SIAM J. APPL. MATH.
Vol. 80, No. 1, pp. 142–159
© 2020 Society for Industrial and Applied Mathematics

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

EQUILIBRIUM SELECTION VIA OPTIMAL TRANSPORT*

SHUI-NEE CHOW†, WUCHEN LI‡, JUN LU†, AND HAOMIN ZHOU†

Abstract. We propose a new dynamical model, inspired by optimal transport theory and Fisher information, for finite player discrete strategy games. For a potential game, the model is a gradient flow, known as Fokker–Planck equations (FPEs), in a probability simplex set equipped with an optimal transport metric. Based on FPEs, we introduce the best-reply Markov processes, which model players' myopia, greed, and uncertainty when making decisions. The stationary measure of the dynamics provides each pure Nash equilibrium a probability by which it is ranked. We extend the proposed dynamical model to rank/select equilibria for nonpotential games as well.

Key words. game theory, optimal transport, gradient flow, entropy, Fisher information

AMS subject classifications. 91Axx; 34Axx

DOI. 10.1137/18M1163828

1. Introduction. Game theory plays vital roles in economics, biology, social network, etc. [15, 21, 17, 18]. It involves models of conflict and cooperation between rational decision makers. Each player in a game optimizes his/her own objective function. Nash equilibrium (NE) describes a status that no player is willing to change his/her strategy unilaterally. A fundamental question in game theory is that if there are multiple pure NEs, how can one select/rank them? This problem has been studied previously using various approaches. For example, in [10, 11], NEs were selected by refining the concept of equilibrium, such as payoff dominance or risk dominance principle. Another class of approaches uses learning dynamics by assuming that the players have bounded knowledge and they need to "learn" from what occurred in previous stages of the game and then respond to other players' strategies. In these settings, irrationalities of individual players are often considered. Such examples include fictitious play, no-regret dynamics, replicator dynamics, logit dynamics, and best-response dynamics [2, 12].

For continuous strategy games, equilibrium selection can be done in a rather natural way by stochastic differential equations (SDEs) and optimal transport theory. Individual players can be modeled to make decisions according to a stochastic process, such as the best-reply process [7], in which players change their pure strategies *locally* and simultaneously in a continuous fashion according to the direction that minimizes their own cost function most rapidly. A player's irrationality is modeled by the Brownian motion with a parameter representing the irrationality level. The time evolution of the probability density of the best-reply process is characterized by a Fokker–Planck equation, which is the learning dynamics of the game. For potential games in which all players have the same cost function named potential, this learning dynamics is the *gradient flow* of the free energy in the probability space equipped with

*Received by the editors January 4, 2018; accepted for publication (in revised form) September 23, 2019; published electronically January 14, 2020.
https://doi.org/10.1137/18M1163828
Funding: This work was partially supported by NSF Awards DMS-1419027, DMS-1620345, and ONR Award N000141310408.
†Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332 (chow@math.gatech.edu, junlu0@icloud.com, hmzhou@math.gatech.edu).
‡Department of Mathematics, University of California, Los Angeles, CA 90095 (wcli@math.ucla.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Wasserstein metric [1, 20]. Here the free energy refers to the average of the potential plus negative of the Shannon–Boltzmann entropy, representing the irrationalities or risks taken by the players. This understanding relates the game learning dynamics with statistical physics [19]. If players become purely rational, by vanishing the parameter in front of the Brownian motion, NEs are stationary points of the players' best-reply process. Thus the invariant measure associated with best-reply dynamics naturally introduces an order of NEs. This ranking method shares many similarities with the Conley–Markov matrix described in [5].

Motivated by the learning dynamics and continuous strategy best-reply processes, we propose a new learning dynamical model for finite player discrete strategy games. Our model is based on the discrete optimal transport theory developed recently [4, 6, 13, 8, 14]. Taking an N-player potential game as an example, if $\phi\colon S \to \mathbb{R}$ is the potential, and $S = S_1 \times \cdots \times S_N$ is the strategy set, where S_i is the finite discrete strategy set of player i, we use the following Fokker–Planck equations (FPE) on the strategy set S to define the dynamics in the probability simplex,

$$\begin{aligned}\frac{d\rho(t,x)}{dt} = &\sum_{y\in\mathcal{N}(x)} \rho(t,y)\Big[\phi(y)-\phi(x)+\beta(\log\rho(t,y)-\log\rho(t,x))\Big]_+ \\ &- \sum_{y\in\mathcal{N}(x)} \rho(t,x)\Big[\phi(x)-\phi(y)+\beta(\log\rho(t,x)-\log\rho(t,y))\Big]_+,\end{aligned} \tag{1.1}$$

where $\beta > 0$ is a positive constant representing the irrationality level, $\rho(t,x)$ is the probability at time t with strategy $x \in S$, $[\cdot]_+ = \max\{\cdot, 0\}$, and $y \in \mathcal{N}(x)$ if y can be achieved by players changing their strategies from x.

The density function $\rho(t,x)$ enjoys many appealing mathematical properties. For a potential game, it can be regarded as a gradient flow that converges to the minimizer of the free energy. It can be shown that the convergence is exponentially fast and the convergence rate can be accurately characterized by *relative Fisher information* [20], a key concept in statistical physics [9]. Also, the dissipation of the free energy along this learning dynamics exactly equals the relative Fisher information.

Treating FPE (1.1) as the Kolmogorov forward equation, we obtain its corresponding Markov jump process on the strategy set S, which is the best-reply dynamics. Details of its construction are given in section 3.4. The best-reply dynamics is used to model the players' decision-making process. It captures the players' behaviors, such as myopia, greed, and risk-taking. In addition, it is shown that the asymptotic distribution of the FPE (1.1) has support on pure NEs. Therefore our model can be naturally employed to select NEs, roughly speaking, by comparing the value of the limit probability distribution at each NE.

The paper is organized in the following order. In section 2, we give a brief introduction to best-reply dynamics and optimal transport theory in continuous spaces. In section 3, we describe the mathematical properties of optimal transport and best-reply dynamics derived for discrete strategy games, including the extension to nonpotential games. The connection of our model and statistical physics is discussed in section 4. In section 5, we illustrate equilibrium selections via the proposed dynamics for some well-known games.

2. Equilibrium selection in continuous strategy game. In this section, we briefly review best-reply dynamics and its connection with FPEs and optimal transport theory in continuous strategy games.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

2.1. Best-reply dynamics. Consider a game consisting of N players $i \in \{1, \dots, N\}$. Each player i chooses a strategy x_i from a Borel strategy set S_i, e.g., $S_i = \mathbb{R}^{n_i}$. Denote $S = S_1 \times \cdots \times S_N$. Let x be the vector of all players' decision variables:

$$x = (x_1, \dots, x_N) = (x_i, x_{-i}) \in S \quad \text{for any } i = 1, \dots, N,$$

where we use the notation

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

Each player i has cost function $u_i : S \to \mathbb{R}$, where $u_i(x)$ is a globally Lipchitz continuous function with respect to x. The objective of each player i is to minimize the cost function

$$\min_{x_i \in \mathbb{R}^{n_i}} u_i(x) = u_i(x_i, x_{-i}).$$

A strategy profile $x^* = (x_1^*, \dots, x_N^*)$ is an NE if no player is willing to change his or her current strategy unilaterally:

$$u_i(x_i^*, x_{-i}^*) \leq u_i(x_i, x_{-i}^*) \quad \text{for any } x_i \in S_i,\ i = 1, \dots, N.$$

It is natural to consider stochastic processes to describe players' decision-making processes in a game. For each player i, instead of finding x_i^* satisfying NE directly, he or she plays the game according to a stochastic process $X_i(t),\ t \in [0, +\infty)$. Here t is an artificial time variable, at which player i selects his or her decision based on the current strategies of all other players $X_j(t), j \in \{1, \dots, N\}$. It is important to note that all players make their decisions simultaneously and without knowing others' decisions. Each player selects a strategy that decreases his or her own cost most rapidly. To model the uncertainties of decision making, an N-dimensional independent Brownian motion is added:

$$dX_i = -\nabla_{X_i} u_i(X_i, X_{-i})dt + \sqrt{2\beta}dB_t^i, \tag{2.1}$$

where $\beta > 0$ controls the magnitude of the noise. SDE (2.1), $X(t) = (X_i(t))_{i=1}^N$, is called the best-reply process. Observe that if an NE exists, it is also an equilibrium distribution for $X(t)$ in (2.1) with $\beta = 0$. It is known that the transition density function $\rho(t, x)$ of the stochastic process $X(t)$ satisfies the FPE

$$\frac{\partial \rho(t, x)}{\partial t} = \nabla \cdot \Big(\rho(t, x)\big(\nabla_{x_i} u_i(x_i, x_{-i})\big)_{i=1}^N\Big) + \beta \Delta \rho(t, x).$$

In the case that the game is a potential game, i.e. there exists a C^1 potential function $\phi : \mathbb{R}^N \to \mathbb{R}$, such that $\nabla_{x_i} u_i(x_i, x_{-i}) = \nabla_{x_i} \phi(x)$, the best-reply process (2.1) becomes

$$dX = -\nabla \phi(X)dt + \sqrt{2\beta}dB_t,$$

which is a perturbed gradient flow, whose density function satisfies

$$\frac{\partial \rho(t, x)}{\partial t} = \nabla \cdot (\rho(t, x) \nabla \phi(x)) + \beta \Delta \rho(t, x). \tag{2.2}$$

The stationary distribution of (2.2) is the Gibbs measure given by

$$\rho^*(x) = \frac{1}{K} e^{-\frac{\phi(x)}{\beta}}, \quad \text{where } K = \int_S e^{-\frac{\phi(x)}{\beta}} dx.$$

It's easily seen that the Gibbs measure introduces an order of NEs in terms of the potential $\phi(x)$. In other words, given two NEs, the one with the larger density value will be considered more stable. One can extend this ranking to general games by studying the stationary solution of the FPE; see details in [5].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

2.2. Optimal transport. Equation (2.2) is closely related to optimal transport. It has a gradient flow interpretation in the geometry of probability space, equipped with optimal transport metric, also named L^2-Wasserstein metric.

Consider the set of smooth and strictly positive densities

$$\mathcal{P}_+(S) = \left\{ \rho \in C^\infty(S) \colon \rho(x) > 0, \ \int_S \rho(x)dx = 1 \right\}.$$

DEFINITION 2.1 (L^2-Wasserstein metric tensor). *Denote*

$$T_\rho \mathcal{P}_+(S) = \left\{ \dot\rho \in C^\infty(S) \colon \int \dot\rho(x)dx = 0 \right\}.$$

The L^2-Wasserstein metric $g_\rho^W \colon T_\rho\mathcal{P}_+(S) \times T_\rho\mathcal{P}_+(S) \to \mathbb{R}$ is defined by

$$g_\rho^W(\dot\rho_1, \dot\rho_2) = \int_S \Big(\dot\rho_1(x), (-\Delta_\rho)^{-1}\dot\rho_2(x) \Big) dx.$$

Here $\dot\rho_1, \dot\rho_2 \in T_\rho\mathcal{P}_+(S)$, $(\cdot,\cdot)$ is the metric on S, and $\Delta_\rho^{-1} \colon T_\rho\mathcal{P}_+(S) \to T_\rho\mathcal{P}_+(S)$ is the inverse of the elliptical operator

$$\Delta_\rho = div(\rho\nabla\cdot).$$

Here $(\mathcal{P}_+(S), g^W)$ is a Riemannian manifold named the *density manifold*. In particular, the Riemannian gradient of a functional $\mathcal{F}(\rho)$ is

$$\mathrm{grad}_W \mathcal{F}(\rho) = \Big((-\Delta_\rho)^{-1} \Big)^{-1} \delta\mathcal{F}(\rho) = -\Delta_\rho \delta\mathcal{F}(\rho) = -\mathrm{div}(\rho\nabla\delta\mathcal{F}(\rho)),$$

where $\mathcal{F} \colon \mathcal{P}_+(S) \to \mathbb{R}$ and δ is the L^2 first-variation operator.

We next demonstrate that the FPE (2.2) is a gradient flow of an informational functional, known as the free energy, in $(\mathcal{P}_+(S), g^W)$. Consider

$$\mathcal{F}(\rho) = \int_S \phi(x)\rho(x)dx + \beta \int_S \rho(x)\log\rho(x)dx; \tag{2.3}$$

then $\delta\mathcal{F}(\rho) = \phi(x) + \beta(\log\rho(x) + 1)$ and

$$\mathrm{grad}_W \mathcal{F}(\rho) = (-\Delta_\rho^{-1})^{-1}\delta\mathcal{F}(\rho) = -\nabla \cdot \Big(\rho\nabla(\phi + \beta\log\rho + \beta) \Big) = -\nabla\cdot(\rho\nabla\phi) - \beta\Delta\rho,$$

where we use the fact $\nabla\log\rho = \frac{1}{\rho}\nabla\rho$. Thus the gradient flow $\partial_t\rho_t = -\mathrm{grad}_W\mathcal{F}(\rho_t)$ forms exactly (2.2).

This derivation reveals that the best-reply dynamics used in modeling finite player continuous strategy games has the FPE (2.2) as its transition equation, which is also the gradient flow in the probability space with L^2-Wasserstein metric. The objective of this paper is to propose the best-reply dynamics for discrete strategy games. Our approach is to build the L^2-Wasserstein metric tensor on the probability simplex on the discrete strategy set, and study its gradient flow as the FPE. We then establish a Markov jump process associated with the FPE and use it as the best-reply dynamics. This approach also allows us to build dynamics for nonpotential games, and further gives a natural order for NEs.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3. Game dynamics and optimal transport. In this section, we introduce the time evolution of the probability density function of the best-reply process for discrete strategy games.

3.1. Wasserstein geometry in a norm form game. We first review some notations in game theory [15]. Consider a game with N players. Each player $i \in \{1, \ldots, N\}$ chooses a strategy x_i in a discrete strategy set $S_i = \{1, \ldots, M_i\}$, where M_i is an integer. Denote the joint strategy set $S = S_1 \times \cdots \times S_N$. Similarly to continuous games, each player i has a cost function $u_i : \; S \to \mathbb{R}$,

$$u_i(x) = u_i(x_i, x_{-i}).$$

If there are only two players ($N = 2$), it is customary to write the cost function in a bimatrix form (A, B^T) with $A = (u_1(i,j))_{M_1 \times M_2}$, $B^T = (u_2(i,j))_{M_1 \times M_2}$, where $(i,j) \in S_1 \times S_2$. This form of representation is called the normal form.

Example 1 (prisoner's dilemma [16]). Two members of a criminal gang are arrested and imprisoned. Each prisoner is given the opportunity either to defect the other by testifying that the other committed the crime, or to cooperate with the other by remaining silent. Their cost matrix is given by

	player 2 C	player 2 D
player 1 C	(1, 1)	(3, 0)
player 1 D	(0, 3)	(2, 2)

In this case, the strategy set is $S = \{C, D\}$, where C represents "cooperate" and D represents "defect." The cost function can be represented as (A, B^T), where

$$A = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}, \quad B^T = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix}.$$

In this example, it is easy to verify that (D, D) is the NE of the game.

For a given finite player game, we construct a corresponding strategy graph as follows. For each strategy set S_i, construct a graph $G_i = (S_i, E_i)$. Two strategies x and y are connected if player i can switch strategy from x to y. If the player is free to switch between any two strategies, it makes G_i a complete graph. Let $G = (S, E) = G_1 \square \cdots \square G_N$ be the Cartesian product of all the strategy graphs. In other words, $S = S_1 \times \cdots \times S_N$ and $x = (x_1, \ldots, x_N) \in S$ and $y = (y_1, \ldots, y_N) \in S$ are connected if their components are different at only one index and these different components are connected in their component graph, see Figure 1. For any $x = (x_1, \ldots, x_N) \in S$, denote its neighborhood

$$\mathcal{N}(x) = \{y \in S \colon \text{edge}(x, y) \in E\},$$

and directional neighborhood

$$\mathcal{N}_i(x) = \{(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_N) \colon y \in S_i, \ \text{edge}(x_i, y) \in E_i\}$$

for $i = 1, \ldots, N$. Here $\mathcal{N}_i(x)$ entails that each player selects his or her strategy with other players' strategies fixed. Notice that

$$\mathcal{N}(x) = \bigcup_{i=1}^{N} \mathcal{N}_i(x).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

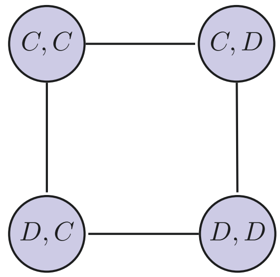


FIG. 1. *An example of the prisoner-dilemma game in Example* 1. *Here* $S_1 = S_2 = \{C, D\}$.

We now introduce an optimal transport distance on the probability space of the strategy graph. The probability space (i.e., a simplex) on all strategies is given by

$$\mathcal{P}(S) = \left\{ (\rho(x))_{x\in S} \in \mathbb{R}^{|S|} \colon \quad \sum_{x\in S} \rho(x) = 1, \quad \rho(x) \geq 0, \quad \text{for any } x \in S \right\},$$

where $\rho(x)$ is the probability at each vertex x, and $|S|$ is total number of strategies. Denote the interior of $\mathcal{P}(S)$ by $\mathcal{P}_o(S)$.

Given any function $\Phi \colon S \to \mathbb{R}$ on strategy set S, define $\nabla\Phi \colon S \times S \to \mathbb{R}$ as

$$\nabla\Phi(x, y) = \begin{cases} \Phi(x) - \Phi(y) & \text{if } (x, y) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Let $m \colon S \times S \to \mathbb{R}$ be an antisymmetric flux function such that $m(x, y) = -m(y, x)$. The divergence of m, denoted as $\text{div}(m) \in \mathbb{R}^{|S|}$, is defined by

$$\text{div}(m)(x) = -\sum_{y\in\mathcal{N}(x)} m(x, y).$$

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

$$m(x, y) = \rho\nabla\Phi := \theta(x, y, \rho)\nabla\Phi(x, y),$$

where $g(x, y, \rho)$ represents the discrete probability (weight) on edge(x, y) and satisfies

$$\theta(x, y, \rho) = \theta(y, x, \rho), \quad \min\{\rho(x), \rho(y)\} \leq \theta(x, y, \rho) \leq \max\{\rho(x), \rho(y)\}. \tag{3.1}$$

A particular choice of $\theta(x, y, \rho)$ will be given shortly in the next subsection.

We can now define

$$(\nabla\Phi, \nabla\Phi)_\rho := \frac{1}{2} \sum_{(x,y)\in E} (\Phi(x) - \Phi(y))^2 \theta(x, y, \rho),$$

which induces the following distance on $\mathcal{P}_o(S)$.

DEFINITION 3.1. *Given two discrete probability functions* ρ^0, $\rho^1 \in \mathcal{P}_o(S)$, *define the optimal transport metric function* $W \colon \mathcal{P}_o(S) \times \mathcal{P}_o(S) \to \mathbb{R}$*:*

$$W(\rho^0, \rho^1)^2 = \inf \left\{ \int_0^1 (\nabla\Phi, \nabla\Phi)_\rho dt \colon \frac{d\rho}{dt} + \text{div}(\rho\nabla\Phi) = 0, \ \rho(0) = \rho^0, \ \rho(1) = \rho^1 \right\}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

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

$$T_\rho \mathcal{P}_o(S) = \left\{ (\sigma(x))_{x\in S} \in \mathbb{R}^n \colon \sum_{x\in S} \sigma(x) = 0 \right\}.$$

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

LEMMA 3.2. *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

$$D_{(i,j)\in E, k\in V} = \begin{cases} 1 & \text{if } i = k,\ i > j, \\ -1 & \text{if } j = k,\ i > j, \\ 0 & \text{otherwise}, \end{cases}$$

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

$$\Theta(\rho)_{(x,y)\in E,(x',y')\in E} = \begin{cases} \theta(x,y,\rho) & \text{if } (x,y) = (x',y') \in E, \\ 0 & \text{otherwise}. \end{cases}$$

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

$$\Phi^{\mathsf{T}} L(\rho)\Phi = \frac{1}{2}\sum_{(x,y)\in E} (\Phi(x) - \Phi(y))^2\theta(x,y,\rho) = 0.$$

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 the above equation. Thus 0 must be the simple eigenvalue of $L(\rho)$ with eigenvector $(1,\dots,1)^{\mathsf{T}}$. Since $\operatorname{Ker}\big(L(\rho)\big) = \{(1,\dots,1)^{\mathsf{T}}\}$,

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

Thus there exists a unique solution of Φ up to a constant shift. ∎

Based on Lemma 3.2, we write

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

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

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

of $L(\rho)$ by

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

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

DEFINITION 3.3. *For any two tangent vectors $\sigma^1, \sigma^2 \in T_\rho \mathcal{P}_o(S)$, define the inner product $g_\rho^W : T_\rho \mathcal{P}_o(S) \times T_\rho \mathcal{P}_o(S) \to \mathbb{R}$ by*

$$g_\rho^W(\sigma, \tilde{\sigma}) := \sigma^\mathsf{T} L(\rho)^{-1} \tilde{\sigma} = \Phi^\mathsf{T} L(\rho) \tilde{\Phi},$$

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

Under this inner product, we can formulate the Wasserstein metric in Definition 3.1 as a geometric action function

$$W(\rho^0, \rho^1)^2 = \inf_{\rho(t) \in \mathcal{C}} \left\{ \int_0^1 \dot{\rho}^\mathsf{T} L(\rho)^{-1} \dot{\rho} dt \colon \rho(0) = \rho^0,\ \rho(1) = \rho^1 \right\}, \tag{3.2}$$

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 [13]. It enables us to define the gradient flow in $\mathcal{P}_o(S)$.

3.2. FPEs for potential games. We first derive the FPE for discrete potential games. Here a potential game means that, *there exists a potential function $\phi : S \to \mathbb{R}$, such that*

$$\phi(x) - \phi(y) = u_i(x) - u_i(y) \quad \textit{for any } x, y \in S_i \textit{ and } i = 1, \ldots, N.$$

As in the continuous case (2.3), our objective functional in $\mathcal{P}(S)$ is the discrete free energy

$$\sum_{x \in S} \phi(x)\rho(x) + \beta \sum_{x \in S} \rho(x) \log \rho(x),$$

where the first term is average of potential and the second one is the linear entropy modeling risk taking.

Using this objective functional, we construct the metric W with an upwind type $\theta(x, y, \rho)$ satisfying (3.1):

$$\theta(x, y, \rho) = \begin{cases} \rho(x) & \text{if } \phi(x) + \beta \log \rho(x) > \phi(y) + \beta \log \rho(y), \\ \rho(y) & \text{if } \phi(x) + \beta \log \rho(x) < \phi(y) + \beta \log \rho(y), \\ \frac{\rho(x) + \rho(y)}{2} & \text{if } \phi(x) + \beta \log \rho(x) = \phi(y) + \beta \log \rho(y). \end{cases}$$

We remark that the so-called upwind scheme is a common numerical method used to compute solutions of conservation laws [3]. It uses either forward or backward finite differences depending on the characteristic lines. In other words, it always uses information coming from the characteristic direction, also known as the upwind direction.

THEOREM 3.4 (gradient flow). *Given a potential game with strategy graph $G = (S, E)$, potential $\phi(x)$, and constant $\beta \geq 0$, we have the following:*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(i) *The gradient flow of*

$$\mathcal{F}(\rho) = \sum_{x\in S}\phi(x)\rho(x) + \beta\sum_{x\in S}\rho(x)\log\rho(x)$$

on the metric space $(\mathcal{P}_o(S), W)$ *is the FPE*

$$\begin{aligned}\frac{d\rho(t,x)}{dt} &= \sum_{y\in\mathcal{N}(x)}\rho(t,y)\Big[\phi(y)-\phi(x)+\beta(\log\rho(t,y)-\log\rho(t,x))\Big]_+ \\ &\quad - \sum_{y\in\mathcal{N}(x)}\rho(t,x)\Big[\phi(x)-\phi(y)+\beta(\log\rho(t,x)-\log\rho(t,y))\Big]_+.\end{aligned} \tag{3.3}$$

(ii) *For* $\beta > 0$, *Gibbs measure*

$$\rho^*(x) = \frac{1}{K}e^{-\frac{\phi(x)}{\beta}}, \quad \textit{where} \quad K = \sum_{x\in S}e^{-\frac{\phi(x)}{\beta}}, \tag{3.4}$$

is the unique stationary measure of ODE (3.3).

(iii) *For any given initial condition* $\rho^0 \in \mathcal{P}_o(S)$, *there exists a unique solution* $\rho(t) : [0,\infty) \to \mathcal{P}_o(S)$ *to* (3.3).

The proof follows [4, 6], so is omitted here.

3.3. FPE for discrete strategy games. For general games, as in the continuous case, the FPE can't be interpreted as gradient flows for some functional. To establish FPEs in a discrete setting, we observe that if the underlying graph is a Cartesian grid partition, (3.3) is the numerical discretization of the continuous FPE using an upwind scheme. This motivates us to define the following FPE.

DEFINITION 3.5. *For a general game with strategy graph* $G = (S, E)$ *with cost functionals* $u_i(x)$ *for* $i \in 1, \ldots, N$, *define its FPE to be*

$$\begin{aligned}\frac{d\rho(t,x)}{dt} &= \sum_{i=1}^{N}\sum_{y\in\mathcal{N}_i(x)}\Big[u_i(y)-u_i(x)+\beta(\log\rho(t,y)-\log\rho(t,x))\Big]_+\rho(t,y) \\ &\quad - \sum_{i=1}^{N}\sum_{y\in\mathcal{N}_i(x)}\Big[u_i(x)-u_i(y)+\beta(\log\rho(t,x)-\log\rho(t,y))\Big]_+\rho(t,x).\end{aligned} \tag{3.5}$$

Notice that $\cup_{i=1}^N N_i(x) = \mathcal{N}(x)$. So when the general game is a potential game, the above FPE coincides with (3.3). Our main result for general games is the following theorem.

THEOREM 3.6 (general flow). *Given an* N*-player game with strategy graph* $G = (S, E)$, *cost functional* u_i, $i = 1, \ldots, N$, *and a constant* $\beta \geq 0$, *we then have the following:*

(i) *For all* $\beta > 0$ *and any initial condition* $\rho(0) \in \mathcal{P}_o(S)$, *there exists a unique solution*

$$\rho(t) : [0,\infty) \to \mathcal{P}_o(S)$$

of (3.5).

(ii) *Given any initial condition* $\rho_0(t)$, *denote* $\rho^\beta(t)$ *as the solutions of* (3.5) *with varying* β*'s. Then for any fixed time* $T \in (0, +\infty)$, *the following convergence is uniform in time:*

$$\lim_{\beta\to 0}\rho^\beta(t) = \rho^0(t), \quad t \in [0, T].$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(iii) *Assume there are k distinct pure NEs $x^1, \dots, x^k \in S$. Let $\rho^*(x)$ be a measure such that*

$$\text{support of } \rho^*(x) \subset \{x^1, \dots, x^k\},$$

then $\rho^*(x)$ is a stationary solution of (3.5) *with $\beta = 0$.*

Proof. (i) This is a slight modification of the results in [3]. (ii) Let's denote ODE (3.5) for $\beta > 0$ as a matrix form

$$\frac{d\rho^\beta(t)}{dt} = Q(\rho, \beta)\rho^\beta(t).$$

We observe that if $\beta = 0$, $Q(\rho, \beta) = Q$ is a constant matrix. By a similar reason proving Theorem 3.4, we know that for any initial condition ρ^0, there exists a compact set $B(\rho^0) \subset \mathcal{P}_o(S)$, such that $\rho^\beta(t) \in B(\rho^0)$ for any β. Hence there exists a constant $M > 0$, such that

$$\|(Q(\rho, \beta) - Q)\rho^\beta(t)\| \le M\beta,$$

where $\|\cdot\|$ is the 2-norm. In other words, the difference between the ODE (3.5)'s solution at $\beta > 0$ and $\beta = 0$ is

$$\begin{aligned}\frac{d(\rho^\beta(t) - \rho^0(t))}{dt} &= Q(\rho^\beta, \beta)\rho^\beta - Q\rho^0 \\ &= Q(\rho^\beta - \rho^0) + (Q(\rho^\beta, \beta) - Q)\rho^\beta.\end{aligned}$$

Hence

$$\begin{aligned}\frac{d\|\rho^\beta(t) - \rho^0(t)\|}{dt} &\le \|Q(\rho^\beta(t) - \rho^0(t))\| + \|(Q(\rho^\beta, \beta) - Q)\rho^\beta\| \\ &\le \|Q\|\|\rho^\beta - \rho^0\| + \beta M.\end{aligned}$$

By Gronwall's inequality, for $t \in [0, T]$, we have

$$\|\rho^\beta(t) - \rho^0(t)\| \le \beta M e^{\|Q\|T},$$

which finishes the proof.

We now prove (iii). Denote $\mathcal{E} = \{x^1, \dots, x^k\}$, then support of $\rho^*(x) \subset \mathcal{E}$ implies

$$\rho^*(x) = \begin{cases} 0 & \text{if } x \notin \mathcal{E}, \\ \ge 0 & \text{if } x \in \mathcal{E}. \end{cases} \tag{3.6}$$

Since $x \in \mathcal{E}$ is an NE, $u_i(y) \ge u_i(x)$ when $y \in \mathcal{N}_i(x)$ for any $i = 1, \dots, d$. For $x \in \mathcal{E}$, we substitute $\rho^*(x)$ into the right-hand side of (3.5), which forms

$$\begin{aligned}&\sum_{i=1}^N \sum_{y \in \mathcal{N}_i(x)} [u_i(y) - u_i(x)]_+ \rho^*(y) - \sum_{i=1}^N \sum_{y \in \mathcal{N}_i(x)} [u_i(x) - u_i(y)]_+ \rho^*(x) \\ &\quad = \sum_{i=1}^N \sum_{y \in \mathcal{N}_i(x)} [u_i(y) - u_i(x)]\rho^*(y) - 0 \\ &\quad = 0,\end{aligned}$$

where the last equality is from the following facts in two cases. (i) If $y \notin \mathcal{E}$, $\rho^*(y) = 0$ from (3.6). (ii) if $y \in \mathcal{E}$, $u_i(y) \ge u_i(x)$, then $u_i(y) - u_i(x) = 0$. Similarly, we can show the case when $x \notin \mathcal{E}$. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3.4. Best reply Markov process. For the continuous strategy games, the FPE can be regarded as the evolution of the density function of best-reply dynamics. In this subsection, we introduce a similar notion to the discrete strategy games. More precisely, we define Markov jumping processes among discrete strategy sets.

We start with an N-player potential game with strategy graph $G = (S, E)$ and potential ϕ. Consider the following time homogenous Markov process $X(t)$ on the set S whose transition probability is

$$\begin{aligned}&\Pr(X(t+h) = y\colon X(t) = x)\\&\quad= \begin{cases}\Big[\phi(x) - \phi(y)\Big]_+ h + o(h) & \text{if } y \in \mathcal{N}(x),\\ 1 - \sum_{y\in\mathcal{N}(x)} \Big[\phi(x) - \phi(y)\Big]_+ h + o(h) & \text{if } y = x,\\ 0 & \text{otherwise,}\end{cases}\end{aligned}$$

where $\lim_{h\to 0} \frac{o(h)}{h} = 0$. Denote $\rho(t, x) = \Pr(X(t) = x)$, the transition probability function. Then the time evolution of $\rho(t, x)$ is given by forward Kolmogorov equation

$$\frac{d\rho(t,x)}{dt} = \sum_{y\in\mathcal{N}(x)} \Big[\phi(y) - \phi(x)\Big]_+ \rho(y) - \sum_{y\in\mathcal{N}(x)} \Big[\phi(x) - \phi(y)\Big]_+ \rho(x). \tag{3.7}$$

Equation (3.7) can be seen as the discrete version of the FPE (2.2) with $\epsilon = 0$ and the Markov process $X(t)$ is the discrete version of the pure gradient flows (2.1) with $\epsilon = 0$. To introduce white noise into the Markov process, by comparing (3.7) and (3.3), one can see that if we replace the potential ϕ with the noisy cost functional

$$\bar{\phi}(x) = \phi(x) + \beta \log \rho(x), \quad x \in S,$$

we will arrive exactly at FPE (3.3). In other words, we define our gradient Markov process $X_\beta(t) \in S$ to be

$$\begin{aligned}&\Pr(X_\beta(t+h) = y\colon X_\beta(t) = x)\\&\quad= \begin{cases}\Big[\bar{\phi}(x) - \bar{\phi}(y)\Big]_+ h + o(h) & \text{if } y \in \mathcal{N}(x),\\ 1 - \sum_{y\in\mathcal{N}(x)} \Big[\bar{\phi}(x) - \bar{\phi}(y)\Big]_+ h + o(h) & \text{if } y = x,\\ 0 & \text{otherwise.}\end{cases}\end{aligned}$$

Remark 3.7. Equation (2.2) in a continuous strategy game can also be rewritten as

$$\frac{\partial\rho(t,x)}{\partial t} = \nabla\cdot\big(\rho(t,x)\big(\nabla\phi(x) + \beta\nabla\log\rho(t,x)\big)\big).$$

Its Lagrange formulation is the the nonlinear ODE in the sense of Mckean type:

$$dX = -\nabla\big(\phi(X) + \beta\log\rho(t,X)\big)dt. \tag{3.8}$$

Notice that process X and its density function are coupled. The term $\log\rho$ corresponds to the Brownian motion in the best-reply SDE. The formulation of (3.8) gives the justification of our definition of *noise cost (payoff)* and motivates the definition of the jump process in discrete strategy games.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Mimicking the connections among (2.1), (3.8), and (2.2), we can extend the formulations to the nonpotential games. Namely, with the noise cost functional

$$\bar{u}_i(x) = u_i(x) + \beta \log \rho(x),$$

the best-reply Markov process $X_\beta(t)$ for a nonpotential game is

$$\begin{aligned} &\Pr(X_\beta(t+h) = y \colon X_\beta(t) = x) \\ &= \begin{cases} \sum_{i=1}^N \left[\bar{u}_i(y) - \bar{u}_i(x)\right]_+ h + o(h) & \text{if } y \in \mathcal{N}(x), \\ 1 - \sum_{i=1}^N \sum_{y \in \mathcal{N}_i(x)} \left[\bar{u}_i(y) - \bar{u}_i(x)\right]_+ h + o(h) & \text{if } y = x, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{3.9}$$

The time evolution $\rho(t) = (\rho(t,x))_{x \in S}$ of Markov process $X_\beta(t)$ is exactly FPE (3.5).

The process $X_\beta(t)$ describes players' behaviors with the following features. The Markovian property of $X_\beta(t)$ reflects players' myopia when making decisions. In other words, players make their decisions based solely on the most recent information. The noisy cost functional reflects players' irrational behaviors (This may be because the player is a *risk-taker*). The decision making is local in our model, meaning players only need local information, including the cost and relative popularity $\log \frac{\rho(t,x)}{\rho(t,y)}$ for its neighboring strategy, to make the next selection. It is easily seen that players select the next strategy to decrease their collective cost functionals with largest probability. This is to say players are greedy during the decision-making process. It's also worth mentioning that the decision process depends on the distribution ρ, which can be interpreted as the collective behavior of infinitely many copies of players playing simultaneously or the game being played by the player repeatedly for infinitely many times. In other words, the proposed model assumes that each player has additional information that stems from repeatedly playing the exact same game.

3.5. NEs selection. If FPEs possess stationary distributions (equilibria) for the dynamics, we can rank different equilibria by comparing their probabilities.

For potential games, the stationary distribution is the Gibbs measure, which provides the same ranking as that given by simply comparing potentials. Denote $x^1, \dots, x^k \in S$ as distinct NEs. A natural order is as follows:

$$x^1 \preceq x^2 \cdots \preceq x^k \quad \text{if} \quad \rho^*(x^1) \le \cdots \le \rho^*(x^k). \tag{3.10}$$

Here, $x \preceq y$ is to say that the strategy y is better(more stable) than strategy x. The above definition is equivalent to looking at $\phi(x^1) \ge \cdots \ge \phi(x^k)$, since $\rho^*(x) = \frac{1}{K} e^{-\frac{\phi(x)}{\beta}}$. In other words, the smaller potential corresponds to the larger probability, and is thus more stable.

For nonpotential games, although there are no potentials, the stationary solution of FPE, $\rho^*(t)$, if it exists, still provides a way of ranking equilibria. We call it the *transport order of NEs*.

DEFINITION 3.8 (transport order of NEs). *Assume* $\rho^*(x) = \lim_{\beta \to 0} \lim_{t \to \infty} \rho(t,x)$ *exists, where* $\rho(t,x)$ *is a solution of* (3.5) *with any initial measure* $\rho^0 \in \mathcal{P}_o(S)$. *We define the order of NE by*

$$x^1 \preceq x^2 \cdots \preceq x^k \quad \textit{if} \quad \rho^*(x^1) \le \cdots \le \rho^*(x^k). \tag{3.11}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

For potential games, $\rho^*(x)$ is always well defined, since it gives the support of a global minimizer, while in nonpotential games, the existence of $\rho^*(x)$ is still an open problem. In our numerical experiments, we choose a suitably small β, and compute the solution of the FPE at a large enough time T. Then we treat $\rho(T,x)$ as the stationary solution, i.e. $\rho(T,x) \approx \rho^*(x)$, and use it to rank the order of NEs. In all numerical examples we tested, the numerical results indicate that the $\rho(t,x)$ becomes invariant when T is large enough.

Compared to the risk and payoff dominance, the proposed approach takes the graph structure into consideration, which impacts the order of NEs. In section 5, we give several examples to illustrate this selection method.

4. Entropy dissipation. In this section, we illustrate the connection between our Markov process and statistical physics, named the discrete H theory. We will mainly focus on potential games. We borrow two "discrete" physical functionals to measure the closeness between two discrete measures, ρ and $\rho^\infty(x) = \frac{1}{K}e^{-\frac{\phi(x)}{\beta}}$. One is the discrete relative entropy (H),

$$\mathcal{H}(\rho|\rho^\infty) := \sum_{x\in S} \rho(x) \log \frac{\rho(x)}{\rho^\infty(x)}.$$

The other is the discrete relative Fisher information (I),

$$\mathcal{I}(\rho|\rho^\infty) := \sum_{(x,y)\in E} \left(\log \frac{\rho(x)}{\rho^\infty(x)} - \log \frac{\rho(y)}{\rho^\infty(y)} \right)_+^2 \rho(x).$$

The H theory states that the relative entropy decreases along a player's decision process. The following theorem can be viewed as the discrete H theorem for finite player games.

THEOREM 4.1 (discrete H theorem). *Suppose $\rho(t)$ is the transition probability of $X_\beta(t)$ in potential games. Then the relative entropy decreases,*

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

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

$$\frac{d}{dt}\mathcal{H}(\rho(t)|\rho^\infty) = -\beta \mathcal{I}(\rho(t)|\rho^\infty). \tag{4.1}$$

Proof. Since $\mathcal{I}(\rho|\rho^\infty) \geq 0$ and equality is achieved if and only if $\rho = \rho^\infty$, we only need to prove (4.1). Substituting $\rho^\infty(x) = \frac{1}{K}e^{-\frac{\phi(x)}{\beta}}$ into the relative entropy, we observe

$$\begin{aligned}
\mathcal{H}(\rho|\rho^\infty) &= \sum_{x\in S} \rho(x) \log \frac{\rho(x)}{\rho^\infty(x)} \\
&= \sum_{x\in S} \rho(x) \log \rho(x) - \sum_{x\in S} \rho(x) \log \rho^\infty(x) \\
&= \sum_{x\in S} \rho(x) \log \rho(x) + \frac{1}{\beta} \sum_{x\in S} \rho(x)\phi(x) + \log K \sum_{x\in S} \rho(x) \\
&= \frac{1}{\beta} \left(\beta \sum_{x\in S} \rho(x) \log \rho(x) + \sum_{x\in S} \rho(x)\phi(x) \right) + \log K.
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

From the explicit formulation of FPE (3.3), we have

$$\begin{aligned}\frac{d}{dt}\mathcal{H}(\rho(t)|\rho^{\infty}) &= \frac{1}{\beta}\frac{d}{dt}\left\{\beta\sum_{x\in S}\rho(t,x)\log\rho(t,x)+\sum_{x\in S}\rho(t,x)\phi(t,x)\right\}\\ &= -\frac{1}{\beta}\sum_{(x,y)\in E}\Big(\phi(x)+\beta\log\rho(t,x)-\phi(y)-\beta\log\rho(t,y)\Big)_+^2\rho(t,x)\\ &= -\frac{1}{\beta}\cdot\beta^2\cdot\sum_{(x,y)\in E}\left(\log\frac{\rho(t,x)}{\rho^{\infty}(x)}-\log\frac{\rho(t,y)}{\rho^{\infty}(y)}\right)_+^2\rho(t,x)\\ &= -\beta\cdot\mathcal{I}(\rho(t)|\rho^{\infty})\leq 0,\end{aligned}$$

which finishes the proof. ∎

Besides the discrete H theorem, there is a deep connection between FPE (3.3) and statistical physics from the mathematical viewpoint. This connection is known as entropy dissipation, i.e., the relative entropy decreases to zero exponentially. We show similar results for the proposed model.

THEOREM 4.2 (entropy dissipation). *Given a potential game with $\beta > 0$, $\rho^0 \in \mathcal{P}_o(S)$, there exists a constant $C = C(\beta, \rho^0, G) > 0$ such that*

$$\mathcal{H}(\rho(t)|\rho^{\infty}) \leq e^{-Ct}\mathcal{H}(\rho^0|\rho^{\infty}). \tag{4.2}$$

The proof of Theorem 4.2 is presented in [6].

5. Examples. We give several examples to illustrate the dynamical model and present the related order of NEs. For comparison purposes, we first recall the ranking results obtained by risk dominance and payoff (cost) dominance, two existing ranking strategies, on the prisoner's dilemma presented in section 3.

Payoff dominance: Strategy pair (C, C) payoff dominates (D, D) if $u_1(C, C) \leq u_1(D, D)$, $u_2(C, C) \leq u_2(D, D)$, and at least one of the two is a strict inequality.

Risk dominance: Strategy pair (C, C) risk dominates (D, D) if the product of the deviation losses is highest for (D, D), i.e., $u_2(C, D) - u_2(C, C) > u_1(D, C) - u_1(D, D)$.

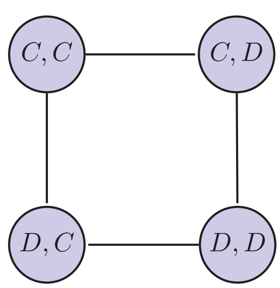
Example 2 (Example 1 continued). Consider a two-player prisoner dilemma (A, B^T) game with cost matrix

$$A = B = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}.$$

Here the strategy set is $S = \{(C, C), (C, D), (D, C), (D, D)\}$. This particular game is a potential game with

$$\phi(x) = -(u_1(x) + u_2(x)), \quad \text{where } x \in S.$$

The strategy graph is $G = K_2 \square K_2$.



Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

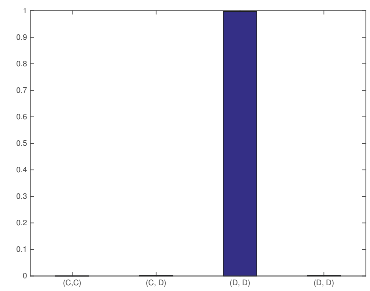


FIG. 2. *The invariant measure ρ^* for prisoner dilemma.*

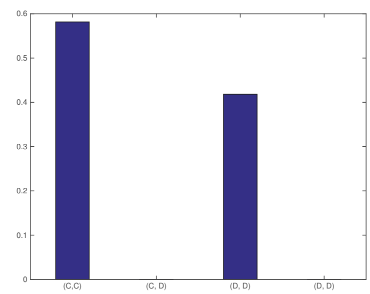


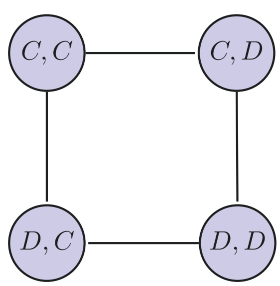
FIG. 3. *The invariant measure ρ^* for an asymmetric game.*

To simplify notations, we denote the transition probability function as

$$\rho(t) = (\rho_{CC}(t), \rho_{CD}(t), \rho_{DC}(t), \rho_{DD}(t))^T,$$

which satisfies FPE (3.5). By numerically solving (3.5) for $\rho^* = \lim_{\beta\to 0} \lim_{t\to\infty} \rho(t)$, we find a unique invariant measure ρ^* for any initial condition $\rho(0)$, which is demonstrated in Figure 2. Indeed, we know that ρ^* is a Gibbs measure and (D, D) is the unique NE. In this case, our method coincides with the payoff (cost) dominance.

Example 3. Consider an asymmetric game (A, B^T), i.e., $A \neq B$. This means players' costs depend on their own identity. Let $A = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 1 & 3 \\ 2 & 1 \end{smallmatrix}\right)$. This game is not a potential game. Again the strategy graph is $G = K_2 \square K_2$.



By solving (3.5) for $\rho^* = \lim_{\beta\to 0} \lim_{t\to\infty} \rho(t)$, we obtain a unique ρ^* for any initial condition $\rho(0)$, which is shown in Figure 3. As we can see, ρ^* only supports at (C, C) and (D, D), both of which are NEs of the game. Moreover, ρ^*_{CC} is larger than ρ^*_{DD}, which implies that (C, C) is more "stable" than (D, D). This is intuitive because player 2 is more willing to change his/her status from (C, D) to (C, C) than player 1 to move the status (D, C) to (D, D), since player 2's cost changes more rapidly than

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

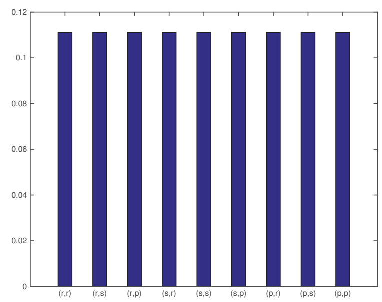


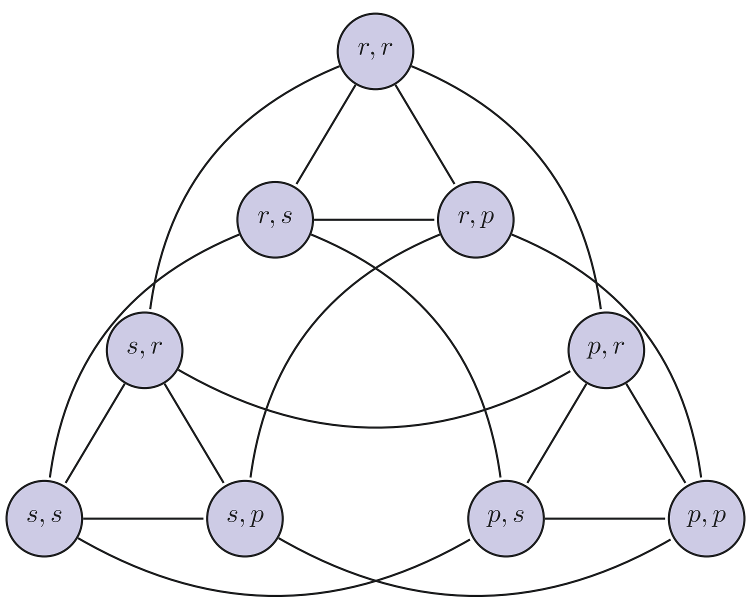
FIG. 4. *The invariant measure ρ^* for rock-scissors-paper.*

the one of player 1: $u_2(C,D)-u_2(C,C)=2>1=u_1(D,C)-u_1(D,D)$. In this case, our method coincides with the risk dominance.

Example 4. Consider a rock-scissors-paper game (A,B^T) with the strategy sets $S_1=S_2=\{r,s,p\}$ and the cost matrix

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

The strategy graph is $G=K_3\square K_3$:



Again, we obtain a unique invariant ρ^* for any initial condition $\rho(0)$ in Figure 4. From the figure, we find that the invariant measure ρ^* is a uniform measure. We conclude that, although each player chooses his/her own strategy depending on others, at the final time, they will arrive at a state that players select strategies uniformly and independently.

Example 5. We consider the same rock-scissors-paper game with constraints, in order to illustrate the effect of the structure of the strategy graph on stationary joint

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

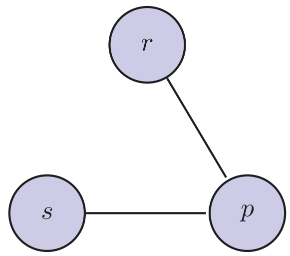


FIG. 5. *Player 1's strategy graph.*

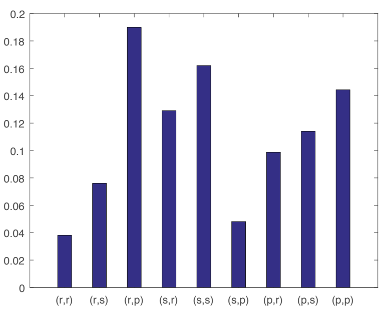


FIG. 6. *The invariant measure ρ^* for rock-scissors-paper with constraints.*

probability ρ^*. Here the constraint is that player 1 is not allowed to play scissors following rock and vice versa. There is no restriction on player 2. The corresponding strategy graph S_1 is in Figure 5 while the strategy graph S_2 is a complete graph. We consider $S_1 \square S_2$ for FPE (3.5) and solve for the invariant measure ρ^*. From Figure 6, we observe several properties that accord with modeling intuitions. First, player 1 is at disadvantage to player 2, since the chance of player 1 winning is less than that of player 2,

$$\rho^*_{(r,s)} + \rho^*_{(p,r)} + \rho^*_{(s,p)} = 0.2228 < 0.4329 = \rho^*_{(s,r)} + \rho^*_{(r,p)} + \rho^*_{(p,s)}.$$

Second, we see that players 1 and 2's probabilities are not independent, meaning that they make decisions depending on each others' choices. Third, from player 1's perspective, by assuming player 2 selected strategies uniformly, player 1 would choose paper more frequently than rock and scissors due to the constraint. Thus in turn by taking advantage of this information, player 2 would have selected paper (0 cost) or scissors (-1 cost). This is reflected by Figure 6 that the top three states with highest probabilities are $(r,p), (s,s)$ and (p,p). In this case, the risk or payoff (cost) dominance does not take the strategy graph information into consideration. While our approach is still able to consider all information and provide a detailed ranking for each situations.

6. Conclusion. We summarize all features of the proposed dynamic framework: First, the model incorporates players' myopia, uncertainty, and greed when making decisions. Second, the model works for both potential and nonpotential games. For potential games, the ranking of NE given by the limit distribution coincides with the ranking given by the potential; For nonpotential games, this ranking relates to the Morse decomposition and Conley–Markov matrix proposed in [5]. Last, but not least,

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

the proposed FPE converges to Gibbs measure for potential games. The convergence is exponentially fast, where the rate is controlled by the relation between discrete entropy and Fisher information [6, 9].

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Math. ETH Zürich, Birkhäuser, Basel, 2008.

[2] L. E. Blume, *The statistical mechanics of strategic interaction*, Games Econom. Behav., 5 (1993), pp. 387–424.

[3] S.-N. Chow, L. Dieci, W. Li, and H. Zhou, *Entropy dissipation semi-discretization schemes for Fokker–Planck equations*, J. Dynam. Differential Equations, 31 (2019), pp. 765–792.

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

[5] S.-N. Chow, W. Li, Z. Liu, and H.-M. Zhou, *A natural order in dynamical systems based on Conley–Markov matrices*, J. Differential Equations, 252 (2012), pp. 3116–3141.

[6] S.-N. Chow, W. Li, and H. Zhou, *Entropy dissipation of Fokker-Planck equations on graphs*, Discrete Contin. Dyn. Syst. A, 38 (2018), pp. 4929–4950.

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

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

[9] B. R. Frieden and B. R. Frieden, *Science from Fisher Information: A Unification*, 2nd ed., Cambridge University Press, Cambridge, 2004.

[10] J. C. Harsanyi, *A new theory of equilibrium selection for games with complete information*, Games Econom. Behav., 8 (1995), pp. 91–122.

[11] J. C. Harsanyi and R. Selten, *A General Theory of Equilibrium Selection in Games*, MIT Press Class., MIT Press, Cambridge, MA, 1992.

[12] M. Kandori, G. J. Mailath, and R. Rob, *Learning, mutation, and long run equilibria in games*, Econometrica, 61 (1993), pp. 29–56.

[13] W. Li, *Geometry of Probability Simplex via Optimal Transport*, arXiv:1803.06360 [math], 2018, https://arxiv.org/abs/1803.06360.

[14] J. Maas, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal., 261 (2011), pp. 2250–2292.

[15] J. F. Nash, *Equilibrium points in n-person games*, Proc. Natl. Acad. Sci. USA, 36 (1950), pp. 48–49.

[16] W. Poundstone, *Prisoner's Dilemma*, Anchor Books, New York, 1993.

[17] W. H. Sandholm, *Evolutionary game theory*, in Encyclopedia of Complexity and Systems Science, R. A. Meyers, ed., Springer, Berlin, 2017, pp. 1–38.

[18] K. Sigmund and M. A. Nowak, *Evolutionary game theory*, Current Biol., 9 (1999), pp. R503–R505.

[19] C. Villani, *A review of mathematical topics in collisional kinetic theory*, in Handbook of Mathematical Fluid Dynamics, Vol. 1, Elsevier, Amsterdam, 2002, pp. 71–74.

[20] C. Villani, *Optimal Transport: Old and New*, Grundlehren Math. Wiss. 338, Springer, Berlin, 2009.

[21] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton Class. Ed., Princeton University Press, Princeton, NJ, 2007.

Downloaded 08/19/20 to 143.215.38.35. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.