Journal of Nonlinear Science (2019) 29:871–896
https://doi.org/10.1007/s00332-018-9507-5

Population Games and Discrete Optimal Transport

Shui-Nee Chow[1] · Wuchen Li[2] · Jun Lu[1] · Haomin Zhou[1]

Received: 14 August 2017 / Accepted: 10 October 2018 / Published online: 24 October 2018

Abstract
We propose an evolutionary dynamics for population games with discrete strategy sets, inspired by optimal transport theory and mean field games. The proposed dynamics is the Smith dynamics with strategy graph structure, in which payoffs are modified by logarithmic terms. The dynamics can be described as a Fokker–Planck equation on a discrete strategy set. For potential games, the dynamics is a gradient flow system under a Riemannian metric from optimal transport theory. The stability of the dynamics is studied through optimal transport metric tensor, free energy and Fisher information.

Keywords Evolutionary game theory · Optimal transport · Mean field games · Fokker–Planck equations

1 Introduction

Population games are introduced as a framework to model population behaviors and study strategic interactions in populations by extending finite player games (Nash 1950; Sigmund and Nowak 1999; Von Neumann and Morgenstern 2007). It has fundamental impact on game theory related to social networks, evolution of biology species, virus and cancer, etc (Huang et al. 2015; Mertikopoulos and Sandholm 2016; Shah and Shin 2010; Wu et al. 2014). Nash equilibrium (NE) describes a status that no player in population is willing to change his/her strategy unilaterally. To investigate stabilities of NEs, evolutionary game theory (Nowak 2006; Sandholm 2012b; Sigmund and Nowak 1999) has been developed in the last several decades. Researchers from various fields

Communicated by Paul Newton.

This work is partially supported by NSF Awards DMS–1419027, DMS-1620345, and ONR Award N000141310408.

✉ Wuchen Li
wcli@math.ucla.edu

1 Georgia Institute of Technology, Atlanta, USA

2 University of California Los Angeles, Los Angeles, USA

(economics, biology, etc) design different dynamics, called mean dynamics or evolutionary dynamics (Hofbauer and Sigmund 2003; Sandholm 2012a), under various assumptions to describe population behaviors. Important examples include Replicator, Best-response, Logit and Smith dynamics (Matsui 1992; Shah and Shin 2010; Smith 1984), just to name a few. A special class of games, named potential games (Hofbauer and Sigmund 1988; Monderer and Shapley 1996; Sandholm 2010), are widely considered. Heuristically, potential games describe the situation that all players face the same payoff function, called potential. Thus, maximizing each player's own payoff is equivalent to maximizing the potential. In this case, NEs correspond to maximizers of the potential, which gives natural connections between mean dynamics and gradient flows obtained from minimizing the negative potential. An important example is the Replicator dynamics, which is a gradient flow of the negative potential in the probability space (simplex) with a Shahshahani metric (Akin 1979; Mertikopoulos and Sandholm 2016; Shahshahani 1979).

To study evolutionary dynamics, modeling uncertainties in individual players' decision processes plays vital roles. Usually such uncertainties are introduced by the notion of noisy potential, i.e., the expected payoff added with Shannon–Boltzmann entropy. One well-known example is the Logit dynamics (Fudenberg and Levine 1998; Hofbauer and Sandholm 2002, 2007; Sandholm 2010), whose solution is forced to converge to critical points of the noisy potential. On the other hand, for population games with continuous strategy sets, there is a natural way to introduce uncertainties by adding white noise, see mean field games introduced by Cardaliaguet (2010), Lasry and Lions (2007) and Best-reply dynamics (Degond et al. 2014). Their results relate to Smith dynamics (1984) originated from studying traffic flows (Smith 1984) by the fact that the Smith dynamics can be viewed as a discrete continuity equation. Mean field games have continuous strategy sets (Blanchet and Carlier 2012, 2014). Each player is assumed to make decisions according to a stochastic process instead of making a one-shot decision. More specifically, individual players change their pure strategies *locally* and simultaneously in a continuous fashion according to the direction that maximizes their own payoff functions most rapidly, and randomness is introduced in the form of white noise perturbation. The resulting dynamics for individual players forms a mean field type stochastic differential equation, whose probability density function evolves according to the Fokker–Planck equation, i.e., continuity equation with diffusion processes. Here, mean field serves as a mediator for aggregating individual players' behaviors. For potential games (Degond et al. 2014), Fokker–Planck equations can also be viewed as gradient flows of negative noisy potential in the probability space.

The aim of this paper is to further the mathematical understandings of optimal transport theory in mean field games and populations games, especially when the strategy set is discrete. We propose an evolutionary equation via gradient flow in discrete optimal transport metric tensor. We note that it is not a straightforward task to transform the theory on games with continuous strategy set directly to discrete settings. This is due to the fact that the discrete strategy set is no longer a length space, a space that one can define length of curves, and morph one curve to another in a continuous fashion. To proceed, we employ key tools developed in Chow et al. (2017a, b), Li (2016) [Similar topics are discussed in Chow et al. (2012), Erbar and Maas (2012), Maas (2011)]. More specifically, we impose a Riemannian metric tensor on the probability space of the strategy. With the Riemannian structure in probability simplex, we derive the gradient flow of the negative noisy potential as mean dynamics.

In detail, let us consider a population game with finite discrete strategy set $S = \{1, \dots, n\}$. Denote the set of population state

$$\mathcal{P}(S) = \left\{ (\rho_i)_{i=1}^n \in \mathbb{R}^n \ : \ \sum_{i=1}^n \rho_i = 1 \ , \ \rho_i \geq 0 \ , \ i \in S \right\} ,$$

and payoff function $F_i : \mathcal{P}(S) \to \mathbb{R}$, for any $i \in S$. The derived mean dynamics is given by

$$\begin{aligned} \frac{d\rho_i}{dt} &= \sum_{j \in N(i)} \frac{1}{d_j} \rho_j [F_i(\rho) - F_j(\rho) + \beta(\log \rho_j - \log \rho_i)]_+ \\ &\quad - \sum_{j \in N(i)} \frac{1}{d_i} \rho_i [F_j(\rho) - F_i(\rho) + \beta(\log \rho_i - \log \rho_j)]_+ \ , \end{aligned} \tag{1}$$

where $\beta \geq 0$ is the strength of uncertainty[1], $\rho_i(t)$ is the probability at time t of strategy $i \in S$, $[\cdot]_+ = \max\{\cdot, 0\}$, $j \in N(i)$ if j can be achieved by players changing their strategies from i and $d_i = \sum_{j \in N(i)} 1$ represents the degree of graph at note i. We call (1) Fokker–Planck equation of a game.

Dynamics (1) has many appealing features. For potential games, the dynamics is shown to be a gradient flow, whose equilibria are discrete Gibbs measures. Their stability properties can also be studied by leveraging two notions, namely *relative free energy* and *relative Fisher information* (Frieden 2004; Villani 2008). Through their relations with optimal transport metric tensor, we show that the relative entropy converges to 0 as t goes to infinity, and the solution converges to the Gibbs measure exponentially fast. For general games, (1) is not a gradient flow, which may exhibit complicated limiting behaviors including Hopf bifurcations. And the noise level becomes a natural parameter to introduce such bifurcations.

When $\beta = 0$ and the strategy graph is complete, then dynamics (1) is exactly the Smith dynamics. When $\beta > 0$, (1) still fits into the Smith dynamics framework with modified payoff functions. From this viewpoint, many mathematical properties of dynamics (1), including the convergence to NEs, can be derived using existing methods for Smith dynamics Sandholm (2010). In addition, one by-product of our model is that the Smith dynamics can be viewed as gradient flows of negative potentials under a optimal transport metric tensor. So many studies in Sandholm (2010) have natural analog or extensions in optimal transport. On the other hand, while both Logit dynamics and the proposed model converge to Gibbs measures, they differ in the following aspects: (i) For Logit dynamics of potential games, the noisy potential is the Lyapunov function, while for (1), it is the objective function of a gradient flow. This additional property gives rise to the exponential convergence results; (ii) in the formulation, the Logit dynamics depends on the information of all strategies (all F_is),

[1] β represents the inverse of temperature.

while (1) only depends on the local information (neighboring F_is). Last but not least, the proposed dynamics depends on the graph structure of strategy set, which is different from the Replicator dynamics (Coucheney et al. 2015; Leslie and Collins 2005), in which all discrete strategies are treated equally.

The arrangement of this paper is as follows. In Sect. 2, we give a brief introduction to population games on discrete sets. In Sect. 3, we derive (1) by an optimal transport metric defined on the simplex set and introduce the Markov process associated with (1) from the modeling perspective. In Sect. 4, we study (1)'s longtime behavior by relative free energy and relative Fisher information. In Sect. 5, we discuss the application of our dynamics by working on some well-known population games.

2 Preliminaries

Consider a game played by a continuum of players. Each player in the population selects a pure strategy from the discrete strategy set $S = \{1, \ldots, n\}$. The aggregated state of the population can be described by the population state $\rho = (\rho_i)_{i=1}^n \in \mathcal{P}(S)$, where ρ_i represents the proportion of players choosing pure strategy i and $\mathcal{P}(S)$ is a probability space (simplex):

$$\mathcal{P}(S) = \left\{ (\rho_i)_{i=1}^n \in \mathbb{R}^n \ : \ \sum_{i=1}^n \rho_i = 1 \ , \ 0 \le \rho_i \le 1 \ , \ i \in S \right\} .$$

The game assumes that each player's payoff is independent of his/her identity (autonomous game). Thus, all players choosing strategy i have the continuous payoff function $F_i : \mathcal{P}(S) \to \mathbb{R}$.

A population state $\rho^* \in \mathcal{P}(S)$ is a Nash equilibrium of the population game if

$$\rho_i^* > 0 \text{ implies that } F_i(\rho^*) \ge F_j(\rho^*) \ , \quad \text{for all } j \in S.$$

The following type of population games has particular importance, in which NEs enjoy various prominent properties.

A population game is named a *potential game*, if there exists a differentiable potential function $\mathcal{F} : \mathcal{P}(S) \to \mathbb{R}$, such that $\frac{\partial}{\partial \rho_i}\mathcal{F}(\rho) = F_i(\rho)$, for all $i \in S$. It is a well-known fact that the NEs of a potential game are the stationary points of $\mathcal{F}(\rho)$.

Example Suppose that a unit mass of agents are randomly matched to play symmetric normal-form game with payoff matrix $A \in \mathbb{R}^{n\times n}$. At population state ρ, a player choosing strategy i receives payoff equal to the expectation of the others, i.e., $F_i(\rho) = \sum_{j\in S} a_{ij}\rho_j$. In particular, if the payoff matrix A is symmetric, then the game becomes a potential game with potential function $\mathcal{F}(\rho) = \frac{1}{2}\rho^{\mathrm{T}} A\rho$, since $\frac{\partial}{\partial \rho_i}\mathcal{F}(\rho) = F_i(\rho)$.

Given a potential game with potential $\mathcal{F}$, define the *noisy potential*

$$\bar{\mathcal{F}}(\rho) := \mathcal{F}(\rho) - \beta \sum_{i=1}^n \rho_i \log \rho_i \ , \quad \beta > 0 \ ,$$

which is the summation of potential and Shannon–Boltzmann entropy. In information theory, it has been known for a long time that the entropy is a way to model uncertainties (Frieden 2004). In the context of population games, such uncertainties may refer to players' irrational behaviors, making mistakes or risk-taking behaviors. In optimal transport theory, the negative noisy potential is usually called the *free energy* (Villani 2003, 2008).

The problem of maximizing each player's payoff with uncertainties is equivalent to maximizing the noisy potential (minimizing the free energy)

$$\min\{-\bar{\mathcal{F}}(\rho) \;:\; \rho \in \mathcal{P}(S)\}.$$

We call the stationary points ρ^* of the above minimization the discrete Gibbs measures, i.e., ρ^* solves the following fixed point problem

$$\rho_i^* = \frac{1}{K} e^{\frac{F_i(\rho^*)}{\beta}}, \text{ for any } i \in S, \text{ where } K = \sum_{j=1}^{n} e^{\frac{F_j(\rho^*)}{\beta}}. \tag{2}$$

3 Evolutionary Dynamics Via Discrete Optimal Transport

In this section, we first introduce an optimal transport metric for population games. Based on the metric, we derive a new interpretation of Smith dynamics with modified payoff function. For potential games, the Smith dynamics can be viewed as gradient flows.

3.1 Optimal Transport Metric for Games

We first introduce the optimal transport metric tensor in probability simplex.

We start with the construction of strategy graphs. A strategy graph $G = (S, E)$ is a *neighborhood* structure imposed on the strategy set $S = \{1, \ldots, n\}$. Two vertices $i, j \in S$ are connected in G if players who currently choose strategy i is able to switch to strategy j. Denote the neighborhood of i by

$$N(i) = \{j \in S \colon (i, j) \in E\}.$$

For many games, every two strategies are connected, making G a complete graph. In other words, $N(i) = S \setminus \{i\}$, for any $i \in S$. For example, the strategy set of Prisoner Dilemma game is either cooperation (C) or defection (D), i.e., $S = \{C, D\}$. Thus, the strategy graph is

For any given strategy graph G, we can introduce an optimal transport metric on the simplex $\mathcal{P}(S)$. Denote the interior of $\mathcal{P}(S)$ by $\mathcal{P}_o(S)$.

Given a function $\Phi\colon S \to \mathbb{R}$, define $\nabla\Phi\colon S \times S \to \mathbb{R}$ as

$$\nabla\Phi_{ij} = \begin{cases} \Phi_i - \Phi_j & \text{if } (i, j) \in E \ ; \\ 0 & \text{otherwise} \ . \end{cases}$$

Let $m\colon S \times S \to \mathbb{R}$ be an anti-symmetric flux function such that $m_{ij} = -m_{ij}$. The divergence of m, denoted as $\mathrm{div}(m)\colon S \to \mathbb{R}$, is defined by

$$\mathrm{div}(m)_i = -\sum_{j\in N(i)} m_{ij} \ .$$

For the purpose of defining our distance function, we will use a particular flux function

$$m_{ij} := \theta_{ij}(\rho)\nabla\Phi_{ij} \ ,$$

where $\theta_{ij}(\rho)$ represents the discrete probability on edge (i, j), defined by

$$\theta_{ij}(\rho) = \begin{cases} \frac{1}{d_j}\rho_j & \bar{F}_j(\rho) < \bar{F}_i(\rho) \ ; \\ \frac{1}{d_i}\rho_i & \bar{F}_j(\rho) > \bar{F}_i(\rho) \ ; \\ \frac{1}{2}\left(\frac{\rho_i}{d_i} + \frac{\rho_j}{d_j}\right) & \bar{F}_j(\rho) = \bar{F}_i(\rho) \ . \end{cases} \tag{3}$$

Here, $d_i = \sum_{j\in N(i)} 1$ is the degree of graph at node i and $\bar{F}_i(\rho) = F_i(\rho) - \beta \log \rho_i$.

Given two potential vector fields $\nabla\Phi$, $\nabla\tilde{\Phi}$, define

$$(\nabla\Phi, \nabla\tilde{\Phi})_\rho := \frac{1}{2}\sum_{(i,j)\in E} (\Phi_i - \Phi_j)(\tilde{\Phi}_i - \tilde{\Phi}_j)\theta_{ij}(\rho) \ , \tag{4}$$

where $\frac{1}{2}$ is applied because each edge is summed twice, i.e., (i, j), $(j, i) \in E$. The above definitions provide the following distance function on $\mathcal{P}_o(S)$.

Definition 1 Given two discrete probability functions $\rho^0, \rho^1 \in \mathcal{P}_o(S)$, the Wasserstein metric W is defined by:

$$W(\rho^0, \rho^1)^2 = \inf\left\{\int_0^1 (\nabla\Phi(t), \nabla\Phi(t))_{\rho(t)}\mathrm{d}t \ : \ \frac{\mathrm{d}\rho}{\mathrm{d}t} + \mathrm{div}(\rho\nabla\Phi) = 0 \right.$$
$$\left. \rho(0) = \rho^0, \ \rho(1) = \rho^1\right\}. \tag{5}$$

Here, the infimum is taken over pairs $(\rho(t), \Phi(t))$ with $\rho \in H^1((0, 1), \mathbb{R}^n)$ and $\Phi\colon [0, 1] \to \mathbb{R}^n$ measurable.

The Wasserstein metric induces a Riemannian metric tensor in the interior of probability simplex. Consider the tangent space at a point $\rho \in \mathcal{P}_o(S)$:

$$T_\rho \mathcal{P}_o(S) = \left\{ (\sigma_i)_{i=1}^n \in \mathbb{R}^n : \sum_{i=1}^n \sigma_i = 0 \right\}.$$

We next identify a potential vector $\Phi \in \mathbb{R}^n$ with a tangent vector $\sigma \in \mathcal{P}_o(S)$.

Lemma 2 *For given $\sigma \in T_\rho \mathcal{P}_o(S)$, there exists a unique function Φ, up to a constant shift, such that*

$$\sigma = -\operatorname{div}(\rho \nabla \Phi).$$

Proof We prove the result by rewriting the operator $-\operatorname{div}(\rho\nabla)$ into a matrix form. Denote

$$L(\rho) = D^{\mathsf{T}} \Theta(\rho) D \in \mathbb{R}^{n\times n},$$

where $D \in \mathbb{R}^{|E|\times n}$ is the discrete gradient operator

$$D_{(i,j)\in E, k\in V} = \begin{cases} \sqrt{\omega_{ij}}, & \text{if } i = k, i > j \\ -\sqrt{\omega_{ij}}, & \text{if } j = k, i > j \\ 0, & \text{otherwise;} \end{cases}$$

$-D^{\mathsf{T}} \in \mathbb{R}^{n\times|E|}$ is the discrete divergence operator, and $\Theta(\rho) \in \mathbb{R}^{|E|\times|E|}$ is a weight matrix

$$\Theta(\rho)_{(i,j)\in E,(k,l)\in E} = \begin{cases} \theta_{ij}(\rho) & \text{if } (i,j) = (k,l) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Using this matrix notation, we prove that $-\operatorname{div}(\rho\nabla\Phi) = L(\rho)\Phi = \sigma$ has a unique solution for Φ up to a constant shrift.

If $\rho \in \mathcal{P}_o(G)$, all diagonal entries of the weight matrix $\Theta(\rho)$ are nonzero. Consider

$$\Phi^{\mathsf{T}} L(\rho)\Phi = \frac{1}{2} \sum_{(i,j)\in E} \omega_{ij}(\Phi_i - \Phi_j)^2 \theta_{ij}(\rho) = 0.$$

Since $\rho_i > 0$ for any $i \in V$ and the strategy graph is connected, $\Phi_1 = \cdots = \Phi_n$ is the only solution of above equation. Thus, 0 must be the simple eigenvalue of $L(\rho)$ with eigenvector $(1, \ldots, 1)^{\mathsf{T}}$. Since $\operatorname{Ker}\big(L(\rho)\big) = \{(1, \ldots, 1)^{\mathsf{T}}\}$,

$$\mathbb{R}^n/\ker(L(\rho)) \cong \operatorname{Ran}(L(\rho)) = T_\rho \mathcal{P}_o(G).$$

Thus, there exists a unique solution of Φ up to a constant shrift. □

Based on Lemma 2, we write

$$L(\rho) = T \begin{pmatrix} 0 & & & \\ & \lambda_{\text{sec}}(L(\rho)) & & \\ & & \ddots & \\ & & & \lambda_{\max}(L(\rho)) \end{pmatrix} T^{-1} ,$$

where $0 < \lambda_{\text{sec}}(L(\rho)) \le \cdots \le \lambda_{\max}(L(\rho))$ are eigenvalues of $L(\rho)$ arranged in ascending order, and T is its corresponding eigenvector matrix. We denote the pseudo-inverse of $L(\rho)$ by

$$L(\rho)^{-1} = T \begin{pmatrix} 0 & & & \\ & \frac{1}{\lambda_{\text{sec}} L(\rho)} & & \\ & & \ddots & \\ & & & \frac{1}{\lambda_{\max} L(\rho)} \end{pmatrix} T^{-1} .$$

Here, the matrix $L(\rho)^{-1}$ endows an inner product on $T_\rho \mathcal{P}_o(G)$.

Definition 3 For any two tangent vectors $\sigma^1, \sigma^2 \in T_\rho \mathcal{P}_o(S)$, define the inner product $g^W : T_\rho \mathcal{P}_o(S) \times T_\rho \mathcal{P}_o(S) \to \mathbb{R}$ by

$$g^W_\rho(\sigma, \tilde{\sigma}) := \sigma^\mathsf{T} L(\rho)^{-1} \tilde{\sigma} = \Phi^\mathsf{T} L(\rho) \tilde{\Phi} = \frac{1}{2} \sum_{(i,j)\in E} \theta_{ij}(\rho)(\Phi_i - \Phi_j)(\tilde{\Phi}_i - \tilde{\Phi}_j) ,$$

where $\sigma = L(\rho)\Phi$ and $\tilde{\sigma} = L(\rho)\tilde{\Phi}$.

Under this inner product, we can formulate the Wasserstein metric (5) as a geometric action function

$$W(\rho^0, \rho^1)^2 = \inf_{\rho(t)\in\mathcal{C}} \left\{ \int_0^1 \dot{\rho}^\mathsf{T} L(\rho)^{-1} \dot{\rho} \mathrm{d}t : \rho(0) = \rho^0, \ \rho(1) = \rho^1 \right\} , \tag{6}$$

where $\mathcal{C}$ is the set of all continuously differentiable curves in $\mathcal{P}_o(S)$. Thus, $(\mathcal{P}_o(S), g^W)$ is a finite-dimensional Riemannian manifold (Li 2018). In particular, we call g^W_ρ the optimal transport metric tensor.

3.2 Fokker–Planck Equations as Evolutionary Dynamics

We shall derive (1) as a gradient flow of the free energy on the Riemannian manifold $(\mathcal{P}_o(S), g^W)$.

Theorem 4 *Given a potential game with strategy graph $G = (S, E)$, potential $\mathcal{F}(\rho) \in C^2(\mathbb{R}^n)$ and a constant $\beta \ge 0$. Then, the gradient flow of free energy*

$$-\mathcal{F}(\rho) + \beta \sum_{i=1}^n \rho_i \log \rho_i$$

on the Riemannian manifold $(\mathcal{P}_o(S), g^W)$ is the Fokker–Planck equation

$$\frac{d\rho_i}{dt} = \sum_{j\in N(i)} \frac{1}{d_j}\rho_j[F_i(\rho) - F_j(\rho) + \beta(\log\rho_j - \log\rho_i)]_+ \\ - \sum_{j\in N(i)} \frac{1}{d_i}\rho_i[F_j(\rho) - F_i(\rho) + \beta(\log\rho_i - \log\rho_j)]_+ \ ,$$

for any $i \in S$. In addition, for any initial $\rho^0 \in \mathcal{P}_o(S)$, there exists a unique solution $\rho(t) : [0,\infty) \to \mathcal{P}_o(S)$. And the free energy is a Lyapunov function. Moreover, if $\rho^\infty = \lim_{t\to\infty}\rho(t)$ exists, ρ^∞ is one of the Gibbs measures satisfying (2).

Remark 1 We note that if $\beta = 0$ and G is a complete graph, the derived Fokker–Planck equation is the Smith dynamic (1984) by dividing a constant ratio n.

Remark 2 The strategy graph G is different from the one in evolutionary graph games studied in Allen and Nowak (2014), Lieberman et al. (2005), Szabo and Fath (2007). They mainly consider a spatial space as the graph, while our graph relates to the strategy set.

Proof We show that (1) is a gradient flow. For any $\sigma \in T_\rho\mathcal{P}_o(S)$, there exists Φ, such that $\sigma = -\mathrm{div}(\rho\nabla\Phi)$. Since $\frac{\mathrm{d}\rho}{\mathrm{d}t} = \left(\frac{\mathrm{d}\rho_i}{\mathrm{d}t}\right)_{i=1}^n$ is in $T_\rho\mathcal{P}_o(S)$. By Definition 3, we have

$$g_\rho^W\left(\frac{\mathrm{d}\rho}{\mathrm{d}t}, \sigma\right) = \sum_{i=1}^n \frac{\mathrm{d}\rho_i}{\mathrm{d}t}\Phi_i \ . \tag{7}$$

On the other hand, we have

$$\mathrm{d}\bar{\mathcal{F}}(\rho)\cdot\sigma = \sum_{i=1}^n \frac{\partial}{\partial\rho_i}\bar{\mathcal{F}}(\rho)\cdot\sigma_i = -\sum_{i=1}^n \bar{F}_i(\rho)\mathrm{div}(\rho\nabla\Phi)_i \\ = (\nabla\bar{F}(\rho), \nabla\Phi)_\rho = -\sum_{i=1}^n \Phi_i\mathrm{div}(\rho\nabla\bar{F}(\rho))_i \ . \tag{8}$$

Combining (7) and (8), and the definition of gradient flow of $-\bar{\mathcal{F}}(\rho)$ on the manifold, we obtain

$$0 = g_\rho^W\left(\frac{\mathrm{d}\rho}{\mathrm{d}t}, \sigma\right) - \mathrm{d}\bar{\mathcal{F}}(\rho)\cdot\sigma \\ = \sum_{i=1}^n \left\{\frac{\mathrm{d}\rho_i}{\mathrm{d}t} + \mathrm{div}(\rho\nabla\bar{F}(\rho))_i\right\}\Phi_i \ .$$

Since the above formula is true for all $(\Phi_i)_{i=1}^n \in \mathbb{R}^n$,

$$\frac{\mathrm{d}\rho_i}{\mathrm{d}t} + \sum_{j\in N(i)} \theta_{ij}(\rho)\big(\bar{F}_j(\rho) - \bar{F}_i(\rho)\big) = 0$$

holds for all $i \in V$. Substituting θ_{ij} defined in (3) into the above formula, we derive (1). The rest of the proof are in Chow et al. (2017a), Li (2016), see details there. □

We can further extend (1) as mean dynamics to model general population games without potential. Although (1) can no longer be viewed as gradient flows of any sort in this case, yet it is a system of well-defined ordinary differential equations in $\mathcal{P}(S)$.

Corollary 5 *Given a population game with strategy graph $G = (S, E)$ and a constant $\beta \geq 0$. Assume payoff function $F : \mathcal{P}(S) \to \mathbb{R}^n$ is continuous. For any initial condition $\rho^0 \in \mathcal{P}_o(S)$, the Fokker–Planck equation*

$$\begin{aligned}\frac{d\rho_i}{dt} &= \sum_{j\in N(i)} \frac{1}{d_j}\rho_j[F_i(\rho) - F_j(\rho) + \beta(\log\rho_j - \log\rho_i)]_+ \\ &\quad - \sum_{j\in N(i)} \frac{1}{d_i}\rho_i[F_j(\rho) - F_i(\rho) + \beta(\log\rho_i - \log\rho_j)]_+ ,\end{aligned}$$

is a well-defined flow in $\mathcal{P}_o(S)$.

The proof is similar to that of Theorem 4 and hence omitted.

It is worth mentioning that, for potential games, there may exist multiple Gibbs measures as equilibria of (1). For non-potential games, there exist even more complicated phenomena than equilibria, for example, invariant sets. We illustrate this by a modified Rock–Scissors–Paper game in Sect. 5, for which Hopf bifurcation exists with respect to the parameter β.

Remark 3 (Links with existing dynamics) In the literature, there are discussions of the relation between the dynamics' rest points and Gibbs measure for various evolutionary dynamics, see Sandholm (2010). For example, Leslie and Collins (2005) study perturbed Best-response dynamics, and Coucheney et al. (2015) discuss the issue for perturbed Replicator dynamics. In these dynamics, the perturbations are driven by entropy.

We compare the proposed dynamics (FPE) with some existing game dynamics (entropy perturbed Replicator dynamics and Logit dynamics). Firstly, the rest points of FPE, entropy perturbed Replicator dynamics (Coucheney et al. 2015) and Logit dynamics, are the same, i.e., Gibbs measures. Secondly, the dynamical property of these dynamics is different. For potential games, (i) the FPE, Replicator dynamics are gradient flows, while the Logit is not; (ii) if the potential is given by entropy only, the Replicator dynamics is the Hessian flow (Newton method) in the probability set, while the FPE is not. This comes from the difference of the geometry of Shahshahani (Fisher–Rao) metric and Wasserstein metric. The Shahshahani metric is given by Hessian operator of entropy. It is a symmetric metric tensor treating all discrete strategy states equally. The Wasserstein metric is built on the transportation of measures on graphs. If the graph is not a complete graph, the Wasserstein metric tensor is not symmetric for discrete strategies, which results in asymmetrical dynamics.

We give an example for illustrating these differences. Let $n = 3$. Denote $\mathcal{F}(\rho) = -\sum_{i=1}^{3} \rho_i \log \rho_i$ and the Gibbs measure $\rho^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The Logit dynamics satisfies

$$\dot{\rho}_i = \frac{1}{3} - \rho_i, \qquad i = 1, 2, 3.$$

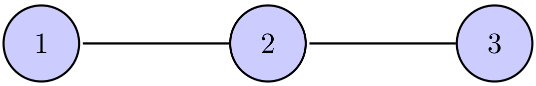
The Replicator dynamics is given by

$$\dot{\rho}_i = \rho_i \left(\log \rho_i + 1 - \sum_{i=1}^{n} \rho_i \log \rho_i \right), \qquad i = 1, 2, 3.$$

The FPE follows

$$\begin{cases} \dot{\rho}_1 &= \frac{1}{2}(\log \rho_2 - \log \rho_1)_+ \rho_2 - (\log \rho_1 - \log \rho_2)_+ \rho_1 \\ \dot{\rho}_1 &= (\log \rho_1 - \log \rho_2)_+ \rho_1 + (\log \rho_3 - \log \rho_2)_+ \rho_3 \\ & \quad -\frac{1}{2}\Big((\log \rho_2 - \log \rho_1)_+ + (\log \rho_2 - \log \rho_3)_+ \Big) \rho_2 \\ \dot{\rho}_3 &= \frac{1}{2}(\log \rho_2 - \log \rho_3)_+ \rho_2 - (\log \rho_3 - \log \rho_2)_+ \rho_3 \end{cases}.$$

for the following asymmetrical strategy graph



The vector fields of the three equations are plotted in the following figures. We observe that the vector fields of Logit and Replicator dynamics are symmetric, while the vector field of FPE depends on the structure of strategy graph. In this case, because strategy (1) and strategy (3) are disconnected, the vector field of FPE is not symmetric, i.e., strategy 2 behaves differently from strategy 1, 3. This demonstrates how the behavior of the dynamics is affected by the underlying strategy graph (Fig. 1).

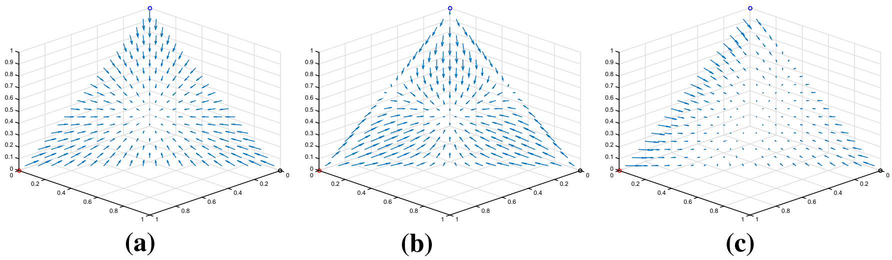


Fig. 1 Comparison of vector fields in different dynamics. **a** Logit, **b** replicator, **c** FPE

Likewise, the structure of the strategy graph determines the behavior of Smith dynamics as well, even if all states' payoff are given equally. See more examples in Sect. 5.

3.3 Markov Process

In this subsection, we look at Fokker–Planck equation (1) from the probabilistic viewpoint. More specifically, we present a Markov process whose transition function is given by (1). From the modeling perspective, such a Markov process models individual player's decision process that is myopic, irrational and locally greedy. The Markov process $X_\beta(t)$ is defined as

$$\begin{aligned}
&\Pr(X_\beta(t+h) = j \mid X_\beta(t) = i) \\
&\quad = \begin{cases} \frac{1}{d_j}(\bar{F}_j(\rho) - \bar{F}_i(\rho))_+ h + o(h)\ , & \text{if } j \in N(i)\ ; \\ 1 - \sum_{j \in N(i)} \frac{1}{d_i}(\bar{F}_j(\rho) - \bar{F}_i(\rho))_+ h + o(h)\ , & \text{if } j = i\ ; \\ 0\ , & \text{otherwise}\ , \end{cases}
\end{aligned} \tag{9}$$

where $\bar{F}_i(\rho) = F_i(\rho) - \beta \log \rho_i$ and $\lim_{h\to 0} \frac{o(h)}{h} = 0$. It can be easily seen that the probability evolution equation of $X_\beta(t)$ is exactly (1).

Process $X_\beta(t)$ characterizes players' decision-making process. Intuitively, players compare their current strategies with local strategy neighbors. If the neighboring strategy has payoff higher than their current payoffs, they switch strategies with probability proportional to the difference between the two payoffs. In addition, $X_\beta(t)$ represents an individual player's irrational behavior. This irrationality may be due to players' mistake or willingness to take risk. The uncertainty of strategy i is quantified by term $\log \rho_i$. The monotonicity of this term intuitively implies that *the fewer players currently select strategy i, the more likely players are willing to take risks by switching to strategy i*. For this interpretation, we call $F_i(\rho) - \beta \log \rho_i$ the noisy payoff of strategy i, where β is the noise level.

4 Stability

In this section, we discuss the longtime behavior of (1) for potential games. We shall study the convergence properties of dynamics (1). Our derivation is based on two concepts, discrete relative free energy and relative Fisher information (Carrillo et al. 2003). They are used to measure the closeness between two discrete measures ρ and ρ^∞, the Gibbs measure defined by (2).

The first concept is the discrete relative free energy ($\mathcal{H}$)

$$\mathcal{H}(\rho|\rho^\infty) := \beta(\bar{\mathcal{F}}(\rho^\infty) - \bar{\mathcal{F}}(\rho))\ .$$

The other is the discrete relative Fisher information ($\mathcal{I}$)

$$\mathcal{I}(\rho|\rho^{\infty}) := \sum_{(i,j)\in E} \left(\log \frac{\rho_i}{e^{F_i(\rho)/\beta}} - \log \frac{\rho_j}{e^{F_j(\rho)/\beta}} \right)_+^2 \frac{1}{d_i} \rho_i \ .$$

We remark that in finite player games, where the potential is a linear function of ρ, $\mathcal{H}$ and $\mathcal{I}$ coincide with the classical relative entropy (Kullback–Leibler divergence) and relative Fisher information respectively, e.g., let $\mathcal{F}(\rho) = \sum_{i=1}^{n} v_i \rho_i$ with $\sum_{i=1}^{n} e^{-\frac{v_i}{\beta}} = 1$, then $\rho_i^{\infty} = e^{-\frac{v_i}{\beta}}$,

$$\begin{aligned}
\mathcal{H}(\rho|\rho^{\infty}) &= \beta \sum_{i=1}^{n} \rho_i \log \frac{\rho_i}{\rho_i^{\infty}} \\
&= \beta \sum_{i=1}^{n} \rho_i \log \rho_i - \beta \sum_{i=1}^{n} \rho_i \log e^{-\frac{v_i}{\beta}} \\
&= \beta \sum_{i=1}^{n} \rho_i \log \rho_i + \sum_{i=1}^{n} v_i \rho_i .
\end{aligned}$$

We shall show that $\mathcal{H}(\rho(t)|\rho^{\infty})$ converges to 0 as t goes to infinity. We will also estimate the speed of convergence and characterize their stability properties. Before that, we state a theorem that connects $\mathcal{H}$ and $\mathcal{I}$ via gradient flow (1).

Theorem 6 *Suppose $\rho(t)$ is the transition probability of $X_\beta(t)$ of a potential game. Then, the relative entropy decreases as a function of t. In other words,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(\rho(t)|\rho^{\infty}) < 0.$$

And the dissipation of relative entropy is β times relative Fisher information

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(\rho(t)|\rho^{\infty}) = -\beta\mathcal{I}(\rho(t)|\rho^{\infty}). \tag{10}$$

Proof The proof is based on the fact that $\mathcal{H}$ (the difference between noisy potentials) decreases along the gradient flow with respect to time. Namely,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(\rho(t)|\rho^{\infty}) &= -\beta\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{F}}(\rho(t)) = \beta(\nabla\bar{F}, \nabla\bar{F})_\rho \\
&= \beta \sum_{(i,j)\in E} \left[(\bar{F}_j(\rho) - \bar{F}_i(\rho))_+\right]^2 \frac{1}{d_i}\rho_i \\
&= \beta \sum_{(i,j)\in E} \left[\left(\log \frac{\rho_i}{e^{F_i(\rho)/\beta}} - \log \frac{\rho_j}{e^{F_j(\rho)/\beta}} \right)_+\right]^2 \frac{1}{d_i} \rho_i \ .
\end{aligned} \tag{11}$$

□

This shows that the noisy potential grows at the rate equal to the relative Fisher information. In other words, the population as a whole always seeks to improve the average noisy payoff at the rate equal to the expected squared benefits. Based on Theorem 6, we show that the dynamics converges to the equilibrium exponentially fast. Here, the convergence is in the sense of $\mathcal{H}$ going to zero. Such phenomenon is called entropy dissipation.

Theorem 7 (Entropy dissipation) *Let $\mathcal{F} \in C^2(\mathcal{P}(S))$ be a concave potential function (not necessary strictly concave) for a given game. Then, there exists a constant $C = C(\rho^0, G) > 0$ such that*

$$\mathcal{H}(\rho(t)|\rho^\infty) \le e^{-Ct}\mathcal{H}(\rho^0|\rho^\infty) . \tag{12}$$

The proof of Theorem 7 is readily available by noticing the fact that

$$\mathcal{I}(\rho|\rho^\infty) \le C\beta\mathcal{H}(\rho|\rho^\infty) ,$$

and an application of Gronwall's inequality. See details in Chow et al. (2017a), Li (2016). In fact, the exponential convergence is naturally expected because (1) is the gradient flow on a Riemannian manifold $(\mathcal{P}_o(S), g^W)$.

In fact, a more precise characterization on the convergence rate C in (12) can be made. This characterization enables us to address the stability issues of Gibbs measures. Define

$$\lambda(\rho) = \min_{\Phi} \; \mathrm{div}(\rho\nabla\Phi)^{\mathrm{T}} \cdot \mathrm{Hess}(-\bar{\mathcal{F}}(\rho)) \cdot \mathrm{div}(\rho\nabla\Phi) , \tag{13}$$

where the infimum is among all $(\Phi_i)_{i=1}^n \in \mathbb{R}^n$, such that $(\nabla\Phi, \nabla\Phi)_\rho = 1$ and Hess represents the Hessian operator in $\mathbb{R}^n$.

Theorem 8 (Stability and asymptotic convergence rate) *For a potential game with potential $\mathcal{F}(\rho) \in C^2$. Denote its Gibbs measure ρ^∞ by (2). If $\lambda(\rho^\infty) > 0$, then ρ^∞ is an asymptotic stable equilibrium for (1). In addition, for any sufficiently small $\epsilon > 0$, there exists a time $T > 0$, such that when $t > T$,*

$$\mathcal{H}(\rho(t)|\rho^\infty) \le e^{-2(\lambda(\rho^\infty)-\epsilon)(t-T)}\mathcal{H}(\rho^0|\rho^\infty) . \tag{14}$$

Theorem 8 can be proved by utilizing standard techniques from dynamical systems. Namely,

(i) Calculate the second order derivative of $\mathcal{F}(\rho(t))$ with respect to time t.

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\bar{\mathcal{F}}(\rho(t)) = 2\mathrm{div}(\rho\nabla\bar{F}(\rho))^{\mathrm{T}} \cdot \mathrm{Hess}\bar{\mathcal{F}}(\rho) \cdot \mathrm{div}(\rho\nabla\bar{F}(\rho)) + o\left(\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\rho(t))\right) . \tag{15}$$

(ii) Compare the first and second derivative to have

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\bar{\mathcal{F}}(\rho(t)) \le -\lambda(\rho^\infty)\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{F}}(\rho(t)) + o\left(\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{F}}(\rho(t))\right) ,$$

and apply Gronwall's inequality to show (14) and (12).

The crucial part of the proof is to establish (15), which is given below. For complete details, see Chow et al. (2017a).

Proof of (15) The first derivative of the free energy along (1) is

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{F}}(\rho(t)) = \sum_{(i,j)\in E} [(\bar{F}_j - \bar{F}_i)_+]^2 \frac{1}{d_i}\rho_i \; .$$

The second derivative of the free energy can be calculated by using the product rule:

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\bar{\mathcal{F}}(\rho(t)) = \sum_{(i,j)\in E} [(\bar{F}_j - \bar{F}_i)_+]^2 \frac{\mathrm{d}\rho_i}{\mathrm{d}t} \tag{T1}$$

$$+ 2\sum_{(i,j)\in E} \left(\frac{\mathrm{d}\bar{F}_j}{\mathrm{d}t} - \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}\right)(\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i \; . \tag{T2}$$

Since $\rho(t)$ is assumed to converge to an equilibrium ρ^∞ and the boundary is a repeller (Theorem 4), we know that $\frac{\mathrm{d}\rho}{\mathrm{d}t} \to 0$, while $\rho_i(t) \geq c(\rho^0) > 0$. Hence, T1 is a high-order term of the first derivative, i.e.,

$$\mathrm{T1} = o\left(\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mathcal{F}}(\rho(t))\right) \; .$$

On the other hand,

$$\begin{aligned}
\mathrm{T2} &= 2\sum_{(i,j)\in E} \left(\frac{\mathrm{d}\bar{F}_j}{\mathrm{d}t} - \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}\right)(\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i \\
&= 2\sum_{(i,j)\in E} \frac{\mathrm{d}\bar{F}_j}{\mathrm{d}t}(\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i - 2\sum_{(i,j)\in E} \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}(\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i \\
&= 2\sum_{(j,i)\in E} \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}(\bar{F}_i - \bar{F}_j)_+ \frac{1}{d_j}\rho_j - 2\sum_{(i,j)\in E} \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}(\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i \\
&= 2\sum_{i=1}^{n} \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t} \sum_{j\in N(i)} \left\{(\bar{F}_i - \bar{F}_j)_+ \frac{1}{d_j}\rho_j - (\bar{F}_j - \bar{F}_i)_+ \frac{1}{d_i}\rho_i\right\} \\
&= 2\sum_{i=1}^{n} \frac{\mathrm{d}\bar{F}_i}{\mathrm{d}t}\frac{\mathrm{d}\rho_i}{\mathrm{d}t} = 2\frac{\mathrm{d}\bar{F}}{\mathrm{d}t} \cdot \frac{\mathrm{d}\rho}{\mathrm{d}t} \\
&= 2 \cdot \left(\frac{\mathrm{d}\rho}{\mathrm{d}t}\right)^{\mathrm{T}} \cdot \mathrm{Hess}\bar{\mathcal{F}}(\rho) \cdot \frac{\mathrm{d}\rho}{\mathrm{d}t} \\
&= 2 \cdot \mathrm{div}(\rho\nabla\bar{F}(\rho))^{\mathrm{T}} \cdot \mathrm{Hess}\bar{\mathcal{F}}(\rho) \cdot \mathrm{div}(\rho\nabla\bar{F}(\rho)) \; ,
\end{aligned}$$

where the third equality is by relabeling i and j and the last equality is from the alternative representation of (1), i.e.,

$$\frac{d\rho_i}{dt} = -\mathrm{div}(\rho \nabla \bar{F}(\rho)) \ .$$

□

The rest of proof is to compare (T1) and (T2), see details in Chow et al. (2017a).

Remark 4 (Link with current works) It is known that the Smith dynamics converge to equilibrium in potential games. See Sandholm (2010) and references therein. We demonstrate that how the convergence rate depends on the graph structure. This result shares many similar properties with continuous cases.

5 Examples

In this section, we investigate (1) by applying it to several well-known population games.

Example 1 *Stag Hunt.* The point we seek to convey in this example is that the noisy payoff reflects the *rationality* of the population. The symmetric normal-form game with payoff matrix

$$A = \begin{pmatrix} h & h \\ 0 & s \end{pmatrix}$$

is known as Stag Hunt game. Each player in a random match needs to decide whether to hunt for a hare (h) or stag (s). Assume $s \geq h$, which means that the payoff of a stag is larger than a hare. This population game has three Nash equilibria: two pure equilibria $(0, 1)$, $(1, 0)$, and one mixed equilibrium $(1 - \frac{h}{s}, \frac{h}{s})$.

In particular, let $h = 2$ and $s = 3$. The population state is $\rho = (\rho_h, \rho_s)^{\mathrm{T}}$ with payoff $F_h(\rho) = 2\rho_h$ and $F_s(\rho) = 3\rho_s$. Then, Fokker–Planck equation (1) becomes

$$\begin{cases} \dot{\rho}_s = \rho_s[2\rho_h - 3\rho_s + \beta \log \rho_s - \beta \log \rho_h]_+ - \rho_h[-2\rho_h + 3\rho_s + \beta \log \rho_h - \beta \log \rho_s]_+ \\ \dot{\rho}_s = \rho_h[3\rho_s - 2\rho_h + \beta \log \rho_h - \beta \log \rho_s]_+ - \rho_s[-3\rho_s + 2\rho_h + \beta \log \rho_s - \beta \log \rho_h]_+ \ . \end{cases}$$

The numerical results are in Fig. 2. One can easily see that if the noise level β is sufficient small, the perturbation does not affect the limit behavior of the mean dynamics. On the other hand, if noise level β is large enough, (1) settles around $(\frac{1}{2}, \frac{1}{2})$. Lastly, if the noise level is moderate, Equation (1) has $(1, 0)$ as the unique equilibrium.

The above observation has practical meanings. Namely, if the perturbation is large enough, it turns out that people always choose to hunt hare (NE $(1, 0)$). This is a safe choice as players can get at least a hare, no matter how the others behave. This appears even more so if comparing with the state $(0, 1)$ for which the player receives nothing. If the perturbation is small and initial population appears to be more cooperative, people will choose to hunt the stag. This is a rational move because stag is definitely better than hare.

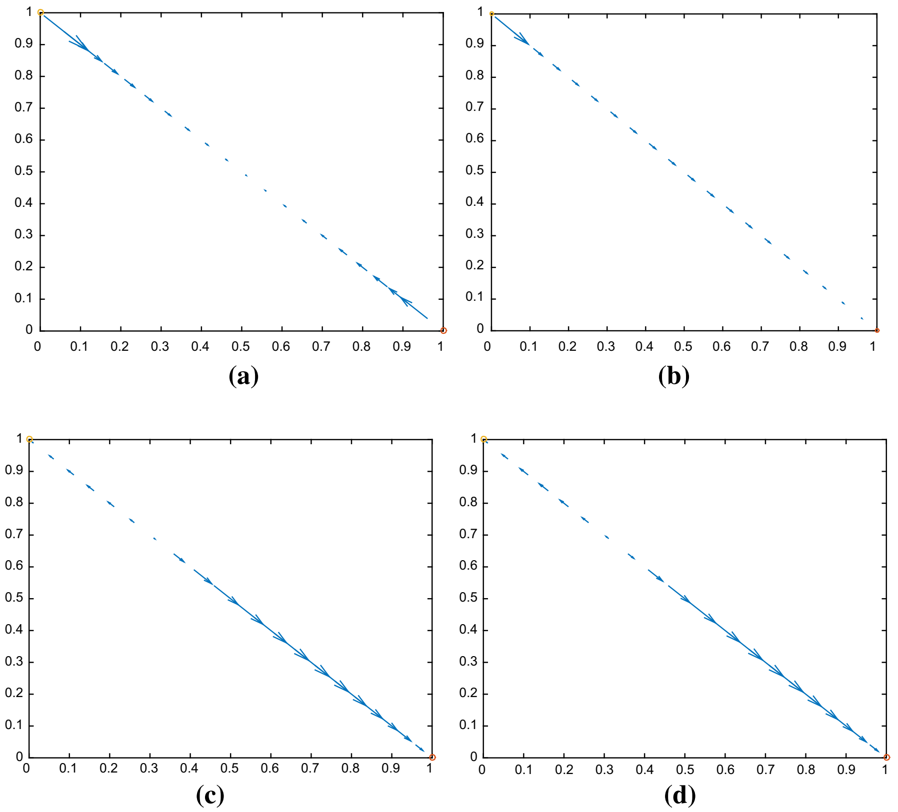


Fig. 2 Stag and Hare. **a** $\beta = 5$. **b** $\beta = 0.5$. **c** $\beta = 0.1$. **d** $\beta = 0$

Example 2 *Rock–Scissors–Paper game*. Rock–Scissors–Paper has payoff matrix

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}.$$

The strategy set is $S = \{r, s, p\}$. The population state is $\rho = (\rho_r, \rho_s, \rho_p)^{\mathrm{T}}$, and the payoff functions are $F_r(\rho) = \rho_s - \rho_p$, $F_s(\rho) = -\rho_r + \rho_p$ and $F_p(\rho) = \rho_r - \rho_s$. By solving (1), we find that there is one unique Nash equilibrium around $\rho^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for various βs. The result can be found in Fig. 3.

Example 3 We show an example with Hopf Bifurcation. Consider a modified Rock–Scissors–Paper game with payoff matrix

$$A = \begin{pmatrix} 0 & 2 & -1 \\ -1 & 0 & 2 \\ 2 & -1 & 0 \end{pmatrix}.$$

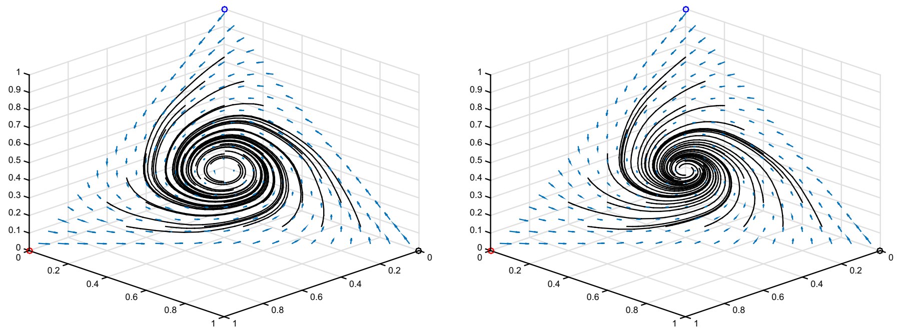


Fig. 3 Rock–Scissors–Paper. **a** $\beta = 0$. **b** $\beta = 0.1$

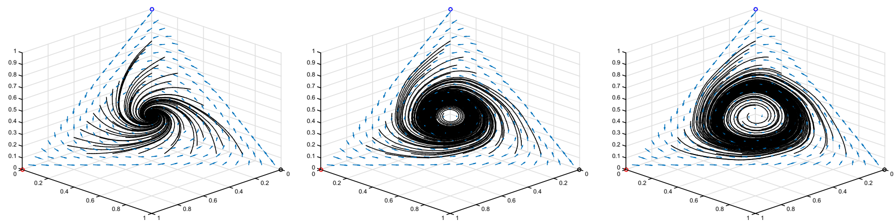
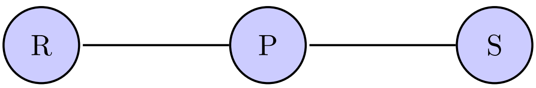


Fig. 4 Modified Rock–Scissors–Paper. **a** $\beta = 0.5$. **b** $\beta = 0.1$. **c** $\beta = 0$

Different from the previous example, the payoff functions are $F_r(\rho) = 2\rho_s - \rho_p$, $F_s(\rho) = -\rho_r + 2\rho_p$ and $F_p(\rho) = 2\rho_r - \rho_s$. We find that there is Hopf bifurcation for Equation (1). If β is large, there is a unique equilibrium around $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^{\mathrm{T}}$. If β goes to 0, the solution approaches to a limit cycle. The results are in Fig. 4.

Also, we illustrate the effect of graph structure in proposed FPEs. In Fig. 5, we consider two strategy graphs: One is the complete graph; the other is the lattice graph



In other words, we cut off one edge for the complete strategy. This consideration results in the asymmetric property of vector fields. It shows the difference among FPEs, Logit and perturbed Replicator dynamics with $\beta = 0.1$ (Fig. 6).

Example 4 We show an example with multiple Gibbs measures. Consider a potential game with payoff matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

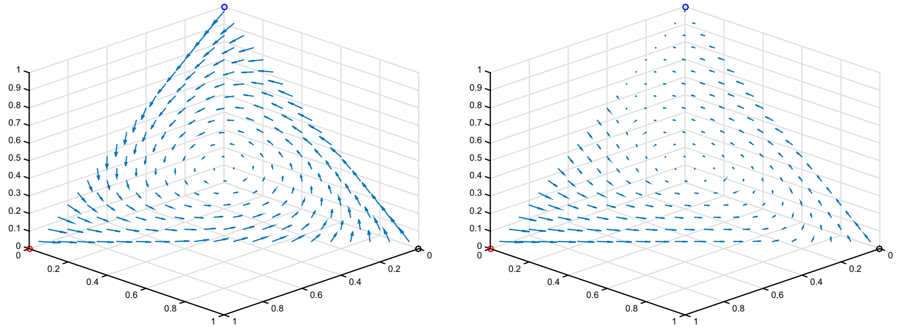


Fig. 5 Left: Fokker–Planck equation on a complete strategy graph, Right: Fokker–Planck equation on a lattice graph

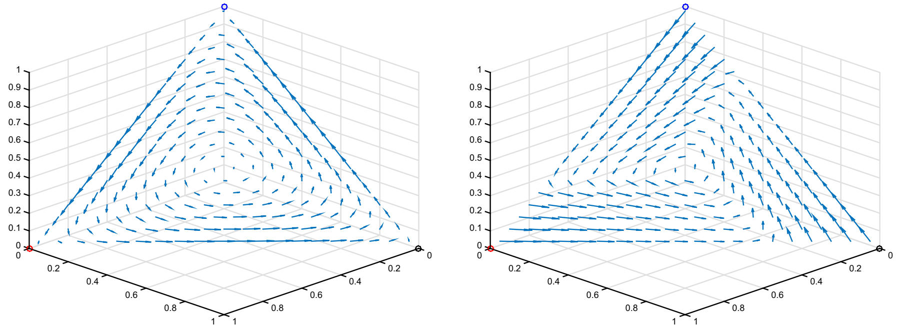


Fig. 6 Left: replicator dynamics with entropy perturbation, Right: logit dynamics

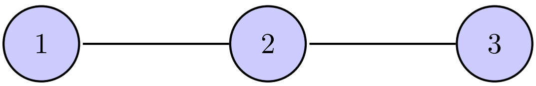
Denote the strategy set as $S = \{1, 2, 3\}$. The population state is $\rho = (\rho_1, \rho_2, \rho_3)^{\mathrm{T}}$, and the payoff functions are $F_1(\rho) = \rho_1$, $F_2(\rho) = \rho_2 + \rho_3$ and $F_3(\rho) = \rho_2 + \rho_3$. We consider three sets of Nash equilibria :

$$\left\{\rho \mid \rho_1 = \frac{1}{2}\right\} \cup \{(1, 0, 0)\} \cup \{\rho \mid \rho_1 = 0\} ,$$

where the first and third ones are lines on the probability simplex $\mathcal{P}(S)$. By applying (1) on a complete graph, we obtain two Gibbs measures

$$\left\{\left(0, \frac{1}{2}, \frac{1}{2}\right)\right\} \cup \{(1, 0, 0)\}$$

as $\beta \to 0$. The vector field is shown in Fig. 7. Similarly, we illustrate the effect of graph structure for potential games. We also consider two strategy graphs: One is the complete graph, the other is the lattice graph



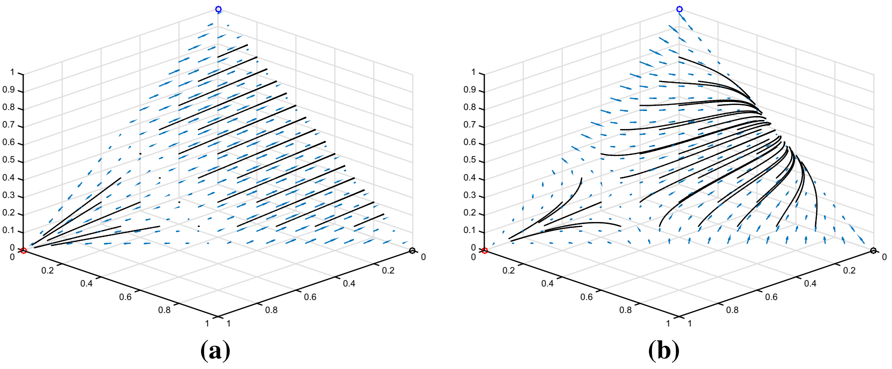


Fig. 7 Multiple Gibbs measures **a** $\beta = 0$. **b** $\beta = 0.1$

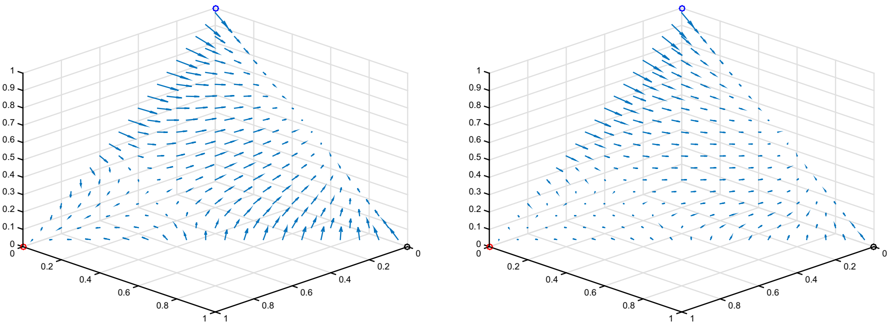


Fig. 8 Left: Fokker–Planck equation on a complete strategy graph, Right: Fokker–Planck equation on a lattice graph

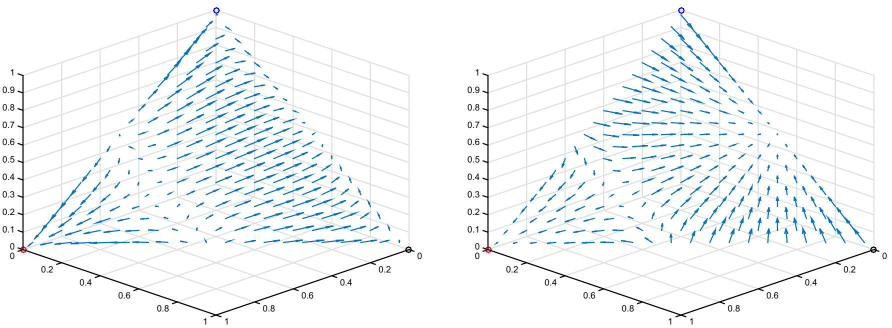


Fig. 9 Left: replicator dynamics with entropy perturbation. Right: logit dynamics

We observe the asymmetric property of vector fields. This property shows the difference among FPEs, Logit and perturbed Replicator dynamics with $\beta = 0.1$. See Figs. 8 and 9.

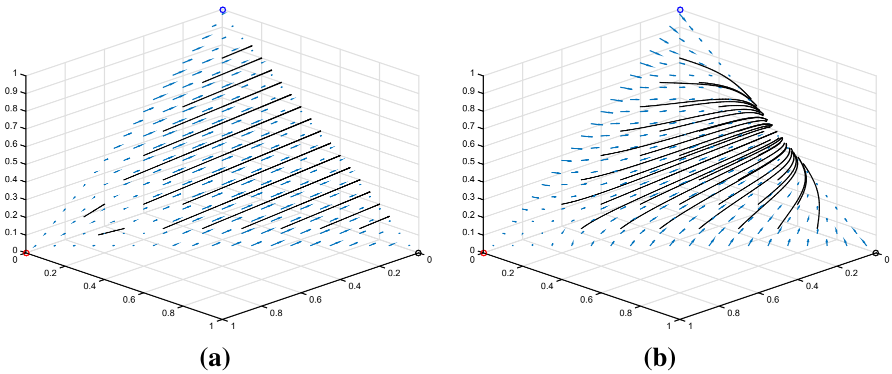


Fig. 10 Unique Gibbs measures. **a** $\beta = 0$. **b** $\beta = 0.1$

Example 5 As a completion, we introduce a game with unique Gibbs measure. Let us consider another potential game with payoff matrix

$$A = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Here, the strategy set is $S = \{1, 2, 3\}$, the population state is $\rho = (\rho_1, \rho_2, \rho_3)^{\mathrm{T}}$, and the payoff functions are $F_1(\rho) = \frac{1}{2}\rho_1$, $F_2(\rho) = \rho_2 + \rho_3$ and $F_3(\rho) = \rho_2 + \rho_3$. There are three sets of Nash equilibria

$$\left\{\rho \mid 1 - \frac{1}{2}\rho_1 = \rho_2 + \rho_3\right\} \cup \{(1, 0, 0)\} \cup \{\rho \mid 1 = \rho_2 + \rho_3\},$$

By applying Fokker–Planck equation (1) on a complete graph, we have a unique Gibbs measure

$$\left(0, \frac{1}{2}, \frac{1}{2}\right)$$

as $\beta \to 0$. See Fig. 10 for the vector fields.

6 Conclusion

In this paper, we proposed a dynamics for population games utilizing optimal transport theory and mean field games. Comparing to existing models, it has the following desirable features. Firstly, the dynamics is the gradient flow of the noisy potential in the probability space endowed with the optimal transport metric. The dynamics can also be seen as the mean field type Fokker–Planck equations. Secondly, the dynamics is the probability evolution equation of a Markov process. Such processes model players' myopicity, greediness and irrationality. In particular, the irrational behaviors or uncertainties are introduced via the notion of noisy payoff. This shares many similarities with the diffusion or white noise perturbation in continuous cases. Last but not least, for potential games, Gibbs measures are equilibria of the dynamics. Their stability properties are easily obtained by the relation between relative free energy and Fisher information. In general, the dynamics may exhibit more complicated limiting behaviors, including Hopf bifurcations.

We would continue to bridge the communities between optimal transport and population games. On the one hand, the evolutionary game theory provides broad application fields for optimal transport. It introduces various modeling perspective. On the other hand, optimal transport introduces the other mathematical structures for games. It gives the game-dependent Riemannian metric tensor. The metric tensor relies on the graph structure of discrete strategy set. Many questions intersecting both communities arise, e.g., what are dynamical properties of FPEs related to this metric? What is the effect of the strategy graph for the stability issues of NEs? We will continue to work on these problems in future.

Acknowledgements This paper is based on Wuchen Li's thesis Li (2016).

Appendix

In this section, we briefly review the Best-reply dynamics and its connection with optimal transport theory. These serve the motivations of the dynamics considered in this paper. For more details see Degond et al. (2014), Villani (2008).

Best-reply dynamics and Fokker–Planck equations We first consider a game consisting N players $i \in \{1, \ldots, N\}$. Each player i chooses a strategy x_i from a same Borel strategy set S. For concreteness, we consider $S = \mathbb{T}^d$, which is a d dimensional torus. Suppose each player receives a payoff function $F_i \in C^\infty(S)$. For notational convenience, we denote $F_i(x_i, x_{-i}) = F_i(x_1, \ldots, x_N)$, where we abuse the notation by

$$x_{-i} = \{x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N\} .$$

We model players' decision-making processes in a game by stochastic process $x_i(t)$, $t \in [0, +\infty)$. Here, t is an artificial time variable, at which player i selects his/her decision based on the current strategies of all other players $x_{-i}(t)$. We note that all players make their decisions simultaneously and without knowing others' decisions. Each player selects his or her strategy that increases the player's payoff most rapidly. In other words, we model the game by the following stochastic differential equations (SDEs)

$$\mathrm{d}x_i = \nabla_{x_i} F_i(x_i, x_{-i})\mathrm{d}t + \sqrt{2\beta} dB_t^i \ , \tag{16}$$

where the independent Brownian motion $(B_t^i)_{i=1}^N$ is added to model the uncertainty of each player and $\beta > 0$ controls the magnitude of the noise.

Under the standard assumptions in population games, i.e., the game is autonomous and the players are symmetric, one can simply encode all the information of players into one probability density $\rho \in \mathcal{P}(S)$ by taking $N \to \infty$. In this limiting processes, each player's cost function is rewritten as $F\colon S \times \mathcal{P}(S) \to \mathbb{R}$, and the limiting stochastic process forms the following mean field SDE

$$\mathrm{d}X_t = \nabla_{X_t} F(X_t, \rho)\mathrm{d}t + \sqrt{2\beta} dB_t^i \ , \tag{17}$$

where X_t has probability law $\rho(t, x)$.

In Degond et al. (2014), SDE (17) is called the Best-reply dynamics, and X_t is the Best-reply decision process. Here, the transition density function $\rho(t, x)$ of the stochastic process $X(t)$ satisfies the FPE

$$\frac{\partial \rho(t, x)}{\partial t} = -\nabla \cdot (\rho(t, x) F(x, \rho)) + \beta \Delta \rho(t, x) \ . \tag{18}$$

The game is called a potential game if there exists a potential function $\mathcal{F}\colon \mathcal{P}(S) \to \mathbb{R}$, such that $\frac{\delta}{\delta \rho(x)} \mathcal{F}(\rho) = F(x, \rho)$. For potential games, the Best-reply SDE (17) becomes

$$\mathrm{d}X_t = \nabla \frac{\delta}{\delta \rho(t, X_t)} \mathcal{F}(\rho)\mathrm{d}t + \sqrt{2\beta} dB_t \ ,$$

which is a perturbed gradient flow and whose transition equation (FPE) forms

$$\frac{\partial \rho(t, x)}{\partial t} = -\nabla \cdot \left(\rho(t, x) \nabla \frac{\delta}{\delta \rho(t, x)} \mathcal{F}(\rho)\right) + \beta \Delta \rho(t, x) \ . \tag{19}$$

From the theory of optimal transport, Equation (19) can be interpreted as a gradient ascend flow of the free energy

$$\bar{\mathcal{F}}(\rho) = \mathcal{F}(\rho) - \beta \int_S \rho(x) \log \rho(x) \mathrm{d}x \ . \tag{20}$$

Optimal transport and density manifold We next review the geometry of optimal transport on the continuous strategy set S.

Consider the set $\mathcal{P}_2(S)$ of Borel measurable probability density functions on S with finite second moment. Given $\rho^0, \rho^1 \in \mathcal{P}_2(S)$, the L^2-Wasserstein distance between ρ^0 and ρ^1 is denoted by $W\colon \mathcal{P}_2(S) \times \mathcal{P}_2(S) \to \mathbb{R}_+$. There are two equivalent ways of defining this distance.

The first definition is the following linear programming formulation:

$$W(\rho^0, \rho^1)^2 = \inf_{\pi \in \Pi(\rho^0, \rho^1)} \int_{\Omega \times \Omega} d_\Omega(x, y)^2 \pi(\mathrm{d}x, \mathrm{d}y) \ , \tag{21}$$

where the infimum is taken over the set Π of joint probability measures on $\Omega \times \Omega$ that have marginals ρ^0, ρ^1.

The second definition considers a probability path $\rho : [0, 1] \to \mathcal{P}_2(S)$ connecting ρ^0, ρ^1. And the distance is defined by a variational problem known as the Benamou–Brenier formula:

$$W(\rho^0, \rho^1)^2 = \inf_{\Phi} \int_0^1 \int_{\Omega} (\nabla\Phi(t,x), \nabla\Phi(t,x))\rho(t,x)\mathrm{d}x\mathrm{d}t\ , \tag{22a}$$

where the infimum is taken over the set of Borel *potential* functions $[0, 1] \times S \to \mathbb{R}$. Each potential function Φ determines a corresponding density path ρ as the solution of the *continuity equation*

$$\frac{\partial\rho(t,x)}{\partial t} + \mathrm{div}(\rho(t,x)\nabla\Phi(t,x)) = 0\ , \quad \rho(0,x) = \rho^0(x)\ , \quad \rho(1,x) = \rho^1(x)\ . \tag{22b}$$

Here, div and ∇ are the divergence and gradient operators in Ω. The continuity equation is known as the probability density transition equation according to the given vector field.

The equivalence between the static (21) and dynamical (22) formulations is well known. Moreover, the variational formulation (22) entails a similar Riemannian structure used in this paper. For simplicity, we only consider the set of smooth and strictly positive probability densities

$$\mathcal{P}_+(S) = \left\{\rho \in C^\infty(\Omega) \colon \rho(x) > 0\ , \ \int_\Omega \rho(x)\mathrm{d}x = 1\right\} \subset \mathcal{P}_2(S)\ .$$

Denote $\mathcal{F}(S) := C^\infty(S)$ the set of smooth real valued functions on S. The tangent space of $\mathcal{P}_+(S)$ is given by

$$T_\rho\mathcal{P}_+(S) = \left\{\sigma \in \mathcal{F}(S) \colon \int_S \sigma(x)\mathrm{d}x = 0\right\}\ .$$

Given $\Phi \in \mathcal{F}(S)$ and $\rho \in \mathcal{P}_+(S)$, define

$$V_\Phi(x) := -\mathrm{div}(\rho(x)\nabla\Phi(x))\ .$$

Thus, $V_\Phi \in T_\rho\mathcal{P}_+(S)$. The elliptic operator $\nabla \cdot (\rho\nabla)$ identifies the function Φ on S modulo additive constants with the tangent vector V_Φ of the space of densities. This gives an isomorphism

$$\mathcal{F}(S)/\mathbb{R} \to T_\rho\mathcal{P}_+(S); \quad \Phi \mapsto V_\Phi\ .$$

Define the Riemannian metric (inner product) on the tangent space of positive densities $g^W \colon T_\rho\mathcal{P}_+(S) \times T_\rho\mathcal{P}_+(S) \to \mathbb{R}$ by

$$g^W_\rho(V_\Phi, V_{\tilde{\Phi}}) = \int_S (\nabla\Phi(x), \nabla\tilde{\Phi}(x))\rho(x)\mathrm{d}x\ ,$$

where $\Phi(x), \tilde{\Phi}(x) \in \mathcal{F}(S)/\mathbb{R}$. The inner product endows $\mathcal{P}_+(S)$ with an infinite-dimensional Riemannian metric tensor. In other words, the variational problem (22) is a geometric action energy in $(\mathcal{P}_+(S), g^W)$.

We are now ready to present the gradient operator of free energy w.r.t. L^2-Wasserstein metric tensor. Following

$$g^W(\mathrm{grad}_W \bar{\mathcal{F}}(\rho), V_\Phi) = \int_S \frac{\delta}{\delta\rho(x)} \bar{\mathcal{F}}(\rho) V_\Phi \mathrm{d}x$$

and $\frac{\delta}{\delta\rho(x)}\mathcal{F}(\rho) = F(x, \rho)$, and noticing $\frac{\delta}{\delta\rho(x)} \int_S \rho(x) \log \rho(x) \mathrm{d}x = \log \rho(x) + 1$, we obtain

$$\mathrm{grad}_W \bar{\mathcal{F}}(\rho) = -\nabla \cdot (\rho \nabla (F(x, \rho) - \beta \log \rho(x))) .$$

From the fact that $\nabla \cdot (\rho \nabla \log \rho) = \nabla \cdot (\nabla \rho) = \Delta\rho$, we derive FPE (19) by the gradient flow of the free energy

$$\frac{\partial\rho}{\partial t} = \mathrm{grad}_W \mathcal{F}(\rho) = -\nabla \cdot (\rho \nabla F(x, \rho)) + \beta \Delta\rho.$$

References

Akin, E.: The Geometry of Population Genetics, vol. 280. Springer, Berlin (1979)

Allen, B., Nowak, M.A.: Games on graphs. EMS Surv. Math. Sci. **1**(1), 113–151 (2014)

Blanchet, A., Carlier, G.: Optimal transport and Cournot-Nash equilibria. arXiv preprint arXiv:1206.6571 (2012)

Blanchet, A., Carlier, G.: From Nash to Cournot–Nash equilibria via the Monge–Kantorovich problem. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. **372**(2028), 20130398 (2014)

Cardaliaguet, P.: Notes on mean field games, Technical report (2010)

Carrillo, J.A., McCann, R.J., Villani, C.: Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. Revista Matematica Iberoamericana **3**(19), 971–1018 (2003)

Chow, S.-N., Huang, W., Li, Y., Zhou, H.: Fokker–Planck equations for a free energy functional or Markov process on a graph. Arch. Ration. Mech. Anal. **203**(3), 969–1008 (2012)

Chow, S.-N., Li, W., Zhou, H.: Entropy dissipation of Fokker–Planck equations on graphs. arXiv:1701.04841 (2017a)

Chow, S.-N., Li, W., Lu, J., Zhou, H.: Game theory and discrete optimal transport. arXiv:1703.08442 (2017b)

Coucheney, P., Gaujal, B., Mertikopoulos, P.: Penalty-regulated dynamics and robust learning procedures in games. Math. Oper. Res. **40**(3), 513–796 (2015)

Degond, P., Liu, J.-G., Ringhofer, C.: Large-scale dynamics of mean-field games driven by local Nash equilibria. J. Nonlinear Sci. **24**(1), 93–115 (2014)

Erbar, M., Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. Arch. Ration. Mech. Anal. **206**(3), 997–1038 (2012)

Frieden, B.R.: Science from Fisher Information: A Unification. Cambridge University Press, Cambridge (2004)

Fudenberg, D., Levine, D.K.: The Theory of Learning in Games. MIT Press, Cambridge (1998)

Hofbauer, J., Sigmund, K.: The Theory of Evolution and Dynamical Systems: Mathematical Aspects of Selection. Cambridge University Press, Cambridge (1988)

Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. Bull. Am. Math. Soc. **40**(4), 479–519 (2003)

Hofbauer, J., Sandholm, W.H.: On the global convergence of stochastic fictitious play. Econometrica **70**, 2265–2294 (2002)

Hofbauer, J., Sandholm, W.H.: Evolution in games with randomly disturbed payoffs. J. Econ. Theory **132**, 47–69 (2007)

Huang, Y., Hao, Y., Wang, M., Zhou, W., Wu, Z.: Optimality and stability of symmetric evolutionary games with applications in genetic selection. Math. Biosci. Eng. **12**(3), 503–523 (2015)

Lasry, J.-M., Lions, P.-L.: Mean field games. Jpn. J. Math. **2**(1), 229–260 (2007)

Li, W.: A study of stochastic differential equations and Fokker–Planck equations with applications. Ph.D thesis (2016)
Li, W.: Geometry of probability simplex via optimal transport. arXiv:1803.06360 (2018)
Lieberman, E., Hauert, C., Nowak, M.A.: Evolutionary dynamics on graphs. Nature **433**(7023), 312–316 (2005)
David, S.L., Collins, E.J.: Individual Q-learning in normal form games. SIAM J. Control Optim. **44**(2), 495–514 (2005)
Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011)
Matsui, A.: Best response dynamics and socially stable strategies. J. Econ. Theory **57**(2), 343–362 (1992)
Mertikopoulos, P., Sandholm, W.H.: Riemannian game dynamics. arXiv preprint arXiv:1603.09173 (2016)
Mertikopoulos, P., Sandholm, W.H.: Learning in games via reinforcement and regularization. Math. Oper. Res. **41**(4), 1297–1324 (2016)
Monderer, D., Shapley, L.S.: Potential games. Games Econ. Behav. **14**(1), 124–143 (1996)
Nash, J.F.: Equilibrium points in n-person games. Proc. Natl. Acad. Sci. **36**(1), 48–49 (1950)
Nowak, M.A.: Evolutionary Dynamics. Harvard University Press, Harvard (2006)
Sandholm, W.H.: Population Games and Evolutionary Dynamics. MIT Press, Cambridge (2010)
Sandholm, W.H.: Evolutionary Game Theory. In: Meyers, R. (ed) Encyclopedia of Complexity and Systems Science, pp. 3176–3205. Springer, New York, NY (2012a)
Sandholm, W.H.: Decompositions and potentials for normal form games. Games Econ. Behav. **70**(2), 446–456 (2010)
Sandholm, W.H.: Local stability of strict equilibria under evolutionary game dynamics. J. Dyn. Games **5**(1), 27–50 (2012b)
Shah, D., Shin, J.: Dynamics in congestion games. In: ACM SIGMETRICS Performance Evaluation Review, vol. 38, pp. 107–118. ACM (2010)
Shahshahani, S.: A new mathematical framework for the study of linkage and selection. Mem. Am. Math. Soc. **211** (1979). https://doi.org/10.1090/memo/0211
Sigmund, K., Nowak, M.A.: Evolutionary game theory. Curr. Biol. **9**(14), R503–R505 (1999)
Michael, J.S.: The stability of a dynamic model of traffic assignment-an application of a method of Lyapunov. Transp. Sci. **18**(3), 245–252 (1984)
Szabo, G., Fath, G.: Evolutionary games on graphs. Phys. Rep. **446**(4), 97–216 (2007)
Villani, C.: Topics in Optimal Transportation, vol. 58. American Mathematical Society, Providence (2003)
Villani, C.: Optimal Transport: Old and New, vol. 338. Springer, Berlin (2008)
Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior (60 Anniversary Commemorative Edition). Princeton University Press, Princeton (2007)
Wu, A., Liao, D., Tlsty, T.D., Sturm, J.C., Austin, R.H.: Game theory in the death galaxy: interaction of cancer and stromal cells in tumour microenvironment. Interface Focus **4**(4), 20140028 (2014)