^i and y^i such as the projection onto the submanifold

$$\tilde{M} = \{y^i = \beta^i + \sum_j U^{(ij)} X^j \mid i \in \mathcal{N}\} \subset M', \quad (\text{C9})$$

then the right-hand side becomes

$$\sum_j U^{(kj)} \nabla h_j^*(y^j) \mid_{y^j=y^j(t)} = \sum_j U^{(kj)} \frac{dX^j(t)}{dt}, \quad (\text{C10})$$

and this Hamiltonian system could be thought of as an equivalent of Eq. (C2). Nonetheless, imposing constraints on the variables, such as Eq. (C9), implies that “the FTRL dynamics is the Hamiltonian system on the manifold on which the solution is the FTRL dynamics,” which is a tautological statement. Overall, the function $H^{\sigma=-1}$ is merely a conserved quantity and not the Hamiltonian function for the FTRL dynamics on the total phase space $(M'; X, y)$.

As Bailey and Piliouras mentioned in Ref. [[6], Sec. 5], the function $H^{\sigma=-1}$ restricted to the submanifold (C9),

$$H^{\sigma=-1} \mid_{\tilde{M}}(X, y) = 2 \sum_{i \in \mathcal{N}} h_i^*(y^i), \quad (\text{C11})$$

is essentially a certain shift of the Fenchel coupling. As a by-product of the FTRL Hamiltonian (18), this result emerges in our Hamiltonian system as simply another conserved quantity through the corresponding symmetry. From the proof of Proposition 2 in Appendix D2, it follows that the FTRL Hamiltonian H_{FTRL} is also invariant under the following canonical transformation for all $i \in \mathcal{N}$, which is a variant of Eq. (27),

$$x'^i = x^i + \nabla h_i^*(y^i), \quad y'^i = y^i. \quad (\text{C12})$$

A generator of this canonical transformation, $g \in C^\infty(M_{\Gamma_{-1}})$, satisfies the following for all $i \in \mathcal{N}$ and for all $a \in A_i$,

$$\{x_a^i, g\}_{\text{P}} = \frac{\partial g(x, y)}{\partial y_a^i} = \frac{\partial h_i^*(y^i)}{\partial y_a^i}, \quad (\text{C13})$$

$$\{y_a^i, g\}_{\text{P}} = -\frac{\partial g(x, y)}{\partial x_a^i} = 0. \quad (\text{C14})$$

Integration of these equations gives us the conserved quantity in the Hamiltonian system of H_{FTRL} ,

$$g(x, y) = \sum_{i \in \mathcal{N}} h_i^*(y^i) + \text{const.}, \quad (\text{C15})$$

which is a part of the Fenchel coupling. This result is consistent with that of Ref. [[6], Sec. 5] and provides a natural explanation from the perspective of the symmetry of our Hamiltonian.

APPENDIX D: PROOFS

We show the complete proofs for Lemma 1, Proposition 2, Proposition 3, and Theorem 2 in order. In the proofs of Lemma 1, Proposition 2, and Proposition 3, we use Noether's theorem. For the details on Hamiltonian dynamics, see Appendix B.

1. Proof of Lemma 1

Proof. Let us suppose that a Hamiltonian function f on $M_{\Gamma_{-1}}$ has translation symmetry in $y^i \in \mathbb{R}^{|A_i|}$; that is, f is invariant under the following canonical transformation:

$$f(x', y') = f(x, y), \quad (\text{D1})$$

$$x' = x, \quad y'^i = y^i + \epsilon^i, \quad (\text{D2})$$

where the other $y^{j(\neq i)}$ are kept intact and ϵ^i is an identical-entry constant vector, $\epsilon^i = \epsilon^i \mathbf{1}$, $\mathbf{1} := (1, \dots, 1)$ is the all-one vector. Then, for all $a \in A_i$, a generator $g \in C^\infty(M_{\Gamma_{-1}})$ of this canonical transformation satisfies

$$\{x_a^i, g\}_{\text{P}} = \frac{\partial g(x, y)}{\partial y_a^i} = 0, \quad (\text{D3})$$

$$\{y_a^i, g\}_{\text{P}} = -\frac{\partial g(x, y)}{\partial x_a^i} = 1. \quad (\text{D4})$$

We thus obtain $g(x, y) = -\sum_a x_a^i + \text{const.}$ By Noether's theorem, the function g is conserved in the Hamiltonian system

generated by the Hamiltonian f , implying that $\sum_a x_a^i$ is conserved. ■

Lemma 1 holds for any Hamiltonian functions that possess translation symmetry in the y^i direction, not only for ours. Roughly speaking, this lemma corresponds to the conservation law of the center-of-mass momentum, which is obtained from Eq. (B13) with $e = \mathbf{1}$. It is noted, however, that the role of canonical variables is interchanged between Lemma 1 and Eq. (B13).

2. Proof of Proposition 2

Proof. The proof is straightforward by applying Noether's theorem to the Hamiltonian (18). We aim to show that our Hamiltonian (18) is invariant under the canonical transformation below, and that its generator is given by G_F . Let us consider the canonical transformation

$$x'^i = x^i + \nabla h_i^*(y^i) - x^{i*}, \quad y'^i = y^i, \quad (\text{D5})$$

for all $i \in \mathcal{N}$. Under this transformation, the FTRL Hamiltonian is invariant as

$$\begin{aligned} H_{\text{FTRL}}(x', y') - H_{\text{FTRL}}(x, y) &= \sum_{i,j} \langle \nabla h_j^*(y^j), U^{(ji)} \nabla h_i^*(y^i) - U^{(ji)} x^{i*} \rangle \\ &= - \sum_j \left\langle \nabla h_j^*(y^j), \sum_i U^{(ji)} x^{i*} \right\rangle \\ &= - \sum_j \lambda^j \sum_a \nabla h_j^*(y^j)_a \\ &= -C \sum_j \lambda^j = 0. \end{aligned} \quad (\text{D6})$$

From the first line to the second, we used the zero-sum property of the game. From the second line to the third, we apply the property of the Nash equilibrium [Eq. (A20)]. Finally, we used the assumption that the elements of $\nabla h_j^*(y^j)$ sum to a constant, and again Eq. (A20). A generator of this canonical transformation, $g \in C^\infty(M_{\Gamma-1})$, satisfies the following for all $i \in \mathcal{N}$ and for all $a \in A_i$,

$$\{x_a^i, g\}_{\text{P}} = \frac{\partial g(x, y)}{\partial y_a^i} = \frac{\partial h_i^*(y^i)}{\partial y_a^i} - x_a^{i*}, \quad (\text{D7})$$

$$\{y_a^i, g\}_{\text{P}} = -\frac{\partial g(x, y)}{\partial x_a^i} = 0. \quad (\text{D8})$$

Thus, integrating these equations, we obtain the generator of this canonical transformation,

$$g(x, y) = \sum_{i \in \mathcal{N}} (h_i^*(y^i) - \langle y^i, x^{i*} \rangle) + \text{const.} \quad (\text{D9})$$

Since $\sum_i h_i(x^{i*})$ is a constant, it follows that the Fenchel coupling G_F is the conserved quantity generating the symmetry [Eq. (D5)] of the FTRL Hamiltonian. ■

3. Proof of Proposition 3

(1) We show that the FTRL Hamiltonian (18) is invariant under the canonical transformation below, and that its

generator is given by G_K . Let us consider the canonical transformation

$$x'^i = R_i x^i, \quad y'^i = R_i y^i, \quad (\text{D10})$$

for all $i \in \mathcal{N}$. Under this transformation, the FTRL Hamiltonian is invariant:

$$\begin{aligned} H_{\text{FTRL}}(x', y') &= \sum_{i,j} \langle \nabla h_j^*(y'^j), U^{(ji)} x'^i \rangle \\ &= \sum_{i,j} \langle R_j \nabla h_j^*(y^j), U^{(ji)} R_i x^i \rangle \\ &= \sum_{i,j} \langle \nabla h_j^*(y^j), R_j^\top U^{(ji)} R_i x^i \rangle \\ &= H_{\text{FTRL}}(x, y). \end{aligned} \quad (\text{D11})$$

From the first line to the second, we used the equivariance assumption for $\nabla h_j^*(y^j)$. From the third line to the fourth, we used the facts that $R_j^\top = R_j^{-1}$ and rotation matrices and the payoff matrices commute. Since $R_i \in \text{SO}(|A_i|)$ are rotation matrices, there exist skew-symmetric matrices $K_i \in \mathfrak{so}(|A_i|)$ such that $R_i = e^{\epsilon K_i}$ and $K_i^\top = -K_i$. So, a generator of this canonical transformation, $g \in C^\infty(M_{\Gamma-1})$, satisfies

$$\{x_a^i, g\}_{\text{P}} = \frac{\partial g(x, y)}{\partial y_a^i} = \sum_b (K_i)_{ab} x_b^i, \quad (\text{D12})$$

$$\{y_a^i, g\}_{\text{P}} = -\frac{\partial g(x, y)}{\partial x_a^i} = \sum_b (K_i)_{ab} y_b^i. \quad (\text{D13})$$

Integrating these equations, we obtain the generator of this canonical transformation,

$$g(x, y) = - \sum_{i \in \mathcal{N}} \langle x^i, K_i y^i \rangle + \text{const.} \quad (\text{D14})$$

From Noether's theorem, it follows that the function G_K is the conserved quantity in the Hamiltonian system of H_{FTRL} .

(2) Similarly to the first statement, we can show the FTRL Hamiltonian (18) is invariant under the canonical transformation,

$$x' = \mathcal{R}x, \quad y' = \mathcal{R}y, \quad (\text{D15})$$

where $\mathcal{R} \in \text{SO}(n)$ acts on the index of $\mathcal{N}$ and rotates the whole spaces $\mathbb{R}_{\geq 0}^{\sum_i |A_i|}$ and $\mathbb{R}^{\sum_i |A_i|}$ for x and y , respectively. Using the notation $\langle \langle \alpha, \beta \rangle \rangle := \sum_{i \in \mathcal{N}} \langle \alpha^i, \beta^i \rangle$, we have

$$\begin{aligned} H_{\text{FTRL}}(x', y') &= \langle \langle \nabla h^*(y'), Ux' \rangle \rangle \\ &= \langle \langle \mathcal{R} \nabla h^*(y), U \mathcal{R}x \rangle \rangle \\ &= \langle \langle \nabla h^*(y), \mathcal{R}^{-1} U \mathcal{R}x \rangle \rangle \\ &= H_{\text{FTRL}}(x, y). \end{aligned} \quad (\text{D16})$$

Thus, with $\mathcal{R} = e^{\epsilon \Omega}$, $\Omega^\top = -\Omega$, we find from Noether's theorem that the function G_Ω is the conserved quantity in the Hamiltonian system of H_{FTRL} :

$$G_\Omega(x, y) = \langle \langle x, \Omega y \rangle \rangle = \sum_{i,j \in \mathcal{N}} \Omega_{ij} \langle x^i, y^j \rangle. \quad (\text{D17})$$

As mentioned in the main text, these two conservation laws correspond to the conservation of the total “angular

momentum” of each agent and the total angular momentum of the system, respectively. To see this more explicitly, one can consider the cases of $|A_i| = 3$ and/or $n = 3$;

$$K_i, \Omega = L_a, \quad (\text{D18})$$

where L_a is either of the standard basis vectors of $\mathfrak{so}(3)$:

$$L_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad L_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$L_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{D19})$$

Then, these “angular momentum” conservation laws show that the following quantities are conserved:

$$G_K(x, y) = \sum_{i \in \mathcal{N}} \langle x^i, L_a y^i \rangle = \sum_{i \in \mathcal{N}} (x^i \times y^i)_a, \quad (\text{D20})$$

$$G_\Omega(x, y) = \langle \langle x, L_a y \rangle \rangle =: \langle \langle x \times y \rangle \rangle_a, \quad (\text{D21})$$

where $\times$ is the three-dimensional exterior product.

One can indeed find an angular-momentum-like conserved quantity in the FTRL dynamics of the two-player Rock-Paper-Scissors game with the Euclidean regularizer. The two-player Rock-Paper-Scissors game has $\mathcal{N} = \{1, 2\}$, $A_i = \{\text{R}, \text{P}, \text{S}\} \simeq \{1, 2, 3\}$, and the payoff matrices are given by

$$U^{(12)} = U^{(21)} = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}. \quad (\text{D22})$$

In this case, the symmetry transformation is given by the following rotation matrices:

$$R_i = e^{\epsilon K}, \quad K = L_1 + L_2 + L_3 = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}, \quad (\text{D23})$$

which represent rotations around the $\mathbf{1} = (1, 1, 1)^\top$ axis, that is,

$$R_i \mathbf{1} = \mathbf{1}, \quad K \mathbf{1} = 0, \quad \mathbf{1}^\top K = 0. \quad (\text{D24})$$

To see that this is a symmetry, we can check that these rotation matrices commute with the payoff matrices, since $U^{(ij)} = -K$:

$$U^{(ij)} R_j = R_i U^{(ij)}. \quad (\text{D25})$$

In addition, the equivariance condition is also satisfied. A point is that for the Euclidean regularizer we have

$$\nabla h_i^*(y^i) = y^i - \frac{1}{3} \mathbf{1} (\mathbf{1}^\top y^i - 1) = P_i y^i + \frac{1}{3} \mathbf{1}, \quad (\text{D26})$$

where $P_i = I_3 - \frac{1}{3} \mathbf{1} \mathbf{1}^\top$ are the projection matrices and commute with the rotation matrices, $P_i R_i = R_i P_i$. In this form, we can readily see the equivariance condition with the

aforementioned properties for the rotation matrices as

$$\begin{aligned} R_i \nabla h_i^*(y^i) &= R_i P_i y^i + \frac{1}{3} R_i \mathbf{1} \\ &= P_i R_i y^i + \frac{1}{3} \mathbf{1} \\ &= \nabla h_i^*(R_i y^i). \end{aligned} \quad (\text{D27})$$

Thus, the FTRL Hamiltonian is invariant under this symmetry transformation, and the conserved quantity is given by

$$\begin{aligned} G_K(x, y) &= \sum_{i=1,2} \langle x^i, K y^i \rangle \\ &= \sum_{i=1,2} (x_1^i (y_2^i - y_3^i) + x_2^i (y_3^i - y_1^i) + x_3^i (y_1^i - y_2^i)). \end{aligned} \quad (\text{D28})$$

We also have the other angular-momentum-like conserved quantity in the FTRL dynamics of the two-player Matching Pennies game with the Euclidean regularizer. This game has $\mathcal{N} = \{1, 2\}$, $A_i = \{\text{H}, \text{T}\} \simeq \{1, 2\}$, and the payoff matrices are given by

$$U^{(12)} = -U^{(21)} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \quad (\text{D29})$$

One finds that this is the case of

$$U = \Sigma \otimes S, \quad \Sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \quad (\text{D30})$$

We then have the following rotation matrix as a symmetry transformation:

$$\mathcal{R} = e^{\epsilon \Sigma}, \quad \Sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (\text{D31})$$

Thus, the FTRL Hamiltonian is invariant under this symmetry transformation, and the conserved quantity is given by

$$G_\Omega(x, y) = \langle \langle x, \Sigma y \rangle \rangle, \quad (\text{D32})$$

$$= \sum_{i,j=1,2} \sigma_{ij} \langle x^i, y^j \rangle, \quad (\text{D33})$$

$$= \langle x^1, y^2 \rangle - \langle x^2, y^1 \rangle. \quad (\text{D34})$$

4. Proof of Theorem 2

Proof. From the relations (42), the learning dynamics is found to satisfy $x^i(t) = \nabla h_i^*(y^i(t))$. Therefore, it suffices to

show that the Fenchel coupling G_F is monotonically decreasing under the modified FTRL dynamics given by Eqs. (41) and (15) together with Eq. (43).

$$\frac{d}{dt} G_F(x(t), y(t)) = \sum_{i \in \mathcal{N}} \left\langle \frac{dy^i(t)}{dt}, \nabla h_i^*(y^i(t)) - x^{i*} \right\rangle = \sum_{i,j} \langle U^{(ij)} x^j(t), x^i(t) - x^{i*} \rangle + \alpha \sum_i \langle g^i(x(t), y(t)), x^i(t) - x^{i*} \rangle. \quad (\text{D35})$$

Note that $x(t)$ and $y(t)$ here are the solutions to Eq. (41). Using the assumptions, we know that the first term vanishes:

$$\sum_{i,j} \langle U^{(ij)} x^j(t), x^i(t) - x^{i*} \rangle = \sum_{i,j} \langle x^j(t), U^{(ji)} x^{i*} \rangle = \sum_j \lambda^j \sum_a x_a^j(t) = 0, \quad (\text{D36})$$

where we used the zero-sum property of the game in the first line, and the property of the Nash equilibrium [Eq. (A20)] in the last. This is nothing but the consequence of Proposition 2. Then, the time evolution of the Fenchel coupling reads as follows:

$$\begin{aligned} \frac{d}{dt} G_F(x(t), y(t)) &= \alpha \sum_{i \in \mathcal{N}} \langle g^i(x(t), y(t)), x^i(t) - x^{i*} \rangle \\ &= \alpha \sum_{i,j} \left\langle U^{(ij)} \text{Hess}(h_j^*(y^j(t))) \sum_{k \in \mathcal{N}} U^{(jk)} x^k(t), x^i(t) - x^{i*} \right\rangle \\ &= -\alpha \sum_j \left\langle \text{Hess}(h_j^*(y^j(t))) \sum_k U^{(jk)} x^k, \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle \\ &= -\alpha \sum_j \left\langle \text{Hess}(h_j^*(y^j(t))) \sum_k U^{(jk)} (x^k(t) - x^{k*}), \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle \\ &\leq 0, \end{aligned} \quad (\text{D37})$$

since $\alpha \geq 0$ and the Hessian of the dual regularizer h^* is positive definite. The equality is attained when $x(t) = x^*$ or $\alpha = 0$, which reduces to the ordinary FTRL dynamics. From the second line to the third, we used the zero-sum property, $(U^{(ij)})^\top = -U^{(ji)}$. From the third line to the fourth, we apply the assumptions and use the fact that

$$\begin{aligned} \sum_j \left\langle \text{Hess}(h_j^*(y^j(t))) \sum_k U^{(jk)} x^{k*}, \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle &= \sum_j \left\langle \text{Hess}(h_j^*(y^j(t))) \lambda^j \mathbf{1}, \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle \\ &= \sum_j \lambda^j \left\langle \nabla_{y^j} \left(\sum_a \nabla_{y^j} h_j^*(y^j(t))_a \right), \sum_i U^{(ji)} (x^i(t) - x^{i*}) \right\rangle \\ &= 0. \end{aligned} \quad (\text{D38})$$

■

APPENDIX E: CONTINUOUS FTRL ALGORITHMS AND PERTURBATIVE EXPANSION

We derive the discrete-to-continuum limit of the last-iterate convergent FTRL algorithms known in the literature. We also demonstrate interesting perturbative expansions thereof and show the proof of Proposition 4 along the way.

First, let us describe the setup for the discrete FTRL dynamics. Let $I = [0, T]$ be the time interval and $L = \{0, \epsilon, 2\epsilon, \dots, N\epsilon = T\}$ be a discrete lattice of I with the step size ϵ ;

$$L \hookrightarrow I \hookrightarrow \mathbb{R}. \quad (\text{E1})$$

The step size ϵ is by definition scaled in terms of the number of lattice points as T/N for the fixed T . We denote the discrete learning dynamics by hatted variables such as $\hat{\phi} : L \rightarrow \mathcal{X}^{N_i}$ or $\mathbb{R}^{|A_i|}$. We assume that there exists a continuous learning dynamics ϕ , to which $\hat{\phi}$ converges at large N ($\epsilon \rightarrow 0$ limit) in an appropriate sense.

1. Optimistic FTRL

The optimistic FTRL algorithm [18–22] is defined by the following update rule [60]:

$$\begin{aligned} \hat{x}^i(t) &= \nabla h_i^*(y^i) \big|_{y^i = \hat{y}^i(t)}, \\ \hat{y}^i(t + \epsilon) &= \hat{y}^i(t) + \epsilon \sum_{j \in \mathcal{N}} U^{(ij)} \hat{x}^j(t) \\ &\quad + \alpha \sum_{j \in \mathcal{N}} U^{(ij)} (\hat{x}^j(t) - \hat{x}^j(t - \epsilon)), \end{aligned} \quad (\text{E2})$$

where we assume $\alpha \in \mathbb{R}_{\geq 0}$. By taking the step size $\epsilon \rightarrow 0$ while keeping the combination $N\epsilon = T$, i.e., $n\epsilon = t$, we obtain the CO FTRL dynamics,

$$\begin{aligned} x^i(t) &= \nabla h_i^*(y^i) \big|_{y^i = y^i(t)}, \\ \frac{dy^i(t)}{dt} &= \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) + \alpha \sum_{j \in \mathcal{N}} U^{(ij)} \frac{dx^j(t)}{dt}. \end{aligned} \quad (\text{E3})$$

For $\alpha \ll 1$ [61], we consider the perturbative expansion of $y^i(t)$ in powers of α ,

$$y^i(t) = \sum_{n=0}^{\infty} \alpha^n y_{(n)}^i(t), \quad (\text{E4})$$

while $x^i(t)$ are kept untouched. Substituting this expression into the CO FTRL dynamics [Eq. (E3)], we obtain (assuming the termwise differentiability)

$$x^i(t) = \nabla h_i^*(y^i(t)), \quad (\text{E5})$$

$$\begin{aligned} \frac{dy^i(t)}{dt} &= \sum_{n=0}^{\infty} \alpha^n \frac{dy_{(n)}^i(t)}{dt} \\ &= \sum_{n=0}^{\infty} \alpha^n U^{(ij_1)} H_{j_1} U^{(j_1 j_2)} \dots H_{j_n} U^{(j_n j_{n+1})} x^{j_{n+1}}(t), \end{aligned} \quad (\text{E6})$$

where $H_{j_l} := \text{Hess}(h_{j_l}^*(y^{j_l}))$ and the repeated indices are summed over all possible values. At order α^1 , we have

$$\begin{aligned} x^i(t) &= \nabla h_i^*(y^i(t)), \\ \frac{dy^i(t)}{dt} &\simeq \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t) \\ &\quad + \alpha \sum_{j,k} U^{(ij)} \text{Hess}(h_j^*(y^j(t))) U^{(jk)} x^k(t), \end{aligned} \quad (\text{E7})$$

which is equivalent to our DFTRL algorithm [Eq. (46)]. This proves half of Proposition 4.

2. Extra-gradient FTRL

The extra-gradient FTRL algorithm [18,23–25] is defined through

$$\begin{aligned} \hat{x}^i(t) &= \nabla h_i^*(y^i) \big|_{y^i=\hat{y}^i(t)}, \\ \hat{y}^i(t+\epsilon) &= \hat{y}^i(t) + \epsilon \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j) \bigg|_{y^j=\hat{y}^j(t)+\alpha \sum_k U^{(jk)} \hat{x}^k(t)}. \end{aligned} \quad (\text{E8})$$

In the literature, this update rule is often expressed using auxiliary variables $\hat{b}^i$ and $\hat{c}^i$ as follows:

$$\hat{x}^i(t) = \nabla h_i^*(y^i) \big|_{y^i=\hat{y}^i(t)}, \quad (\text{E9})$$

$$\hat{b}^i(t) = \hat{y}^i(t) + \alpha \sum_{j \in \mathcal{N}} U^{(ij)} \hat{x}^j(t), \quad (\text{E10})$$

$$\hat{c}^i(t) = \nabla h_i^*(y^i) \big|_{y^i=\hat{b}^i(t)}, \quad (\text{E11})$$

$$\hat{y}^i(t+\epsilon) = \hat{y}^i(t) + \epsilon \sum_{j \in \mathcal{N}} U^{(ij)} \hat{c}^j(t). \quad (\text{E12})$$

By eliminating the auxiliary variables $\hat{b}^i$ and $\hat{c}^i$, one finds that these equations reduce to the aforementioned extra-gradient FTRL update rule [Eq. (E8)]. In the above four equations, the second and fourth equations imply that during one time step ϵ from $\hat{y}^i(t)$ to $\hat{y}^i(t+\epsilon)$, it acquires an “extra gradient” due to the additional term proportional to α :

$$\alpha \sum_{j \in \mathcal{N}} U^{(ij)} \hat{x}^j(t). \quad (\text{E13})$$

Now, by taking the step size $\epsilon \rightarrow 0$, we derive the CEG FTRL dynamics from Eq. (E8),

$$\begin{aligned} x^i(t) &= \nabla h_i^*(y^i) \big|_{y^i=y^i(t)}, \\ \frac{dy^i(t)}{dt} &= \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j) \bigg|_{y^j=y^j(t)+\alpha \sum_k U^{(jk)} x^k(t)}. \end{aligned} \quad (\text{E14})$$

For $\alpha \ll 1$, we again consider the perturbative expansion of $y^i(t)$ in powers of α . It should be noted that for a multivariable function f , its Taylor expansion is given by

$$f(x+c) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{a_1, \dots, a_n} \left(\frac{\partial^n}{\partial x_{a_1} \dots \partial x_{a_n}} f(x) \right) c_{a_1} \dots c_{a_n}. \quad (\text{E15})$$

Then, we find that

$$\begin{aligned} \frac{dy_a^i(t)}{dt} &= \sum_j \sum_{a'} U_{aa'}^{(ij)} \sum_{n=0}^{\infty} \frac{1}{n!} \\ &\quad \times \sum_{a_1, \dots, a_n} \left(\frac{\partial^{n+1}}{\partial y_{a_1}^j \dots \partial y_{a_n}^j \partial y_{a'}^j} h_j^*(y^j(t)) \right) \\ &\quad \times \left(\alpha \sum_k U^{(jk)} x^k(t) \right)_{a_1} \dots \left(\alpha \sum_k U^{(jk)} x^k(t) \right)_{a_n}. \end{aligned} \quad (\text{E16})$$

Thus, at order α^1 , the CEG FTRL dynamics [Eq. (E14)] becomes

$$\begin{aligned} x^i(t) &= \nabla h_i^*(y^i(t)), \\ \frac{dy^i(t)}{dt} &\simeq \sum_{j \in \mathcal{N}} U^{(ij)} \nabla h_j^*(y^j(t)) \\ &\quad + \alpha \sum_{j,k} U^{(ij)} \text{Hess}(h_j^*(y^j(t))) U^{(jk)} x^k(t), \end{aligned} \quad (\text{E17})$$

which is equivalent to our DFTRL algorithm [Eq. (46)]. This proves the other half of Proposition 4.

3. Negative-momentum FTRL

The negative-momentum FTRL algorithm [26–30] is defined by the following update rule:

$$\begin{aligned} \hat{x}^i(t) &= \nabla h_i^*(y^i) \big|_{y^i=\hat{y}^i(t)}, \\ \hat{y}^i(t+\epsilon) &= \hat{y}^i(t) + \epsilon \sum_{j \in \mathcal{N}} U^{(ij)} \hat{x}^j(t) - \alpha(\hat{y}^i(t) - \hat{y}^i(t-\epsilon)). \end{aligned} \quad (\text{E18})$$

By taking the step size $\epsilon \rightarrow 0$, we obtain the continuous negative-momentum FTRL dynamics,

$$\begin{aligned} x^i(t) &= \nabla h_i^*(y^i) \big|_{y^i=y^i(t)}, \\ \frac{dy^i(t)}{dt} &= (1+\alpha)^{-1} \sum_{j \in \mathcal{N}} U^{(ij)} x^j(t). \end{aligned} \quad (\text{E19})$$

This dynamics is essentially the same as the ordinary FTRL dynamics, since an overall constant to y^i does not affect the qualitative behavior of the dynamics. In fact, one can readily

confirm that the Fenchel coupling G_F is conserved under Eq. (E19):

$$\begin{aligned} \frac{d}{dt} G_F(x(t), y(t)) &= \sum_{i \in \mathcal{N}} \left\langle \frac{dy^i(t)}{dt}, \nabla h_i^*(y^i(t)) - x^{i*} \right\rangle \\ &= (1 + \alpha)^{-1} \sum_{i,j} \langle U^{(ij)} x^j(t), x^i(t) - x^{i*} \rangle \\ &= 0, \end{aligned} \quad (\text{E20})$$

where in the last step we used the fact given in Eq. (D36). Thus, the solution to Eq. (E19) does not converge in general, regardless of the value of α . As this example demonstrates, it is generically nontrivial whether the continuum limit of a discrete convergent dynamics turns out to be convergent again, and vice versa.

APPENDIX F: ADDITIONAL RESULTS

1. Extra conservation laws in the FTRL Hamiltonian system

We provide two extra conservation laws in the FTRL Hamiltonian system. These two represent the conservation laws under the translation symmetry and the dilation symmetry of the FTRL Hamiltonian, respectively. In this subsection, we always have in mind that we have a game Γ_{-1} with the Hamiltonian (18).

Proposition F1. If there exist $c^i \in \mathbb{R}^{|A_i|}$, $i \in \mathcal{N}$, such that

$$\sum_{i \in \mathcal{N}} U^{(ji)} c^i = 0, \quad (\text{F1})$$

for all $j \in \mathcal{N}$, then the following function is a conserved quantity of the Hamiltonian system:

$$G_C(x, y) = \sum_{i \in \mathcal{N}} \langle c^i, y^i \rangle. \quad (\text{F2})$$

Proof. The FTRL Hamiltonian (18) is invariant under the canonical transformation,

$$x^i = x^i + \epsilon c^i, \quad y^i = y^i, \quad (\text{F3})$$

for all $i \in \mathcal{N}$:

$$\begin{aligned} H_{\text{FTRL}}(x', y') - H_{\text{FTRL}}(x, y) &= \epsilon \sum_j \left\langle \nabla h_j^*(y^j), \sum_i U^{(ji)} c^i \right\rangle \\ &= 0. \end{aligned} \quad (\text{F4})$$

A generator $g \in C^\infty(M_{\Gamma_{-1}})$ of this canonical transformation satisfies

$$\{x_a^i, g\}_P = \frac{\partial g(x, y)}{\partial y_a^i} = c_a^i, \quad (\text{F5})$$

$$\{y_a^i, g\}_P = -\frac{\partial g(x, y)}{\partial x_a^i} = 0. \quad (\text{F6})$$

We thus obtain $g(x, y) = \sum_{i \in \mathcal{N}} \langle c^i, y^i \rangle + \text{const.}$ By Noether's theorem, the function g is conserved in the FTRL Hamiltonian system, implying that G_C is conserved. ■

When $c^i = \mathbf{1}$, this proposition shows the conservation of the total cumulative payoff, which is quite reasonable for zero-sum games.

Proposition F2. If the regularizers satisfy the “homogeneity” condition,

$$\text{Hess}(h_i^*(y^i)) y^i = \nabla h_i^*(y^i), \quad (\text{F7})$$

for all i and y^i , then the following function is a conserved quantity of the FTRL Hamiltonian system:

$$G_D(x, y) = \sum_{i \in \mathcal{N}} \langle x^i, y^i \rangle. \quad (\text{F8})$$

Proof. Under the assumption, the FTRL Hamiltonian (18) has the dilation symmetry; namely, it is invariant under the canonical transformation,

$$x^i = x^i + \epsilon x^i, \quad y^i = y^i - \epsilon y^i, \quad (\text{F9})$$

for all $i \in \mathcal{N}$:

$$\begin{aligned} H_{\text{FTRL}}(x', y') - H_{\text{FTRL}}(x, y) &= \sum_{i,j} [\langle \nabla h_j^*(y^j), U^{(ji)} \epsilon x^i \rangle \\ &\quad + \langle \text{Hess}(h_j^*(y^j))(-\epsilon y^i), U^{(ji)} x^i \rangle] \\ &= \epsilon \sum_{i,j} \langle \nabla h_j^*(y^j) - \text{Hess}(h_j^*(y^j)) y^i, U^{(ji)} x^i \rangle \\ &= 0. \end{aligned} \quad (\text{F10})$$

A generator $g \in C^\infty(M_{\Gamma_{-1}})$ of this canonical transformation satisfies

$$\{x_a^i, g\}_P = \frac{\partial g(x, y)}{\partial y_a^i} = x_a^i, \quad (\text{F11})$$

$$\{y_a^i, g\}_P = -\frac{\partial g(x, y)}{\partial x_a^i} = -y_a^i. \quad (\text{F12})$$

We thus obtain $g(x, y) = \sum_{i \in \mathcal{N}} \langle x^i, y^i \rangle + \text{const.}$ By Noether's theorem, the function g is conserved in the FTRL Hamiltonian system, implying that G_D is conserved. ■

2. A generalization of DFTRL

By extending the argument in the proof of Theorem 2, i.e., Appendix D4, we identify a potential generalization of the DFTRL algorithm.

Proposition F3. For $m \in \mathbb{Z}_{\geq 0}$, the DFTRL algorithm (Theorem 2) can be generalized to the $(4m+1)$ -power perturbations:

$$\begin{aligned} g^i(x, y) &= U^{(i \blacktriangle)} (HU)^{4m+1} x^\blacktriangledown \\ &:= U^{(i j_1)} H_{j_1} U^{(j_1 j_2)} \dots H_{j_{4m+1}} U^{(j_{4m+1} j_{4m+2})} x^{j_{4m+2}}, \end{aligned} \quad (\text{F13})$$

where $H_{j_i} := \text{Hess}(h_{j_i}^*(y^{j_i}))$ and the repeated indices are summed over all possible values.

Proof. The proof follows straightforwardly by expanding the argument in the proof of Theorem 2. Under the modified FTRL dynamics [Eqs. (15) and (41)] together with Eq. (F13),

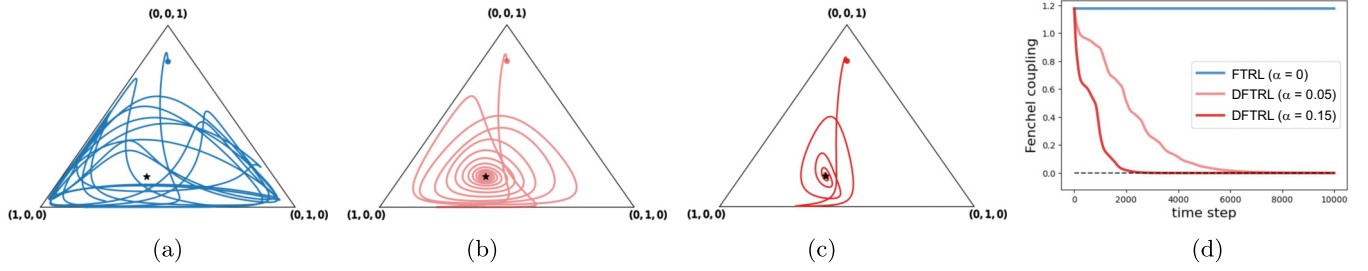


FIG. 3. Two-player weighted Rock-Paper-Scissors. (a) FTRL ($\alpha = 0$); (b) DFTRL ($\alpha = 0.05$); (c) DFTRL ($\alpha = 0.15$), and (d) Fenchel coupling G_F .

the time evolution of the Fenchel coupling reads as

$$\begin{aligned} \frac{d}{dt} G_F(x(t), y(t)) &= \alpha \sum_{i \in \mathcal{N}} \langle g^i(x(t), y(t)), x^i(t) - x^{i*} \rangle \\ &= \alpha \langle U^{(i_{j_1})} H_{j_1} U^{(j_1 j_2)} \dots H_{j_{4m+1}} \\ &\quad \times U^{(j_{4m+1} j_{4m+2})} x^{j_{4m+2}}(t), x^i(t) - x^{i*} \rangle. \end{aligned} \quad (\text{F14})$$

It is understood that the repeated indices are summed over all possible values. By moving the $2m + 1$ payoff matrices $U^{(i_{j_1})}, \dots, U^{(j_{2m} j_{2m+1})}$ to the opposite side in the inner product, along with the $2m$ Hessians, we have

$$\begin{aligned} \frac{d}{dt} G_F(x(t), y(t)) &= -\alpha \langle H_{j_{2m+1}} (UH)^{2m} U^{(\blacktriangle j)} (x^j(t) - x^{j*}), \\ &\quad \times (UH)^{2m} U^{(\blacktriangledown i)} (x^i(t) - x^{i*}) \rangle. \end{aligned} \quad (\text{F15})$$

The overall minus sign arises from the zero-sum property of the $2m + 1$ payoff matrices. The following abbreviations are introduced:

$$(UH)^{2m} U^{(\blacktriangle j)} = U^{(j_{2m+1} j_{2m+2})} H_{j_{2m+2}} \dots U^{(j_{4m} j_{4m+1})} H_{j_{4m+1}} U^{(j_{4m+1} j)}, \quad (\text{F16})$$

$$(UH)^{2m} U^{(\blacktriangledown i)} = U^{(j_{2m+1} j_{2m})} H_{j_{2m}} \dots U^{(j_2 j_1)} H_{j_1} U^{(j_1 i)}. \quad (\text{F17})$$

Since the repeated indices are summed over and hence act as the dummies, one finds that the vectors on both sides of the Hessian $H_{j_{2m+1}}$ are the same. Thus, from the positive

definiteness of the Hessian, we obtain

$$\frac{d}{dt} G_F(x(t), y(t)) \leq 0. \quad (\text{F18})$$

The equality is attained when $x(t) = x^*$ or $\alpha = 0$, for which the system reduces to the ordinary FTRL dynamics.

This generalized DFTRL algorithm reduces to Theorem 2 when $m = 0$. As discussed in the main text, the Hamiltonian (18) admits a one-parameter deformation of the dynamics that preserves the Hamiltonian. As an example, we examined the damping of the Fenchel coupling, for which Eq. (43) was designed. While we have not yet been able to demonstrate its uniqueness, it is likely that any one-parameter deformation without derivative terms that damps the Fenchel coupling while conserving the Hamiltonian (18) will, up to a constant, be equivalent to Eq. (F13). The generalized perturbation (F13) might have additional nice properties for learning in games, such as an improvement in the convergence rate. However, we do not discuss these aspects here and leave them for future work.

APPENDIX G: EXPERIMENTAL DETAILS AND ADDITIONAL SIMULATIONS

In Sec. IV, we numerically solve the FTRL/DFTRL dynamics for the well-known examples. Beyond the typical setups for experimental simulations given in the main text, we provide two additional examples to validate our theoretical results. To this end, we begin with more detailed descriptions of the experimental setups.

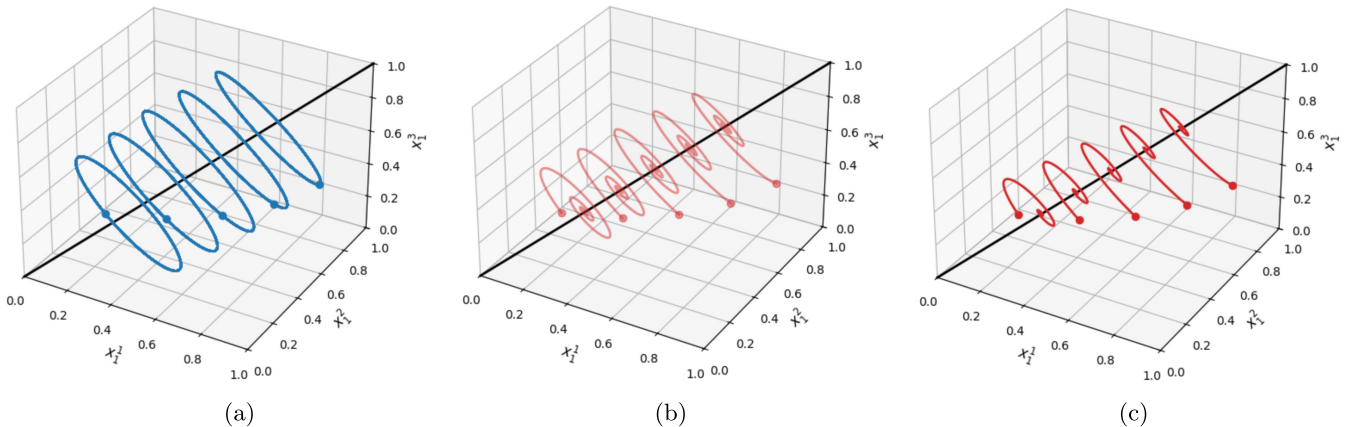


FIG. 4. Three-player Matching Pennies. (a) FTRL ($\alpha = 0$), (b) DFTRL ($\alpha = 0.05$), and (c) DFTRL ($\alpha = 0.1$).

Let us examine the two-player *weighted* Rock-Paper-Scissors game, where $\mathcal{N} = \{1, 2\}$ and $A_i = \{R, P, S\} \simeq \{1, 2, 3\}$ [62]. This game is characterized by the payoff matrices,

$$U^{(12)} = U^{(21)} = \begin{pmatrix} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{pmatrix}, \quad (\text{G1})$$

where $a, b, c \in \mathbb{R}$. In this game, the Nash equilibrium is explicitly given by

$$x^{i*} = \left(\frac{c}{a+b+c}, \frac{b}{a+b+c}, \frac{a}{a+b+c} \right), \quad i = 1, 2. \quad (\text{G2})$$

In addition to these data, a choice of regularizer functions defines the dynamical system [Eq. (41)]. Given initial conditions $x^1(0)$ and $x^2(0)$, the dynamical system with the relations [(42) and (43)] generates the solution. In Sec. IV, we take $a = b = c = 1$, the entropic regularizer $h_i^*(y^i) = \text{lse}(y^i)$, and the initial conditions $x^1(0) = x^2(0) = (0.1, 0.1, 0.8)$, which result in Fig. 2. We now consider the case of $a = 1, b = 2, c = 3$, and the initial conditions $x^1(0) = (0.1, 0.1, 0.8)$ and $x^2(0) = (0.2, 0.6, 0.2)$, where the entropic regularizer is kept. Figure 3 depicts the results. As one can see, the solution trajectory for $\alpha = 0$ (FTRL) exhibits nonconvergent behavior, and its Fenchel coupling is conserved over time. Our DFTRL

algorithm yields convergent dynamics and we see that the larger perturbation coefficient α promotes the faster convergence, as expected.

We also consider the three-player Matching Pennies, with different regularizer functions. This game is characterized by $\mathcal{N} = \{1, 2, 3\}$ and $A_i = \{H, T\} \simeq \{1, 2\}$. We set the payoff matrices as

$$U^{(12)} = U^{(23)} = U^{(31)} = \begin{pmatrix} a & -1 \\ -1 & a \end{pmatrix}, \quad U^{(ij)} = -(U^{(ji)})^\top, \quad (\text{G3})$$

where $a \in \mathbb{R}$. In this game, we have a continuum of Nash equilibria distributed along a straight line:

$$(x^{1*}, x^{2*}, x^{3*}) = ((p, 1-p), (p, 1-p), (p, 1-p)), \quad (\text{G4})$$

where $0 \leq p \leq 1$. Again, building on these data with a choice of regularizer functions, we solve the dynamical system [Eq. (41)] for a given initial condition. In Fig. 1, we take $a = 1$, the entropic regularizer, and initial conditions chosen at random. Here, we employ the Euclidean regularizer, while the other setups are kept intact. The result is shown in Fig. 4. Similarly to the previous example, we observe that the solution trajectories for $\alpha = 0$ (FTRL) exhibit cyclic behavior, and that the DFTRL dynamics for larger α converges faster to the Nash equilibrium.

-
- [1] D. Fudenberg and D. K. Levine, *The Theory of Learning in Games* (MIT press, Cambridge, 1998), Vol. 2.
 - [2] J. F. Nash, Equilibrium points in n -person games, *Proc. Natl. Acad. Sci. USA* **36**, 48 (1950).
 - [3] P. Mertikopoulos, C. Papadimitriou, and G. Piliouras, Cycles in adversarial regularized learning, in *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* (SIAM, Philadelphia, 2018), pp. 2703–2717.
 - [4] S. Shalev-Shwartz and Y. Singer, Convex repeated games and fenchel duality, in *Proceedings of the 20th International Conference on Neural Information Processing Systems* (MIT Press, Cambridge, MA, USA, 2006), pp. 1265–1272.
 - [5] J. Hofbauer, Evolutionary dynamics for bimatrix games: A Hamiltonian system? *J. Math. Biol.* **34**, 675 (1996).
 - [6] J. P. Bailey and G. Piliouras, Multi-agent learning in network zero-sum games is a Hamiltonian system, in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (2019), pp. 233–241.
 - [7] C. Daskalakis and C. H. Papadimitriou, On a network generalization of the minmax theorem, in *International Colloquium on Automata, Languages, and Programming* (Springer, 2009), pp. 423–434.
 - [8] Y. Cai and C. Daskalakis, On minmax theorems for multiplayer games, in *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms* (SIAM, Philadelphia, 2011), pp. 217–234.
 - [9] Y. Cai, O. Candogan, C. Daskalakis, and C. Papadimitriou, Zero-sum polymatrix games: A generalization of minmax, *Math. Oper. Res.* **41**, 648 (2016).
 - [10] We make comments on the paper by Bailey and Piliouras in Appendix C.
 - [11] Y. Sato, E. Akiyama, and J. D. Farmer, Chaos in learning a simple two-person game, *Proc. Natl. Acad. Sci. USA* **99**, 4748 (2002).
 - [12] G. Ostrovski and S. van Strien, Piecewise linear Hamiltonian flows associated to zero-sum games: Transition combinatorics and questions on ergodicity, *Regul. Chaot. Dyn.* **16**, 128 (2011).
 - [13] S. van Strien, Hamiltonian flows with random-walk behaviour originating from zero-sum games and fictitious play, *Nonlinearity* **24**, 1715 (2011).
 - [14] D. Balduzzi, S. Racaniere, J. Martens, J. Foerster, K. Tuyls, and T. Graepel, The mechanics of n -player differentiable games, in *International Conference on Machine Learning* (PMLR, 2018), pp. 354–363.
 - [15] A. Letcher, D. Balduzzi, S. Racaniere, J. Martens, J. Foerster, K. Tuyls, and T. Graepel, Differentiable game mechanics, *J. Mach. Learn. Res.* **20**, 3032 (2019).
 - [16] H. N. Alishah and P. Duarte, Hamiltonian evolutionary games, *J. Dyn. Games* **2**, 33 (2015).
 - [17] A. H. Najafi, Conservative replicator and lotka-volterra equations in the context of Dirac\big-isotropic structures, *J. Geom. Mech.* **12**, 149 (2020).
 - [18] Y. Cai, A. Oikonomou, and W. Zheng, Finite-time last-iterate convergence for learning in multi-player games, in *Advances in Neural Information Processing Systems* (2022).
 - [19] C. Daskalakis and I. Panageas, Last-iterate convergence: Zero-sum games and constrained min-max optimization, in *10th Innovations in Theoretical Computer Science* (2019).

- [20] C. Daskalakis, A. Ilyas, V. Syrgkanis, and H. Zeng, Training gans with optimism, in *International Conference on Learning Representations* (2018).
- [21] V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire, Fast convergence of regularized learning in games, in *Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 2* (MIT Press, Cambridge, MA, USA, 2015), pp. 2989–2997.
- [22] S. Rakhlin and K. Sridharan, Optimization, learning, and games with predictable sequences, in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2* (Curran Associates Inc., Red Hook, NY, USA), pp. 3066–3074.
- [23] S. Lee and D. Kim, Fast extra gradient methods for smooth structured nonconvex-nonconcave minimax problems, *Adv. Neural Inf. Process. Syst.* **34**, 22588 (2021).
- [24] P. Mertikopoulos, B. Lecouat, H. Zenati, C.-S. Foo, V. Chandrasekhar, and G. Piliouras, Optimistic mirror descent in saddle-point problems: Going the extra(-gradient) mile, in *International Conference on Learning Representations* (2019).
- [25] G. M. Korpelevich, The extragradient method for finding saddle points and other problems, *Ekon. mat. metod.* **12**, 747 (1976).
- [26] R. A. Hemmat, A. Mitra, G. Lajoie, and I. Mitliagkas, LEAD: Min-max optimization from a physical perspective, *ICML Workshop on New Frontiers in Learning, Control, and Dynamical Systems* (2023).
- [27] G. Zhang and Y. Wang, On the suboptimality of negative momentum for minimax optimization, in *International Conference on Artificial Intelligence and Statistics* (PMLR, 2021), pp. 2098–2106.
- [28] N. B. Kovachki and A. M. Stuart, Continuous time analysis of momentum methods, *J. Mach. Learn. Res.* **22**, 1 (2021).
- [29] G. Gidel, R. A. Hemmat, M. Pezeshki, R. Le Priol, G. Huang, S. Lacoste-Julien, and I. Mitliagkas, Negative momentum for improved game dynamics, in *The 22nd International Conference on Artificial Intelligence and Statistics* (PMLR, 2019), pp. 1802–1811.
- [30] B. T. Polyak, Some methods of speeding up the convergence of iteration methods, *USSR Comput. Math. Math. Phys.* **4**, 1 (1964).
- [31] Y. Feng, H. Fu, Q. Hu, P. Li, I. Panageas, B. Peng, and X. Wang, On the last-iterate convergence in time-varying zero-sum games: Extra gradient succeeds where optimism fails, *Adv. Neural Inf. Process. Syst.* **36**, 21933 (2023).
- [32] Y. Feng, P. Li, I. Panageas, and X. Wang, Last-iterate convergence separation between extra-gradient and optimism in constrained periodic games, *Uncertainty in Artificial Intelligence* (PMLR, 2024), pp. 1339–1370.
- [33] The $j = i$ term can be consistently included in the sum, since by our definition of Γ the corresponding payoff matrix is null; $U^{(ii)} \equiv 0$, which implies that agents do not engage in self-play (do not have self-interaction). The same applies in the subsequent sections generally, although we occasionally remove the $j = i$ term from the sum for clarity.
- [34] In this paper, we focus on games of Γ_σ . However, the discussion in this subsection is generically applicable to Γ , not only to Γ_σ .
- [35] S. Shalev-Shwartz, Online learning and online convex optimization, *Found. Trends Mach. Learn.* **4**, 107 (2012).
- [36] P. D. Taylor and L. B. Jonker, Evolutionary stable strategies and game dynamics, *Math. Biosci.* **40**, 145 (1978).
- [37] P. D. Taylor, Evolutionarily stable strategies with two types of player, *J. Appl. Probab.* **16**, 76 (1979).
- [38] P. Schuster and K. Sigmund, Replicator dynamics, *J. Theor. Biol.* **100**, 533 (1983).
- [39] D. Friedman, Evolutionary games in economics, *Econometrica* **59**, 637 (1991).
- [40] A. Nagurny and D. Zhang, Projected dynamical systems in the formulation, stability analysis, and computation of fixed-demand traffic network equilibria, *Transp. Sci.* **31**, 147 (1997).
- [41] R. Lahkar and W. H. Sandholm, The projection dynamic and the geometry of population games, *Games Econ. Behav.* **64**, 565 (2008).
- [42] W. H. Sandholm, E. Dokumacı, and R. Lahkar, The projection dynamic and the replicator dynamic, *Games Econ. Behav.* **64**, 666 (2008).
- [43] P. Mertikopoulos and W. H. Sandholm, Learning in games via reinforcement and regularization, *Math. Oper. Res.* **41**, 1297 (2016).
- [44] P. Mertikopoulos and Z. Zhou, Learning in games with continuous action sets and unknown payoff functions, *Math. Program.* **173**, 465 (2019).
- [45] The code is available at <https://github.com/Toshihiro-Ota/dftrl>.
- [46] J. Nash, Non-cooperative games, *Ann. Math.* **54**, 286 (1951).
- [47] E. Tewolde, B. H. Zhang, C. Oesterheld, T. Sandholm, and V. Conitzer, Computing game symmetries and equilibria that respect them, *Proc. AAAI Conf. Artif. Intell.* **39**, 14148 (2025).
- [48] J. Eisert, M. Wilkens, and M. Lewenstein, Quantum games and quantum strategies, *Phys. Rev. Lett.* **83**, 3077 (1999).
- [49] J. Eisert and M. Wilkens, Quantum games, *J. Mod. Opt.* **47**, 2543 (2000).
- [50] R. Jain, G. Piliouras, and R. Sim, Matrix multiplicative weights updates in quantum zero-sum games: Conservation laws & recurrence, *Adv. Neural Inf. Process. Syst.* **35**, 4123 (2022).
- [51] K. Lotidis, P. Mertikopoulos, and N. Bambos, Learning in quantum games, [arXiv:2302.02333](https://arxiv.org/abs/2302.02333).
- [52] W. Lin, G. Piliouras, R. Sim, and A. Varvitsiotis, Learning in quantum common-interest games and the separability problem, *Quantum* **9**, 1689 (2025).
- [53] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, Neural ordinary differential equations, *Adv. Neural Inf. Process. Syst.* (2018).
- [54] R. T. Q. Chen, Torchdiff, GitHub (2018), <https://github.com/rtqichen/torchdiff>.
- [55] D. P. Bertsekas, Nonlinear programming, *J. Oper. Res. Soc.* **48**, 334 (1997).
- [56] R. Bergmann and R. Herzog, Intrinsic formulation of KKT conditions and constraint qualifications on smooth manifolds, *SIAM J. Optim.* **29**, 2423 (2019).
- [57] P. Libermann and C.-M. Marle, *Symplectic Geometry and Analytical Mechanics* (Springer, Dordrecht, 1987).
- [58] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics* (Springer, New York, NY, 1989).
- [59] In a narrow sense, the term “Hamiltonian dynamics” may be referred to as the solution to the flow equation (B2) or (B3). However, we use this term collectively including the phase space, the Hamiltonian system, and its solutions, etc.
- [60] In the literature, each time step of ϵ is implicitly assumed. Compared to our notation, our $\hat{\phi}(n\epsilon)$ corresponds to their x_n , etc.

- [61] To be more precise, it should be $\alpha/T \ll 1$, since the perturbative coefficient α is a dimensionful parameter.
- [62] Since this paper has established the Hamiltonian structure for learning dynamics in games, also known as evolutionary dynamics, applying or extending our results to various analyses of Rock-Paper-Scissors games using physics-based approaches (such as Refs. [63–65]) could yield insights.
- [63] K. M. A. Kabir and J. Tanimoto, The role of pairwise nonlinear evolutionary dynamics in the rock–paper–scissors game with noise, [Appl. Math. Comput.](#) **394**, 125767 (2021).
- [64] M. Tenorio, E. Rangel, and J. Menezes, Adaptive movement strategy in rock-paper-scissors models, [Chaos Solitons Fractals](#) **162**, 112430 (2022).
- [65] R. R. Justino, B. F. de Oliveira, and F. E. A. dos Santos, Critical phenomena in the rock-paper-scissors model, [Physica A](#) **680**, 131018 (2025).