GAME THEORY AND DISCRETE OPTIMAL TRANSPORT

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

Abstract. We propose a new dynamic framework for finite player discrete strategy games. By utilizing tools from optimal transportation theory, we derive Fokker-Planck equations (FPEs) on finite graphs. Furthermore, we introduce an associated Best-Reply Markov process that models players' myopicity, greedy and uncertainty when making decisions. The model gives rise to a method to rank/select equilibria for both potential and non-potential games.

1. Introduction

Game theory plays a vital role in economics, biology, social network, etc. [10, 21, 24, 25, 22]. 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) is related to describe a status that no player is willing to change his/her strategy unilaterally. For any given game, a fundamental question is that if there are multiple pure Nash equilibria, how can one select/rank them? This problem has been studied previously using various approaches. For example, in [11, 12], NEs were selected by refinements of equilibrium concepts, namely payoff dominance or risk dominance principle. These deductive principles assume players have consistent beliefs with the equilibria. Another class of approaches use evolutionary dynamics to study the equilibrium by assuming that the game is played by a large population. The players in the population are matched to play the finite game and update their strategy according to different rules. Such dynamics include Replicator dynamics, Logit dynamics, and Best-response dynamics [2, 14, 18].

The aim of this paper is to propose a dynamic model utilizing optimal transport theory [1, 28], which also gives rise to a natural order on pure Nash equilibria [16]. The proposed dynamics assumes that at each time of playing, players "simulate" playing the game infinite many times in their mind and the collective behavior, i.e. the distribution of such infinite plays is known to the players. This is similar to fictitious play [3, 20], with distinctive features that (i) the strategy at each time of playing in our dynamics is not stationary (ii) the distributions the players best responds to are different between the dynamics and fictitious play. Following the core theory of optimal transport, the proposed system has pure Nash equilibria as stationary points and incorporates randomness, which models players' uncertainty during the decision-making process. In short, players are modeled to make decisions according to a stochastic process. This idea is similar to the Best-Reply

Key words and phrases. Game theory; Optimal transport; Gradient flow; Gibbs measure; Entropy; Fisher information.

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

dynamics introduced in [7]. There the strategy sets are continuous and players change their pure strategies *locally* and simultaneously in a continuous fashion according to the direction that minimizes their own cost function most rapidly. Randomness is also introduced in the form of white noise perturbation. The resulting Best-Reply dynamics becomes a stochastic differential equation (SDE), whose probability density function evolves according to the Fokker-Planck equation (FPE). However, this theory can't be parallelly applied to our problem, in which the strategy sets are discrete. This is mainly due to the fact that the discrete strategy sets are no longer a length space (a space that one can define length of curves). This difference comes into play when one wants to define white noise on a discrete set, more specifically, a Markov process.

We overcome the obstacles using the theory recently developed in [4, 5], known as discrete optimal transport. Similar ideas have been studied in [8, 17]. More specifically, FPE can be regarded as either the time evolution of probability density function of a stochastic process, or the gradient flow of the free energy in probability space. In the context of games, the free energy is the average of cost plus an entropy term, which represents the amount of risk taken by the players. We sketch the process for N-player potential games, i.e. all players share the same cost function $\phi\colon S \to \mathbb{R}$, named potential. Let $S = S_1 \times \cdots \times S_N$ be the strategy set where S_i is the finite discrete strategy set of player i. In discrete settings, although one can't define a gradient stochastic process directly, the probability space on S and the discrete version of the free energy are well defined. This allows us to derive the gradient flow by the following ODE

$$\begin{aligned}\frac{d\rho(t,x)}{dt} = &\sum_{y\in\mathcal{N}(x)} \rho(t,y)[\phi(y)-\phi(x)+\beta(\log\rho(t,y)-\log\rho(t,x))]_+ \\ &-\sum_{y\in\mathcal{N}(x)} \rho(t,x)[\phi(x)-\phi(y)+\beta(\log\rho(t,x)-\log\rho(t,y))]_+ \ ,\end{aligned} \tag{1}$$

where $\rho(t,x)$ 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. We call (1) the FPE of a game and the nonlinear log term in (1) the log-Laplacian, which is different from the graph Laplacian. The derivation of log-Laplacian is through discrete entropy [4, 26], which is known to measure uncertainties of a system in information theory. As an analogy of continuous space that entropy introduces a white noise, the log-Laplacian is to represent discrete "white noise" perturbation.

FPE (1) can be extended to non-potential games, in which the resulting ODE is a non-gradient system:

$$\begin{aligned}\frac{d\rho(t,x)}{dt} = &\sum_{i=1}^{N}\sum_{y\in\mathcal{N}_i(x)} \rho(t,y)[u_i(y)-u_i(x)+\beta(\log\rho(t,y)-\log\rho(t,x))]_+ \\ &-\sum_{i=1}^{N}\sum_{y\in\mathcal{N}_i(x)} \rho(t,x)[u_i(x)-u_i(y)+\beta(\log\rho(t,x)-\log\rho(t,y))]_+ \ .\end{aligned} \tag{2}$$

Here $u_i\colon S \to \mathbb{R}$ is the cost function for player i, and $\mathcal{N}_i(x)$ is the strategy neighborhood of player i. See details in 3.3.

From FPE (1) or (2), we obtain its corresponding Markov process through Kolomogrov forward equation. The process is the dynamics governing the players' decisions. Intuitively, the Markov process models the behaviors of game players that are more realistic. Namely, players are myopic, greed and sometimes irrational or more risk-taking. In addition, it is shown that the limit distribution of FPE (1) or (2) has support on pure NEs. Therefore the aforementioned model can be naturally employed for selecting NEs, roughly speaking, by comparing the value of the limit probability distribution at each NE. This ranking strategy shares many similarities to Morse decomposition and Conley-Markov Matrix in [16].

Our 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 derive FPEs by optimal transport theory for discrete strategy games. In section 4, we derive the Best-Reply Markov process. The connection of our model and statistical physics is also discussed. In section 5, we illustrate several examples of new dynamics (2) for some well-known games.

2. Reviews on continuous strategy game

In this section, we briefly review Best-Reply dynamics and its connection with optimal transportation theory, see details in [7, 28].

2.1. Best Reply dynamics. Consider a game consisting N players $i \in \{1, \cdots, N\}$. 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:

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

where we use the notation

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

Each player i has his/her own 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 his/her cost function

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

Definition 1. *A strategy profile $x^* = (x_1^*, \cdots, x_N^*)$ is a Nash equilibrium (NE) if no player is willing to change his/her current strategy unilaterally*

$$u_i(x_i^*, x_{-i}^*) \leq u_i(x_i, x_{-i}^*) \quad \textit{for any } x_i \in S_i \ , \ i = 1, \cdots, N \ . \tag{3}$$

It is natural to consider stochastic processes to describe players' decisions-making processes in a game. For each player i, instead of finding x_i^* satisfying (3) 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/her decision based on the current strategies of all other players $x_j(t), t \in \{1, \cdots, N\}$. It is important to note that all players make their decisions simultaneously and without knowing others' decisions. Each player selects his/her strategy that decreases the player's own cost most rapidly. In other words,

$$dx_i = -\nabla_{x_i} u_i(x_i, x_{-i}) dt \ .$$

To model the uncertainties of decision making, an N-dimensional independent white noise is added

$$dx_i = -\nabla_{x_i} u_i(x_i, x_{-i})dt + \epsilon dW_t^i \ ,$$

where ϵ controls the magnitude of the noise. Putting all players' process together $x(t) = (x_i(t))_{i=1}^N$ and denoting $f(x) = \big(\nabla_{x_i} u_i(x_i, x_{-i})\big)_{i=1}^N$, one gets

$$dx = -f(x)dt + \epsilon dW_t \ . \tag{4}$$

SDE (4) is called the Best-reply dynamics and $x(t)$ the Best-Reply decision process. Observe that if a Nash equilibrium exists, it is also the equilibrium of (4) with $\epsilon = 0$. It is known that the transition density function $\rho(t, x)$ of the stochastic process $x(t)$ satisfies the FPE

$$\frac{\partial \rho(t,x)}{\partial t} = \nabla \cdot (\rho(t,x) f(x)) + \beta \Delta \rho(t,x) \ , \quad \text{where } \beta = \frac{\epsilon^2}{2} \ . \tag{5}$$

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 $f(x) = \nabla \phi(x)$. The Best-Reply SDE (4) becomes

$$dx = -\nabla \phi(x)dt + \epsilon dW_t \ , \tag{6}$$

which is a perturbed gradient flow, whose transition equation, FPE, forms

$$\frac{\partial \rho(t,x)}{\partial t} = \nabla \cdot (\rho(t,x) \nabla \phi(x)) + \beta \Delta \rho(t,x) \ . \tag{7}$$

2.2. **Optimal transport.** Equation (7) connects with the Optimal transport theory, and has a gradient flow interpretation in geometry. We explain it by quoting a sentence in Villani's book [28]:

> The density of gradient flow (FPE) is a gradient flow in density spaces.

To view this connection, we briefly review optimal transport theory. The theory introduce a distance, known as the Wasserstein metric, on the probability density space, through which the probability set $\mathcal{P}(\mathbb{R}^d)$ forms an infinite dimensional Riemannian manifold. On this manifold, FPE (7) is a gradient flow of an informational functional, known as free energy in statical physics:

$$\int_{\mathbb{R}^d} \phi(x)\rho(x)dx + \beta \int_{\mathbb{R}^d} \rho(x) \log \rho(x) dx \ . \tag{8}$$

Many properties of statistical physics, such as entropy dissipation [5] and Fisher information [9], can be understood from this gradient flow interpretations.

Besides above connections, the stationary distribution of (7), or the invariant measure of (4) is the Gibbs measure given by

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

It's easily seen that the Gibbs measure introduces an order of Nash equilibria in terms of the potential $\phi(x)$. In other words, given two Nash equilibria, the one with larger density value will be considered more stable. One can extend this ranking to general Best-Reply dynamics, see recently studies in Morse decomposition and Conley-Markov matrix [16].

Here our goal is to use the above ideas to build dynamics and to rank Nash equilibria in a discrete setting. However, the above derivation highly depends on the structure of the strategy space S. The most general space on which one can extend the derivations is the so called length space, which unfortunately excludes the discrete space we are interested in.

3. Fokker-Planck equations of Discrete strategy Games

In this section, we introduce an optimal transport distance for discrete strategy games. Based on such a distance, we derive FPE for modeling players' behaviors.

3.1. **Optimal transport distance for normal-form games.** We first review some facts and notations in game theory [21]. Consider a game with N players. Each player $i \in \{1, \cdots, N\}$ chooses a strategy x_i in a discrete strategy set

$$S_i = \{1, \cdots, M_i\}$$

where M_i is an integer. Denote the joint strategy set

$$S = S_1 \times \cdots \times S_N.$$

Similar 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 bi-matrix 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 normal form.

Example 1. *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 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 $x = (x_1, \cdots, x_N) \in S$ and $y = (y_1, \cdots, y_N) \in S$ are connected if their components are different at only one index and these different components are connected in their component graph. For any $x = (x_1, \cdots, x_N) \in S$, denote its neighborhood to be $\mathcal{N}(x)$

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

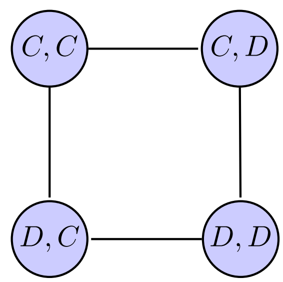
and directional neighborhood to be

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

for $i = 1, \cdots, N$. Notice that

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

Example 2. *Consider a two player Prisoner-Dilemma game, where $S_1 = S_2 = \{C, D\}$. The strategy graph is the following.*



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:

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

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 a anti-symmetric flux function such that $m(x, y) = -m(y, x)$. The divergence of m, denoted as $\text{div}(m) \colon S \to \mathbb{R}^{|S|}$, 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

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

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

$$g(x, y, \rho) = g(y, x, \rho) \ , \quad \min\{\rho(x), \rho(y)\} \leq g(x, y, \rho) \leq \max\{\rho(x), \rho(y)\} \ . \tag{9}$$

A particular choice of $g(x, y, \rho)$ is of up-wind scheme type, whose explicit formulation will be given shortly.

We can now define the discrete inner product on $\mathcal{P}_o(S)$:

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

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

Definition 2. *Given two discrete probability function ρ^0, $\rho^1 \in \mathcal{P}_o(S)$, consider the metric function $\mathcal{W}$:*

$$\mathcal{W}(\rho^1, \rho^2)^2 = \inf\{\int_0^1 (\nabla\Phi, \nabla\Phi)_\rho dt \ : \ \frac{d\rho}{dt} + \mathrm{div}(\rho\nabla\Phi) = 0 \ , \ \rho(0) = \rho^0, \ \rho(1) = \rho^1\} \ .$$

$(\mathcal{P}_o(S), \mathcal{W})$ is a well defined finite dimensional Riemannian manifold [4, 17], which enables us to define the gradient flow (FPE) 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*

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

As in the continuous case [28], our objective functional in $\mathcal{P}(S)$ is

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

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 $\mathcal{W}$ with a upwind type $g(x, y, \rho)$ satisfying (9):

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

Theorem 3 (Gradient flow)**.** *Given a potential game with strategy graph $G = (S, E)$, potential $\phi(x)$ and constant $\beta \geq 0$.*

(i) *The gradient flow of*

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

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

$$\begin{aligned} \frac{d\rho(t,x)}{dt} = & \sum_{y\in\mathcal{N}(x)} \rho(t, y)[\phi(y) - \phi(x) + \beta(\log \rho(t, y) - \log \rho(t, x))]_+ \\ & - \sum_{y\in\mathcal{N}(x)} \rho(t, x)[\phi(x) - \phi(y) + \beta(\log \rho(t, x) - \log \rho(t, y))]_+ \ . \end{aligned} \tag{10}$$

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

$$\rho^*(x) = \frac{1}{K} e^{-\frac{\phi(x)}{\beta}} \ , \quad \textit{where} \quad K = \sum_{x \in S} e^{-\frac{\phi(x)}{\beta}} \ , \tag{11}$$

is the unique stationary measure of ODE (10).

(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 equation* (10).

The proof follows [4, 5], so omitted here.

3.3. **FPE for general game.** For general games, as in the continuous case, the FPE can't be interpreted as gradient flows for some functional on some probability space. To establish FPEs for discrete settings, we observe that in (10), if the underlying graph corresponds to the Cartesian grid partition, (10) is exactly the numerical discretization of the continuous FPE using upwind scheme, see [6]. This motivates us to define the discrete Fokker-Planck equation.

Definition 4. *For a general game with strategy graph $G = (S, E)$ with cost functionals $u_i(x)$ for $i \in 1, \cdots, N$, define its FPE to be*

$$\begin{aligned} \frac{d\rho(t,x)}{dt} =& \sum_{i=1}^{N} \sum_{y \in \mathcal{N}_i(x)} [u_i(y) - u_i(x) + \beta(\log \rho(t,y) - \log \rho(t,x))]_+ \rho(t,y) \\ &- \sum_{i=1}^{N} \sum_{y \in \mathcal{N}_i(x)} [u_i(x) - u_i(y) + \beta(\log \rho(t,x) - \log \rho(t,y))]_+ \rho(t,x) \ . \end{aligned} \tag{12}$$

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 (10).

Our main result for general games is the following theorem.

Theorem 5 (General flow)**.** *Given a N-player game with strategy graph $G = (S, E)$, cost functional u_i, $i = 1, \cdots, N$ and a constant $\beta \geq 0$.*

(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 (12).

(ii) *Given any initial condition $\rho_0(t)$, denote $\rho^\beta(t)$ the solutions of* (12) *with varying β's. Then for any fixed time $T \in (0, +\infty)$*

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

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

$$\textit{Support of } \rho^*(x) \subset \{x^1, \cdots, x^k\} \ ,$$

then $\rho^*(x)$ is the stationary solution of (12) *with $\beta = 0$.*

Proof. (i) is a slight modification of results in [6]. (ii) Let's denote ODE (12) 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 the similar reason in proving Theorem 3, 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

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

where $\|\cdot\|$ is the 2-norm. In other words, the difference of the ODE (12)'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

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

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

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

which finishes the proof.

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

$$\rho^*(x) = \begin{cases} 0 & \text{if } x \notin \mathcal{E}; \\ \geq 0 & \text{if } x \in \mathcal{E}. \end{cases} \tag{13}$$

Since $x \in \mathcal{E}$ is a NE, $u_i(y) \geq u_i(x)$ when $y \in \mathcal{N}_i(x)$, for any $i = 1, \cdots, d$. For $x \in \mathcal{E}$, we substitute $\rho^*(x)$ into the R.H.S. (12), which forms

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

where the last equality is from the following facts in two cases. (i) If $y \notin \mathcal{E}$, $\rho^*(y) = 0$ from (13). (ii) if $y \in \mathcal{E}$, $u_i(y) \geq u_i(x)$, then $u_i(y) - u_i(x) = 0$. Similarly, we can show the case when $x \notin \mathcal{E}$. □

3.4. **Natural Order of Nash equilibria.** FPE gives the stationary distributions (equilibrium) for the dynamics [16]. It allows us to rank different equilibria by comparing the probabilities.

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

$$x^1 \prec x^2 \cdots \prec x^k, \quad \text{if} \quad \rho^*(x^1) \leq \cdots \leq \rho^*(x^k) \ . \tag{14}$$

Here $x \prec y$ is to say that the strategy y is better(more stable) than strategy x. The above definition is equivalent to look at $\phi(x^1) \geq \cdots \geq \phi(x^k)$, since $\rho^*(x) = \frac{1}{K} e^{-\frac{\phi(x)}{\beta}}$.

For non-potential games, although there is no potentials, the stationary solution of FPE $\rho^*(t)$ still provides a way of ranking equilibria.

Definition 6 (Natural order of NE)**.** *Assume*

$$\rho^*(x) = \lim_{\beta \to 0} \lim_{t \to \infty} \rho(t, x)$$

exits, where $\rho(t,x)$ is the solution of (12) *with any initial measure $\rho^0 \in \mathcal{P}_o(S)$. We define the order of NE by*

$$x^1 \prec x^2 \cdots \prec x^k, \quad \textit{if} \quad \rho^*(x^1) \leq \cdots \leq \rho^*(x^k) \ . \tag{15}$$

We will give several examples to illustrate (15) in Section 5.

4. Best-Reply Markov process

For continuous games, the FPE can be regarded as the evolution of the density function of the stochastic differential equations (6). In this section, we introduce the similar notion to discrete games, which are Markov processes. By doing so, we demonstrate the appropriate way to add white noise to a Markov process.

We start with a 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

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

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:

$$\frac{d\rho(t,x)}{dt} = \sum_{y \in \mathcal{N}(x)} [\phi(y) - \phi(x)]_+ \rho(y) - \sum_{y \in \mathcal{N}(x)} [\phi(x) - \phi(y)]_+ \rho(x) \ . \tag{16}$$

Equation (16) can be seen as the discrete version of the FPE (7) with $\epsilon = 0$ and the Markov process $X(t)$ is the discrete version of the pure gradient flows (6) with $\epsilon = 0$. To introduce white noise into the Markov process, by comparing (16) and (10), 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 (10). In other words, we define our gradient Markov process $X_\beta(t) \in S$ to be

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

The same reasoning can be applied to non-potential 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 non-potential game

$$\begin{aligned}
&\Pr(X_\beta(t+h) = y \mid X_\beta(t) = x) \\
&= \begin{cases} \sum_{i=1}^N (\bar{u}_i(y) - \bar{u}_i(x))_+ h + o(h) & \text{if } y \in \mathcal{N}(x); \\ 1 - \sum_{i=1}^N \sum_{y \in \mathcal{N}_i(x)} (\bar{u}_i(y) - \bar{u}_i(x))_+ h + o(h) & \text{if } y = x; \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{17}$$

The time evolution $\rho(t) = (\rho(t,x))_{x \in S}$ of Markov process $X_\beta(t)$ is exactly FPE (5).

The process $X_\beta(t)$ describes players' behaviors with the following distinctive features. The Markovian property of $X_\beta(t)$ reflects players' myopicity 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 the neighboring strategy, to make the next selection. Lastly, it is easily seen that players select next strategy that decrease their collective cost functionals with largest probability. This is to say players are greedy during the decision-making process.

It's 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. In addition, since ρ evolves through time, the update rule of our dynamics is not stationary, contrary to fictitious play.

4.1. **Connection with statistical physics.** In this section, we illustrate the connection between our Markov process and statistics physics by the discrete H theory [9]. 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)

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

The H theory states that for a physical phenomenon (meaning model), the relative entropy decreases along the particle's motion (player's decision process). The following theorem can be viewed as discrete H theorem for finite player games.

Theorem 7 (Discrete H theorem)**.** *Suppose $\rho(t)$ is the transition probability of $X_\beta(t)$ in potential games. Then the relative entropy decrease*

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

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

$$\frac{d}{dt}\mathcal{H}(\rho(t)|\rho^\infty) = -\beta \mathcal{I}(\rho(t)|\rho^\infty) \ . \tag{18}$$

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 (18). Substituting $\rho^\infty(x) = \frac{1}{K} e^{-\frac{\phi(x)}{\beta}}$ into the relative entropy, we observe

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

From the explicit formulation of FPE (10), we have

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

which finishes the proof. □

Besides the discrete H theorem, there is a deep connection between FPE (10) 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 8 (Entropy dissipation)**.** *Given a potential game with $\beta > 0$, $\rho^0 \in \mathcal{P}_o(S)$, there exists a constant $C = C(\rho^0, G) > 0$ such that*

$$\mathcal{H}(\rho(t)|\rho^\infty) \leq e^{-Ct}\mathcal{H}(\rho^0|\rho^\infty) \ . \tag{19}$$

The proof of Theorem 8 is presented in [5].

It is worth mentioning that the dissipation rate C in (19) connects to many interesting concepts related to geometry on finite graphs. See Ricci curvature's lower bound [8, 28] and Yano's formula reported in [5].

5. Examples

We give several examples to illustrate the model.

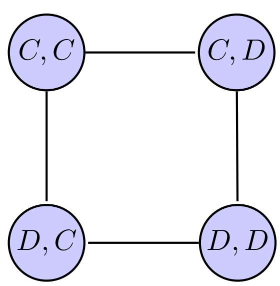
Example 1: Consider a two-player Prisoner Dilemma (A, B^T) game with cost matrix

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

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 $G = K_2 \Box K_2$.



To simplify notation, 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 (2). By numerically solving (2) 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 1.

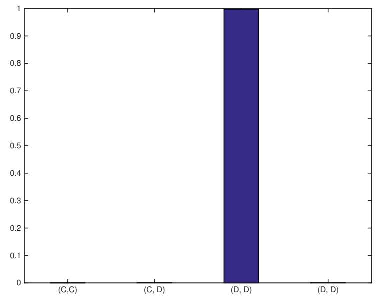
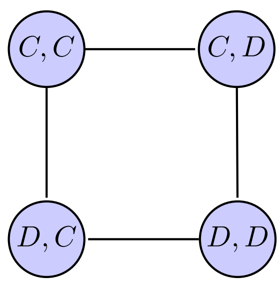


FIGURE 1. The invariant measure ρ^* for Prisoner Dilemma.

Indeed, we know that ρ^* is a Gibbs measure and (D, D) is the unique Nash equilibrium.

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



By solving (2) 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 2.

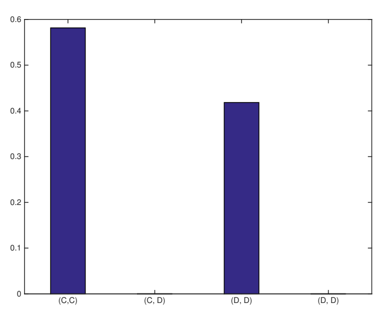


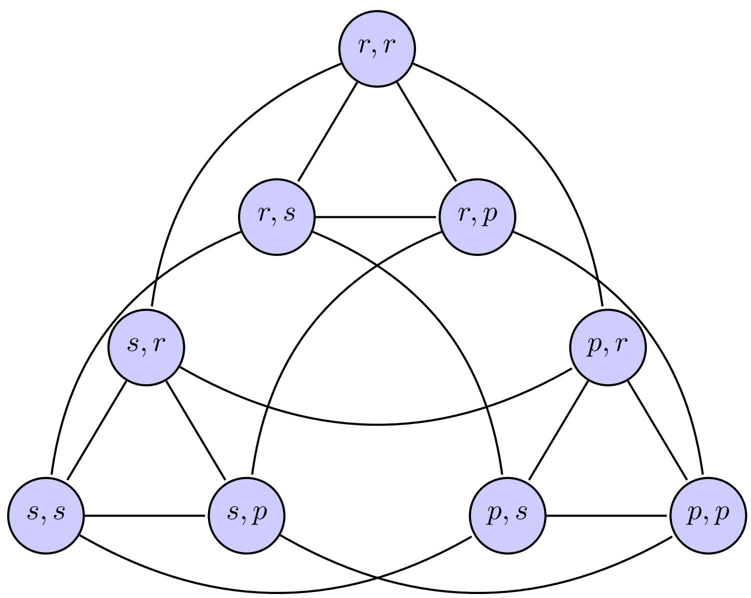
FIGURE 2. The invariant measure ρ^* for asymmetric game.

As we can see, ρ^* only supports at (C, C) and (D, D), both of which are Nash equilibria 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 the one of player 1: $u_2(C, D) - u_2(C, C) = 2 > 1 = u_1(D, C) - u_1(D, D)$.

Example 3: 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 3.

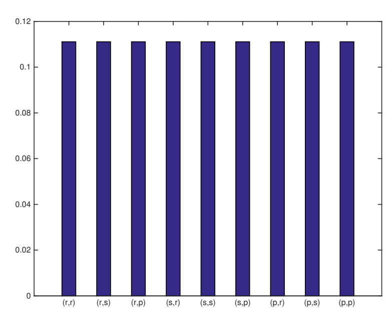


FIGURE 3. The invariant measure ρ^* for Rock-Scissors-Paper.

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 each others, at the final time, they will arrive at a state that players select strategies uniformly and independently.

Example 4. 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 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 4 while the strategy graph S_2 is a complete graph. We consider $S_1 \Box S_2$ for FPE (2) and solve for the invariant measure ρ^*.

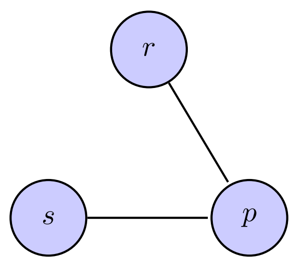


FIGURE 4. Player 1's strategy graph

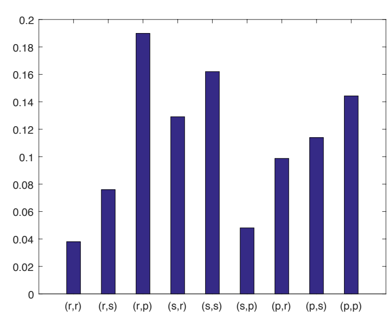


FIGURE 5. The invariant measure ρ^* for Rock-Scissors-Paper with constraints

From Figure 5, we observe several properties that accord with modeling intuitions. Firstly, player 1 is at disadvantage to player 2, since the chance of player 1 winning is less than that of player 2,

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

Secondly, we see that player 1 and 2's probabilities are not independent, meaning that they make decisions depending on each others' choices. Thirdly, 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 5 that the top three states with highest probabilities are $(r, p), (s, s)$ and (p, p).

6. Conclusion

We summary all features of the proposed dynamic framework: First, the model incorporates players myopicity, uncertainty and greedy when making decisions; Second, the model works for both potential and non-potential games. For potential games, the ranking of Nash equilibria given by the limit distribution coincides with the ranking given by the potential; For non-potential games, this ranking relates to the Morse decomposition and Conley-Markov matrix proposed in [16]; Last but not least, the FPE converges to Gibbs measure for potential games. The convergence is exponentially fast, whose rate is controlled by the relation between discrete entropy and Fisher information [5, 9].

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2006.

[2] L.E. Blume. *The statistical mechanics of strategic interaction. Games and Economic Behavior*, 1993.

[3] George W. Brown. *Iterative Solutions of Games by Fictitious Play. Activity Analysis of Production and Allocation*, 1951.

[4] Shui-Nee Chow, Wen Huang, Yao Li, and Haomin Zhou. Fokker–Planck equations for a free energy functional or Markov process on a graph. *Archive for Rational Mechanics and Analysis*, 203(3):969–1008, 2012.

[5] Shui-Nee Chow, Wuchen Li and Haomin Zhou. Nonlinear Fokker-Planck equations and their asymptotic properties, *arXiv:1701.04841*, 2017.

[6] Shui-Nee Chow, Luca Dieci, Wuchen Li and Haomin Zhou. Entropy dissipation semi-discretization schemes for Fokker-Planck equations, *arXiv:1608.02628* , 2016.

[7] Pierre Degond, Jian-Guo Liu, and Christian Ringhofer. Large-scale dynamics of mean-field games driven by local Nash equilibria. *Journal of Nonlinear Science*, 24(1):93–115, 2014.

[8] Matthias Erbar and Jan Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Archive for Rational Mechanics and Analysis*, 206(3):997–1038, 2012.

[9] B. Roy Frieden. *Science from Fisher Information: A Unification*, Cambridge University Press, 2004.

[10] Itzhak Gilboa, Larry Samuelson, and David Schmeidler. No-Betting-Pareto Dominance. *Econometrica*, 1405-1442, 2014.

[11] John C. Harsanyi. *A New Theory of Equilibrium Selection for Games with Complete Information. Games and Econonmic Behavior*, 1995.

[12] John C. Harsanyi and Reinhard Selten. *A General Theory of Equilibrium Selection in Games. MIT Press*, 1988.

[13] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.

[14] Michihiro Kandori and George Mailath and Rafael Rob. *Learning, Mutation, and Long-run equilibria in Games. Econometrica*, 1993.

[15] Wuchen Li. A study of stochastic differential equations and FPEs with applications. *PhD thesis*, 2016. Georgia Institute of Technology.

[16] Shui-Nee Chow, Weiping Li, Zhenxin Liu and Haomin Zhou. A natural order in dynamical systems based on Conley–Markov matrices. *Journal of Differential Equations*, 252, 2012.

[17] Jan Maas. Gradient flows of the entropy for finite Markov chains. *Journal of Functional Analysis*, 261(8):2250–2292, 2011.

[18] D.L. Mcfadden. Conditional Logit analysis of quantitative choice behavior. *Frontiers in Econometrics*, 1974.

[19] Dov Monderer and Lloyd Shapley. Potential games. *Games and economic behavior*, 14(1):124–143, 1996.

[20] Dov Monderer and Lloyd Shapley. Fictitious play property for games with identical interests. *Journal of economic theory*, 258-265, 1996.

[21] John Nash. Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1):48–49, 1950.

[22] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior (60th Anniversary Commemorative Edition)*. Princeton university press, 2007.

[23] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations* , 2001.

[24] William H Sandholm. Evolutionary game theory. In *Encyclopedia of Complexity and Systems Science*, pages 3176–3205. Springer, 2009.

[25] Karl Sigmund and Martin A Nowak. Evolutionary game theory. *Current Biology*, 9(14):R503–R505, 1999.

[26] Cédric Villani. A review of mathematical topics in collisional kinetic theory. *Handbook of mathematical fluid dynamics*, 1:71–305, 2002.

[27] Cédric Villani. *Topics in optimal transportation*, Number 58. American Mathematical Soc., 2003.

[28] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.