A Mean Field Game Model of Spatial Evolutionary Games

Vincent Hsiao
University of Maryland
College Park, Maryland, United States
vhsiao@umd.edu

Dana Nau
University of Maryland
College Park, Maryland, United States
nau@umd.edu

ABSTRACT

Evolutionary Game Theory (EGT) studies evolving populations of agents that interact through normal form games whose outcome determines each individual's evolutionary fitness. In many applications, EGT models are extended to include spatial effects in which the agents are located in a structured population such as a graph. We propose a Mean Field Game (MFG) generalization, denoted Pair-MFG, of the spatial evolutionary game model such that the behavior of a given spatial evolutionary game (or more specifically the behavior of its pair approximation) is a special case trajectory of the corresponding MFG. The proposed Pair-MFG model also allows for the formulation of the spatial evolutionary game as a control problem, opening up additional avenues of research into controlling the outcomes of these games.

The state evolution equations of the proposed model are highly nonlinear and none of the equations in the system are necessarily convex. This necessitates different numerical methods as compared to those for traditional Linear Quadratic Gaussian MFGs. We provide a method for solving this new Pair-MFG model using fixed point iteration with time-dependent proximal terms and show empirically that this method is capable of finding a solution to a selection of EGT games.

KEYWORDS

Mean Field Game, Spatial Evolutionary Games, Moment Closure, Pair Approximation

ACM Reference Format:

Vincent Hsiao and Dana Nau. 2022. A Mean Field Game Model of Spatial Evolutionary Games. In *Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), Online, May 9–13, 2022*, IFAAMAS, 9 pages.

1 INTRODUCTION

Evolutionary Game Theory (EGT) is a branch of game theory that studies evolving populations of agents that interact through normal form games whose outcome determines each individual's evolutionary fitness. *Spatial evolutionary games* model populations in which the players are located at the nodes of a graph and play repeatedly with their neighbors [31]. Spatial evolutionary game models have been widely used to model both biological and cultural evolution, e.g., [5, 8, 10, 29, 36] and a variety of multi-agent systems topics [1, 21, 27, 28, 34]. Most work on evaluating spatial evolutionary games is performed with computationally expensive simulations, an approach that often does not scale well for large graphs [30]. Other limitations on validation [14] and variability [19] make it difficult to apply simulations towards the analysis of many evolutionary games.

A well known approximation technique for avoiding the problems inherent to simulations is known as pair approximation. Pair approximation [10, 20, 24] (see Appendix A of [10] for detailed walkthrough) is a technique for analyzing spatial evolutionary games through the use of a system of differential equations, in which the effects that the spatial structure of the population has on individual i are approximated by the effects of a single neighbor j and the local neighborhoods of i and j.

A different paradigm of multi-agent game theory, Mean Field Games (MFG) were first introduced independently by Huang et al. [15] and Lasry and Lions [17] as a mathematical formalism for analyzing large populations of individually optimizing agents. The key idea behind this model is that interactions between agents can be computed as interactions between a representative agent and an aggregate population distribution. The MFG formalism has found widespread use in the modeling of various optimization problems such as wireless network control [23], crowd evacuation [6], vaccine distribution [18], and swarm robotics [9].

While EGT and MFG models each have many applications, there is a modeling gap in the types of problems they can be used on. Specifically, while spatial EGT models allow for the evaluation of strategy evolution on networks, they can not be used to optimize those strategies. Conversely, while MFG models are naturally optimization problems, they have not been used to model strategy evolution for populations that interact on networks.

Contribution. In this paper, we propose a Mean Field Game (MFG) generalization of the spatial evolutionary game model such that the behavior of a given spatial evolutionary game (or more specifically the behavior of its pair approximation) is a special case trajectory of the corresponding MFG. The proposed Pair-MFG model allows for the formulation of the spatial evolutionary game as a control problem, opening up additional avenues of research into controlling the outcomes of these games. The new model, denoted Pair-MFG, involves many nonlinear functions. This nonlinearity allows for complex behavior characteristic of spatial evolutionary games that simple dynamics such as the replicator equation are unable to model. However, due to this nonlinearity, it is necessary to employ different numerical methods than the ones for Linear Quadratic Gaussian MFGs.

We show how to solve the Pair-MFG model as a fixed point problem. It has been shown in previous work [3] that naive application of fixed point iteration to an initial solution pair converges in certain cases but can get stuck in others. We propose the addition a few heuristics such as time-dependent proximal terms that allow for convergence in evolutionary games where normal fixed point iteration fails. Finally, we provide empirical evaluations of the effectiveness of our approach towards solving a Pair-MFG model.

Proc. of the 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2022), P. Faliszewski, V. Mascardi, C. Pelachaud, M.E. Taylor (eds.), May 9–13, 2022, Online.

2 EVOLUTIONARY GAME THEORY BACKGROUND

Evolutionary Game Theory (EGT) provides a framework for modeling the time evolution of a population of agents that interact through strategic games whose outcome determines each individual's evolutionary fitness. These models disregard any game-theoretic assumptions of rationality and instead let individuals reproduce or change strategies based on a population update rule.

2.1 Spatial Evolutionary Game

Consider a set of agents $\{1, \ldots, M\}$ that are placed on points in a lattice/grid with a wrap-around boundary condition in which each agent has 4 neighbors. Each agent can choose some strategy $s_i \in S$ where S is some discrete state space (ex: $S = \{C, D\}$, where C is cooperate, D is defect). The spatial evolutionary game model consists of two phases:

Interaction phase. At each generation[1] each agent chooses some action $s_i \in S$ and receives a payoff π from playing game with its neighbors. A common setting for this involves each agent playing a normal-form game with each of its neighbors and obtaining a payoff equal to the sum or average of its interactions.

	C	D
C	a	b
D	c	d

Update phase. Each time the update phase occurs, a percentage of agents γ in the population use an update rule to decide whether to change strategies or how to reproduce on the grid. There are a number of different update rules used in EGT models. Some commonly studied update rules include the following:

- *Death-Birth rule*: Each agent has a non-zero chance of dying after the interaction phase. When an agent dies, it leaves behind an empty space in the grid that can be populated by new agents. Afterwards, agents that have not died have a probability of p_b to reproduce into adjacent empty spaces on the grid that depends on their fitness. [for a more detailed account, see the methods section of [4]]
- *Fermi rule*: Each agent compares its payoff π from the interaction phase with the payoff π' of a randomly chosen neighbor and switches to the neighbor's strategy with probability $p_f(\pi, \pi')$, where

$$p_f(\pi, \pi') = \frac{1}{(1 + e^{-s(\pi' - \pi)})} \tag{1}$$

  where $s > 0$ is a constant called the selection strength.
- *Best response*: Different from the two aforementioned rules, the agent does not use the payoff at the current time to inform its strategy updating. Instead an agent will evaluate:

$$\arg\max_{s_i} \pi(s_i, N(i)) \tag{2}$$

  where $\pi(s_i)$ is the payoff obtained playing strategy s_i versus its current neighbors $N(i)$.

[1]In the EGT literature, the successive steps of an EGT simulation are often called iterations [5, 35], but also are called generations [11], time steps [24], and rounds [31]. In this paper we call them generations, to avoid confusion with the iterations of the fixed-point algorithm introduced in Section 4.4.

In the evolutionary game model, the interaction and update phases are repeated iteratively until a steady state or some predetermined time horizon is reached.

2.2 EGT Approximations

EGT models have frequently been studied using systems of differential equations termed as evolutionary game dynamics. These dynamics are a continuous time approximation of the original discrete time process specified by the evolutionary game. The resulting systems of differential equations can be viewed as population-level models as they model the evolutionary game by analyzing the time evolution of the proportion of agents playing each strategy in the population p_i. For evolutionary games defined on a well-mixed population, these dynamics (such as the well known replicator dynamics) can be derived by using the master equation that corresponds to the Markov process specifying the underlying microscopic dynamics [33].

For spatial evolutionary games, applying the master equation does not produce a closed system of equations. The equations that specify the time evolution of the proportion of agents playing each strategy p_i in the population depend on the proportion of pairs of agents p_{ij} playing different strategy pairs in the population. In turn, the equations that specify the time of evolution of pairs p_{ij} will depend on the proportion of triples p_{ijk}. This leads to a hierarchy of equations defined up to proportions of groups of agents the size of the entire population:

$$\begin{aligned} \dot{p}_i &= F(p_i, p_{ij}) \\ \dot{p}_{ij} &= G(p_i, p_{ij}, p_{ijk}) \\ \dot{p}_{ijl} &= H(p_i, p_{ij}, p_{ijk}, p_{ijkl}) \\ &\vdots \end{aligned} \tag{3}$$

These systems of equations are intractable to solve given a large enough population size. Consequently, there is much work in past literature [10, 13, 22, 25, 26, 31, 32] in which higher order proportions are approximated using lower order proportions. For example, [10] defines a pair approximation where second order conditional probabilities are approximated using first order conditional probabilities as follows:

$$\begin{aligned} p_{i|jk} &\approx p_{i|j} \\ p_{i|jk} p_{jk} p_j &\approx p_{i|j} p_{jk} p_j \\ p_{ijk} &\approx \frac{p_{ij} p_{jk}}{p_j} \end{aligned} \tag{4}$$

This idea of approximating higher order terms by lower order terms is an approximation technique known as moment closure [16].

2.3 Pair Approximation

In pair approximation, the time evolution of the population is modeled using a set of differential equations that use global (p_i) and local (p_{ij}) density terms:

$$\begin{aligned} \dot{p}_i &= F_1(p_i, p_{ij}) \\ \dot{p}_{ij} &= F_2(p_i, p_{ij}) \end{aligned} \tag{5}$$

The functions are best expressed as summations over configurations cf of neighborhood strategy assignments. More concretely, for an arbitrary agent x and its neighborhood $\{1, \ldots, N\}$, a configuration $cf = \{s_1, \ldots, s_N\}$ denotes the assignment of strategies to all neighbors of the agent x. Using the master equation, we have:

$$\dot{p}_i = \sum_{cf} \left[\sum_{j \in S} P_{cf}(j) P(j \to i, cf) - P_{cf}(i) P(i \to j \mid cf) \right]$$
$$P(i \to j \mid cf) = P\big(x(t+1) = x_j \mid x(t) = x_i, cf\big) \tag{6}$$

where an update rule (see section 2.1) is used to calculate the probability of an arbitrary agent changing its strategy conditioned on its local neighborhood. The key idea in pair approximation is the method for evaluating P_{cf} in Eq. 6. The true value of P_{cf} is a joint distribution over N variables comprising the neighborhood of an arbitrary agent x. For a configuration $cf = \{s_1, \ldots, s_N\}$, this is the probability $P(s_1, \ldots, s_N)$. In pair approximation, we approximate this probability using first and second order terms. For example, one possible second order approximation is:

$$P(cf) = P(s_1, \ldots, s_N) = p_{s_2|s_1} p_{s_3|s_1} \ldots, p_{s_N|s_1} p_{s_1} \tag{7}$$

This idea of conditioning on local configurations will be a key component of our proposed MFG model.

3 MEAN FIELD GAMES BACKGROUND

Mean Field games (MFG) are a formalism that can be used to study interactions between large populations of agents. Unlike EGT models, each agent directly optimizes a cost function with respect to their control variables (similar in concept to strategies in the EGT model). However, in EGT models, agents do not perform this optimization directly and even in best response dynamics where some optimization is performed, it is not performed over the entire time horizon. As a result, MFGs provide a different solution concept with respect to EGTs that can be more useful in certain applications.

We emphasize that our goal is to fill the MFG modeling gap for populations that have a spatial structure but interact using a discrete strategy set. This is notably different from diffusion based MFGs in which the spatial component is used as the state space. For the purpose of this paper, we provide an overview of the Linear Quadratic MFG framework from which we will use as a guide for developing our proposed model.

3.1 Continuous Linear Quadratic Gaussian MFG

In the Linear Quadratic MFG model [2], we are interested in a population of agents $\{1, \ldots, n\}$ where each agent i individually minimizes a cost function:

$$E\Big[\frac{1}{2}\int_0^T x_t^T F_t x_t + (x_t - \overline{x}_t)^T \overline{F}_t (x_t - \overline{x}_t) + v_t^T L_t v_t dt$$
$$+ \frac{1}{2} x_T^T G_T x_T + (x_T - \overline{x}_T)^T \overline{G}_t (x_T - \overline{x}_T)\Big] \tag{8}$$

where:

- x_t is a vector specifying the agent's state
- $\overline{x}_t$ is a vector specifying the average of all agent states
- v_t is a vector specifying the agent's control
- F_t is a matrix specifying the quadratic running cost of being at state x_t
- L_t is a matrix specifying a quadratic transport cost of using control v_t
- $\overline{F}_t$ is a matrix specifying the quadratic running cost of being at state x_t with respect to the population average $\overline{x}_t$
- G_T and $\overline{G}_T$ being the terminal cost versions of F

The position of each agent changes according to a state evolution equation:

$$dx_t = \left(A_t x_t + B_t v_t + \overline{A}_t \overline{x}_t\right) dt + \sigma_t dW_t \tag{9}$$

where $A_t, \overline{A}_t, B_t$ are some matrices that describe the movement of the agent as a function of its control v_t, current state x_t and the average of other agent states $\overline{x}_t$. The last term $\sigma_t dW_t$ can be used to describe additional Gaussian noise present in the state evolution equation.

3.2 Discrete Linear Quadratic MFG

Consider the discrete version of the LQG-MFG model presented in section 2 of [12]. Suppose each agent x can be in a set of states $s_i \in S$ and the state of a player evolves according to a controlled Markov process:

$$P\big(x(t+h) = s_j \mid x(t) = s_i\big) = (\alpha_{ij}(t) + \eta_{ij})h + o(h) \tag{10}$$

for some $h > 0$. Each agent specifies transition controls α_{ij} for all $s_i, s_j \in S, i \neq j$ and each agent tries to minimize the objective function:

$$\min_{\alpha} E\left[\int_0^T C(x(t), \theta(t), \alpha(t)) dt + G(x(T), \theta(T))\right] \tag{11}$$

where $\theta(t)$ describes the distribution of other players in each state at time t and the cost function C is separated into the running cost F and a quadratic energy cost:

$$C(x(t), \theta(t), \alpha(t)) = F(x(t), \theta(t)) + \frac{1}{L}\frac{\|\alpha(t)\|^2}{2} \tag{12}$$

where L is some constant and

$$\|\alpha(t)\|^2 = \sum_{i \neq j} \alpha_{i,j}(t)^2 \tag{13}$$

3.2.1 *Well-mixed EGT.*

Given the above discrete framework, it is easy to formulate an MFG model analogous to an EGT model for a well mixed population. We define $F(x(t), \theta(t))$ to be:

$$F(x(t), \theta(t)) = \sum_i \mathbf{P}_{xi} \theta_i \tag{14}$$

where $\mathbf{P}_{xi}$ is the (x, i)-th entry of the payoff matrix $\mathbf{P}$. The running cost is computed against $\theta(t)$ which describes the average state distribution of the entire population. In essence, an agent will receive a cost from playing a game with a well-mixed population.

4 PAIR APPROXIMATION MFG

We make an observation that, in the discrete MFG framework, the mean field $\theta(t)$ represents the distribution of a single agent. In EGT literature, this is the same property as evolutionary dynamics defined on a well-mixed population such as the replicator dynamics. As mentioned in Section 2.2, pair approximation is a natural extension of evolutionary dynamics to structured populations where the equations use higher order distributions.

Suppose then, in a manner similar to pair approximation, we define a pair-level mean field distribution $\theta(t)$ over $S \times S$. We denote $\theta_{ij}(t)$ to be the proportion of pairs of agents that are playing strategy i and j at time t. We define our running cost F for a single agent in the state x to be:

$$F(x(t), \theta(t)) = \sum_i \mathbf{P}_{xi} \frac{\theta_{xi}(t)}{\theta_x(t)}$$
$$\theta_x(t) = \sum_j \theta_{xj}(t) \tag{15}$$

Like in the discrete LQ-MFG model we define a state evolution equation based on a Markov process. Unlike the well-mixed model, we define a more complicated state evolution equation consistent with our pair-level mean field:

$$P(x(t+1) = s_j | x(t) = s_i) = \sum_{cf \in C_d} P_{cf}(i,t) \cdot (\alpha_{ij}(t, cf) + \eta_{ij}) \cdot \gamma \tag{16}$$

where C_d is the set of possible local neighborhood assignment configurations for an agent with d neighbors and γ is the percentage of agents that can change their strategies during each generation. More rigorously, we define C_d as the set of tuples:

$$C_d = \{(x_1, x_2, \ldots, x_{|S|}) \ s.t. \ \sum x_i = d\} \tag{17}$$

where x_i denotes the number of neighbors of the agent x that are playing strategy i and d is the degree of the network. We constrain the control variables $\alpha_{ij}(t, cf) \in [0,1]$ to ensure that equation 16 produces a valid probability distribution. $P_{cf}(i)$ denotes the probability of the configuration cf and is a function of $\theta(t)$ and we can approximate it by assuming all neighbors are independently distributed:

$$P_{cf}(i,t) = \binom{d}{x_1, x_2, \ldots, x_{|S|}} \prod_j \left(\frac{\theta_{ij}(t)}{\theta_i(t)}\right)^{x_i}$$
$$cf = (x_1, x_2, \ldots, x_{|S|}) \tag{18}$$

In this setting, we define controls $\alpha_{ij}(cf)$ for each transition $i \to j$ that is some function of cf. From Eq. 16, using pair approximation, we can derive the time evolution of joint probabilities:

$$\theta_{ij}(t+1) = \theta_{ij}(t) + \Delta\theta_{ij}(\alpha, t)$$
$$\Delta\theta_{ij}(\alpha, t) = \sum_k \sum_{cf \in C_d} \Big[\theta_k(t) P_{cf}(k,t) \frac{N(j, cf)}{d} \alpha_{ki}(t, cf)\gamma$$
$$- \theta_i(t) P_{cf}(i,t) \frac{N(j, cf)}{d} \alpha_{ik}(t, cf)\gamma\Big] \tag{19}$$

where $N(j, cf)$ denotes the number of nodes assigned j in neighborhood configuration cf or the x_j value in the cf tuple. To simplify notation for further discussion, let M denote the mean field evolution equations that produce $\theta(t)$ according to Eq. 19:

$$\theta = M(\alpha, \theta(0)) \tag{20}$$

4.1 An Example

Let us define a pair-MFG for modeling an evolutionary game on a square lattice grid ($d = 4$) with two strategies $[0,1]$ and the following payoff matrix:

$$\mathbf{PAY} = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \tag{21}$$

For our MFG model we want low costs for an agent to correspond to high payoffs in the original EGT. A valid transformation for this mapping is to subtract **PAY** from the maximum value of **PAY** and add a small constant cost, $\mathbf{P} = \max(\mathbf{PAY}) - \mathbf{PAY} + 1$:

$$\mathbf{P} = \begin{bmatrix} 2 & 5 \\ 1 & 4 \end{bmatrix} \tag{22}$$

Our mean field distribution is defined as four values: $\theta_{00}, \theta_{01}, \theta_{10}, \theta_{11}$ that sum to 1 and two additional computed values: $\theta_0 = \theta_{00} + \theta_{01}$, $\theta_1 = \theta_{10} + \theta_{11}$ Our running cost F is:

$$F(x(t), \theta(t)) = \begin{cases} \frac{1}{\theta_0}(2\theta_{00} + 5\theta_{01}) & x = 0 \\ \frac{1}{\theta_1}(1\theta_{10} + 4\theta_{11}) & x = 1 \end{cases} \tag{23}$$

and our terminal cost is the same: $G(x(T), \theta(T)) = F(x(T), \theta(T))$

Since our model only has two strategies, a configuration cf can be represented with just one variable $c \in [0,4]$ that denotes the number of agents in the neighborhood playing the first strategy. As an example, we have that θ_{00} evolves as such:

$$P\big(x(t+1) = 1 \mid x(t) = 0\big) = \sum_{c=0}^{4} P_c(0) \cdot \alpha_{01}(t,c)\gamma$$
$$\Delta\theta_{00} = \sum_{c=0}^{4} \theta_1 P_c(0) \frac{c}{4} \alpha_{10}(t,c)\gamma - \theta_0 P_c(0) \frac{c}{4} \alpha_{01}(t,c)\gamma$$
$$P_c(0) = \binom{4}{c} \left(\frac{\theta_{00}}{\theta_0}\right)^c \left(\frac{\theta_{01}}{\theta_0}\right)^{4-c} \tag{24}$$

The transition probabilities and equations for other θ_{ij} are similarly defined. Notice that because of how $P_c(0)$ is computed using higher order terms of θ_{00}, the resulting dynamics are nonlinear and our model will consequently lie outside of the scope of typical LQG-MFG models.

4.2 Features

4.2.1 Generalization of best response. The model is a generalization of a spatial evolutionary game using the best response update rule. To see how this is the case, suppose we limit our pair-MFG model to one generation $t \in \{0, 1\}$.

Claim. There exists a set of controls α (and cost functions C, G) such that the pair-MFG model is equivalent to a pair approximation model of a best response spatial evolutionary game.

Proof. Consider a best response spatial evolutionary game. A given agent x playing strategy i will change its strategy based on the strategies of its neighbors. This is a deterministic function R of a given neighborhood configuration cf:

$$R(i, j, cf) = \begin{cases} 1 & Pay(j, cf) > Pay(k, cf) \ \forall k \in S \\ 0 & \text{otherwise} \end{cases} \tag{25}$$

The above function is the probability in the best response model that a given agent x playing i will change its strategy to j. In our pair-MFG formulation, the control term α models the probability of an agent changing its strategy given a neighborhood configuration cf. If we let our control terms be:

$$\alpha_{ij}(cf) = R(i, j, cf) \tag{26}$$

the equations 16 and 19 are then equivalent to the pair approximation equations for a best response spatial evolutionary game. Now consider the cost function:

$$\alpha^* = \arg\min_{\alpha} E\left[C(x(0), \theta(0), \alpha(0))dt + G(x(1), \theta(1))\right] \tag{27}$$

Replacing the C and G terms:

$$\begin{aligned}\alpha^* &= \arg\min_{\alpha} E\left[F(x(0), \theta(0)) + \frac{1}{L}\frac{\|\alpha(t)\|^2}{2} + G(x(1), \theta(1))\right] \\ &= \arg\min_{\alpha} E\left[\frac{1}{L}\frac{\|\alpha(t)\|^2}{2} + F(x(1), \theta(1))\right]\end{aligned} \tag{28}$$

If we make two additional assumptions for the terms in C, G:

- let L be a large enough such that $\frac{1}{L}\frac{\|\alpha(t)\|^2}{2} << G(x(1), \theta(1))$ or equivalently that the energy term does not significantly impact the minimum of our cost function,
- let $\gamma << 1$, so that $\theta(1) \approx \theta(0)$,

we have that

$$\alpha^* \approx \arg\min_{\alpha} E\left[F(x(1), \theta(0))\right] \tag{29}$$

which is the decision rule for a best response spatial evolutionary game. As $\gamma \longrightarrow 0$, the closer the pair-MFG model approaches the best-response spatial evolutionary game.

4.2.2 Beyond best response. A key concept in the pair-MFG model is that the agent's control is their transition probability given a certain neighborhood configuration. The original spatial evolutionary game is a special case of the pair-MFG where the "control" probabilities are determined using the update rule. Following this logic, we can model any update rule that can be specified by local configurations of strategy assignments. In section 4, we defined our configuration space as:

$$cf \in C_d,\ C_d = \{(x_1, x_2, \ldots, x_{|S|})\ s.t.\ \sum x_i = d\} \tag{30}$$

For the Fermi rule, we can define a configuration space:

$$cf \in S \times C^2_{d-1},\ S \times C^2_{d-1} = \{(o, n_1, n_2) | o \in S, n_1, n_2 \in C_{d-1}\} \tag{31}$$

The new cf defines a neighbor o and the neighborhood distributions of the other $d-1$ neighbors of our agent and the $d-1$ neighbors of o. By replacing C_d in equations 16 and 19 with the new configuration space, we obtain a pair-MFG model that generalizes spatial evolutionary games that use the Fermi rule. Like for best response, there exists values of the control $\alpha_{ij}(cf)$ such that the pair-MFG model is equivalent to the pair approximation equations for the Fermi rule spatial evolutionary game.

4.3 Solving the Model

Observe that the Pair-MFG model is not necessarily monotone. Existing mean field game models frequently make the assumption that the interactions between agents are monotone so that the model has a unique solution. In spatial models this translates to agents being crowd-averse. Some evolutionary games such as coordination games have crowd-following behavior which can result in more than one equilibria. As a result, the solution to a Pair-MFG may not necessarily be unique.

We define a solution to our pair-MFG model as a pair: (θ^*, α^*)

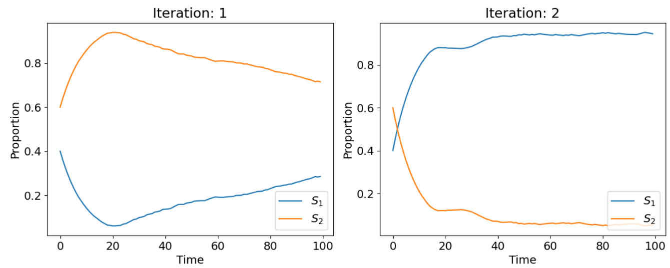

Figure 1: Alternating iterations: the left graph is the trajectory induced by the optimal response to the mean field of the right graph, and vice versa.

such that:

$$\begin{aligned}\alpha^* = J(\theta^*) &= \arg\min_{\alpha} E\left[\sum_t^{T-1} C(x(t), \theta^*(t), \alpha(t)) + G(x(T), \theta(T))\right] \\ P(x(t+1) = s_j \mid x(t) = s_i) &= \sum_{cf \in C_d} P_{cf}(i, t) \cdot \alpha_{ij}(t, cf) \cdot \gamma \\ \theta^* &= M(\alpha^*, \theta(0))\end{aligned} \tag{32}$$

given an initial condition $\theta(0)$.

Existence. Our optimization problem is not necessarily convex given certain specifications of the payoff matrix **PAY** and subsequently the cost function C, but because both θ and α are defined on convex compact sets there exists some (θ, α) pair such that the above equations are satisfied. Using the Brouwer/Kakutani fixed point theorem, it is easy to see that (θ^*, α^*) is a fixed point of the function:

$$(\theta, \alpha) \longrightarrow (M(J(\theta)), J(\theta)) \tag{33}$$

which ensures that there exists a solution to the Pair-MFG model.

4.4 Fixed Point Iteration

Since we know that the optimal pair (θ^*, α^*) must be a fixed point of the mapping in Eq. 33, we can try iterating the function until convergence. However, this approach is not guaranteed to converge, since we have no guarantee that Eq. 33 is a contraction mapping. In fact, in [3], it was shown that in general, Eq. 33 is not a contraction mapping for any finite MFGs. As a result, while on some evolutionary games (such as the Prisoner's Dilemma), naive fixed point iteration converges to a solution, there are many other situations where the fixed point iterations will start to alternate between two different paths. An example of this alternating behavior is shown in Fig. 1 for the Hawk-Dove type payoff matrix:

	C	D
C	2	1
D	3	0

4.4.1 Adjustments. We propose several novel adjustments that can be made to the naive fixed point iteration method that can improve its convergence.

Step-size. The problem encountered by the fixed-point iteration is similar to the problem occasionally encountered by hill-climbing optimization routines without a step-size adjustment (eg. line search or backtracking). Assume that some black-box iterative optimization routine is used to optimize $J(\theta)$. Instead of running this optimization routine until convergence for each fixed point iteration, we can limit the number of iterations in the inner optimization loop and instead re-evaluate M more often. The number of inner black-box optimizer iterations then serves as the "step-size" for the outer fixed point iteration.

Proximal terms. A common technique studied in general nonlinear optimization is the proximal point method [7] which is a method for generating optimization sub-problems that are well conditioned and thus easier to solve through the addition of a quadratic penalty term relative to the previous iterate. Suppose we consider each fixed point iteration as a separate sub-problem in an algorithm for finding the solution of Eq. 32. At each iteration we add a proximal term to the objective function J that penalizes iterations α^n that move too far from the previous iteration α^{n-1}:

$$\alpha^{n+1} = J'(\theta^n) = \arg\min_{\alpha} E\big[\big(\sum_{t}^{T-1} C(x(t), \theta^n(t), \alpha(t)) + \frac{1}{2}\|\alpha(t) - \alpha^n(t)\|\big) + G(x(T), \theta(T))\big] \quad (34)$$

However, since we apply a proximal term at each iteration, the proximal terms closer to $t = 0$ have a much larger effect than those later in time. Small changes in the trajectory early on can disproportionally impact the later trajectory and cost obtained.

Time-dependent Proximal terms. Due to the property where small perturbations near $t = 0$ can drastically change later trajectories, simply adding a proximal term at each iteration is not likely to help the method converge. Since changes earlier in the time interval have more impact on the final trajectory, we will want to penalize early changes more than later changes. We define a new time-dependent proximal term:

$$\frac{1}{2W(t)}\|\alpha(t) - \alpha^n(t)\| \quad (35)$$

such that the function $W(t)$ is a monotonically decreasing function of time.

Accumulating Proximal terms. An alternative approach for penalizing drastic changes is to accumulate differences backwards in time. If large changes have already been made at a later time, earlier changes should be penalized more. Several approaches are possible such as:

- A new function $W(t, \delta(i))\ \forall i > t$, where $\delta(i) = \alpha(t) - \alpha^{n-1}(t)$ are the differences computed from the current backward pass
- A new function $W_A(acc)$ where $acc = \sum_\delta(i)$ multiplied with the proximal term: $\frac{1}{2W(t)}\|\alpha(t) - \alpha^n(t)\|W_A(acc)$

5 EVALUATION

We evaluate the fixed point iteration method with proximal heuristics on the Pair-MFG model for several common games used in EGT models. The EGT payoff matrices for each game are listed in Table 1. In order to evaluate whether the fixed point iteration method can obtain a solution we compute the optimal response difference Δ_{opt}. Given iterates (θ^i, α^i) at iteration i of the fixed point algorithm, the optimal response difference Δ_{opt} is the difference between the two costs:

- Cost O_m obtained by an agent following the current controls α^i with respect to the induced mean field θ^i
- Cost O_d obtained by an optimal defecting agent who uses strategy $J(\theta^i)$ (where J is the Bellman equation with no proximal heuristics) against the mean field θ^i

$$\Delta_{opt} = O_m - O_d \quad (36)$$

It is easy to see that $\Delta_{opt} = 0$ only when a solution is found as this means a defecting agent has no incentive to unilaterally deviate from the current control iterate (the same concept as a Nash equilibrium). This quantity is similar to the notion of policy exploitability

Algorithm 1: Fixed-Point Iteration with Time-dependent Proximal terms
Input: α, θ_0
Output: α^*, θ^*

```
for t ← 1 to T do
    θ_t = M(α_{t−1}, θ_{t−1}) ;
end
Initialize V(x,t), ∀t ;
while α, θ not converged do
    V(x,T) = G(x, θ_T) ;
    α_p = α, acc = 0 ;
    for t ← T − 1 to 0 do // Bellman backwards
        let: J_{x,t}(α) = C(x, θ_t, α) + Σ_y P(y|x, α)V(y, t + 1) +
            1/(2W(t)) ||α − α_{t,p}|| W_A(acc) ;
        α_t = arg min_α J_{x,t}(α) ;
        V(x,t) = J_{x,t}(α_t) ;
        acc = acc + 1/2 ||α_t − α_{t,p}||
    end
    for t ← 1 to T do // Pair Approximation forwards
        θ_t = M(α_{t−1}, θ_{t−1}) ;
    end
end
```

Table 1: Evolutionary Game Payoff Matrices

Game Name	Payoff Matrix
Prisoner's Dilemma	$\begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix}$
Hawk-Dove	$\begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$
Snowdrift	$\begin{pmatrix} 0 & -3 \\ 3 & -10 \end{pmatrix}$
Rock-Paper-Scissors	$\begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix}$

in section 6 of [3]. We will now demonstrate empirical results where the fixed point iteration method with proximal heuristics can reduce this quantity.

5.1 Experimental Setup

We evaluate the Pair-MFG model over configuration space C_4 over a discrete time span of $[0, 100]$. The initial proportions of each strategy in each game is set to non-equilibrium values. For comparison, we provide the results of best-response dynamics on a regular grid of degree 4 over a time span of $[0, 100]$. We set $\gamma = 0.1$, so it is expected that the trajectories found as a solution to the Pair-MFG model will differ from the EGT model. For the fixed point iteration solver, we set our proximal heuristics to be:

$$W(t) = \left(\frac{1}{t/300+1}\right)^2$$
$$W_A(acc) = acc + 1$$

5.2 Results

A solution to a MFG model (and thus the Pair-MFG model) is different compared to the solution of a spatial evolutionary game. Agents in the evolutionary game do not directly optimize their payoff over the entire time-horizon. Intuitively, the trajectories obtained by the Pair-MFG model are more optimal with respect to the possible trajectories of a rogue defecting/invading agent.

It is well known that Evolutionary Stable Strategies (ESS) are an equilibrium refinement of Nash Equilibrium. It is the same case here where only under certain conditions (as discussed in section 4.2.1) is it true that the models have equivalent trajectories. However, many of the behaviors of the spatial EGT simulation are preserved in the corresponding MFG model.

For example, in the Snowdrift game, a pull back after an initial increase in the proportion of the first strategy is observed in the Pair-MFG trajectory. This pull back is a result that cannot be obtained in a well-mixed population model for two strategies, but can frequently appear in spatial models.

Another example of the preservation of spatial effects can be seen in the Rock-Paper-Scissors game. For the specific RPS payoff matrix used, the replicator equation and other well-mixed models predict a limit cycle where the magnitude of the cyclic waves for each strategy remains constant. In a spatial population, the population will instead converge to the Nash Equilibrium of $\{1/3\ R, 1/3\ P, 1/3\ S\}$ in which the magnitude of the cycle waves quickly dampen as the population evolves towards equilibrium.

As seen in the optimal response difference curves in the bottom row of Fig. 2, the proximal heuristics ensure that the optimal response difference values do not get stuck in a limit cycle. By adjusting the optimization rate for the control variables, we can avoid one of the major problems found in the application of fixed point iteration to solving mean field games.

6 MEAN FIELD TYPE CONTROL

Suppose we directly optimize Eq. 32 for α under the assumption that the joint state θ is controlled by the optimizer:

$$\alpha^* = \arg\min_\alpha E\left[\sum_t^{T-1} C(x(t), \theta(t), \alpha(t)) + G(x(T), \theta(T))\right]$$
$$\theta = M(\alpha, \theta(0)) \tag{37}$$

This transforms the problem into a single agent optimal control problem where the state consists of both the agent state x and the global joint state θ. This formulation is also known as a mean field type control (MFTC) problem. Note that if we let α' be the solution to the above optimal control problem and $\theta' = M(\alpha')$, we are not guaranteed that $\alpha' = J(\theta')$. The cost quantity in the objective function is likely to obtain a lower value than the one in Eq. 32. The MFTC model is analogous to having a central planner dictating individual agent strategies and is more akin to the idea of a pareto optimality rather than a nash equilibrium.

This relaxation of a Pair-MFG model, which we would denote Pair-MFTC, is applicable to many multi-agent optimization problems. One possible application is reactive risk management for large-scale intrusion detection. For example, using a Pair-MFTC model, it becomes possible to efficiently optimize over a spatial versions of the susceptible-infected-susceptible (SIS) model that more accurately captures network connectivity compared to existing well-mixed SIS models. A solution to the Pair-MFTC can be used to find optimal decentralized security policies for each agent to follow based on the infection status of their local network dependencies.

6.1 Reinforcement Learning

Suppose we want to frame the Pair-MFG model as reinforcement learning problem. In the case of Pair-MFTC, we can define a reward function r simply by using the negative cost function from the optimal control problem:

$$r(i, (x, \theta_t), \alpha) = -C(x, \theta_t, \alpha) \tag{38}$$

This is a natural reformulation of the MFTC problem as a reinforcement learning problem. However, if we want to solve the Pair-MFG instead of the Pair-MFTC, simply optimizing Eq. 37 can result in rogue agents that defect from the central strategy. If we want to correct for this issue, we will need to subtract the the reward obtained by the rogue agent from the reward obtained from using central planner's solution. This essentially turns the MFG problem into a two player game between a central planner and the optimal defector. However, this presents a few issues:

- Since the Bellman equation must be solved from time T backwards for an optimal strategy, the optimal defecting strategy cannot be computed until time T.
- Solving the optimization problem over the time horizon for the optimal defecting strategy can be expensive (relative to the cost of running the forward pass).

Myopic Relaxation Given the above issues, instead of computing an exact cost function with an optimal defector, we can approximate the optimal defecting strategy with a best response defector.

$$r(i, (x, \theta_t), \alpha) = -C(x, \theta_t, \alpha) + \inf_x C(x, \theta_{t-1}, \alpha) \tag{39}$$

Recall that as $\gamma \to 0$, the solution to the pair-MFG model approaches that of a best response evolutionary game.

7 CONCLUSION

We have proposed Pair-MFG, a MFG generalization of the spatial evolutionary game models. The proposed Pair-MFG model allows for the formulation of the spatial evolutionary game as a control problem, opening up additional avenues of research into controlling

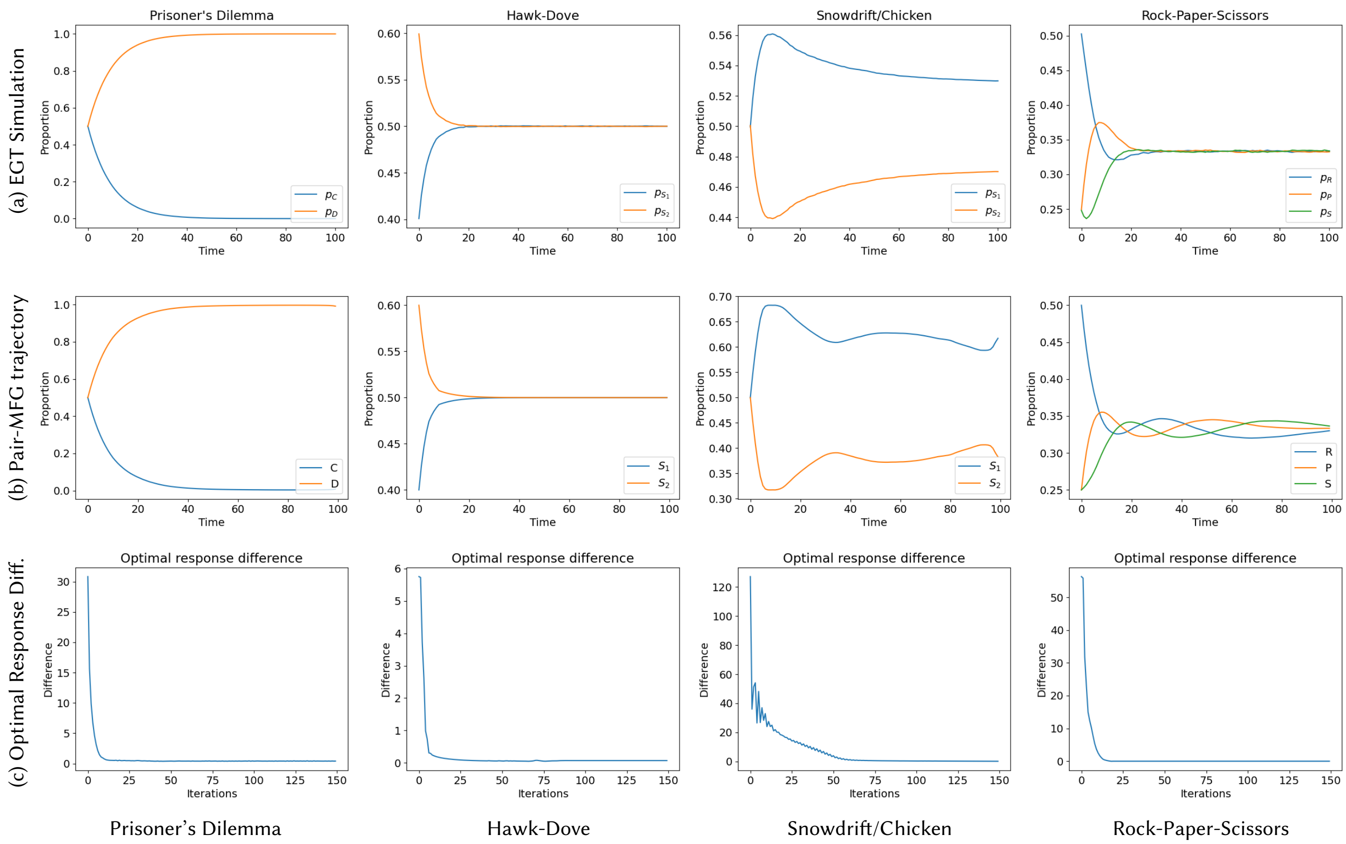

Figure 2: Comparisons of (a) best-reponse EGT simulations with (b) Pair-MFG trajectories after 150 iterations of the fixed-point algorithm, in four evolutionary games. To demonstrate convergence of the fixed-point algorithm, row (c) shows the optimal-response difference at each of its iterations. The x-axes in row (c) denote iterations of the fixed-point algorithm (as opposed to iterations of the evolutionary game in parts (a) and (b)).

the outcomes of these games. We have provided a walkthrough of the derivation of a Pair-MFG model from the equivalent EGT model and shown that the behavior of a given spatial evolutionary game (or more specifically the behavior of its pair approximation) is a special case trajectory of the corresponding MFG.

We have provided a method for solving this new Pair-MFG model using fixed point iteration with time-dependent proximal terms and show empirically that this method is capable of finding a solution to a selection of EGT games. However, since we cannot ensure the mappings are contractive, we can make no guarantees on the computational complexity required to solve a Pair-MFG model. Nevertheless, for two-strategy games, with low inner optimizer iterations, we have empirically observed that solving a Pair-MFG model is faster than multiple runs of a simulation on a decently sized population. While this is not the case with larger games, this is an area for future research. For example, we can define transition functions parameterized by a small number of agent controls to reduce the dimension of the Bellman equation. Such action space projections for computational improvements on large games is a direction we intend to explore.

Alternatively, it would be also interesting to investigate solving Pair-MFG models through extensions of existing methods with convergence guarantees such as Differential Dynamic Programming or through saddle point optimization on the derived Hamilton-Jacobi Bellman equation. For games with a small number of strategies, it may be possible to even explicitly derive equations for a Primal-Dual Hybrid Gradient (PDHG) method to solve a Pair-MFG model. These other methods present challenges in that they are typically formulated with several assumptions that a Pair-MFG model can violate. While our model does not satisfy any of the assumptions needed for nice guarantees (monotonicity, convexity, etc.), there has been work [18] in which PDHG has been applied to models that likewise violate these assumptions with reasonable effectiveness.

ACKNOWLEDGMENTS

This work has been supported in part by AFOSR grant 1010GWA357 and ONR grant N000142012257. The information in this paper does not necessarily reflect the position or policy of the funders, and no official endorsement should be inferred.

REFERENCES

[1] Robert L Axtell. 2002. Non-cooperative dynamics of multi-agent teams. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3*. 1082–1089.

[2] Alain Bensoussan, KCJ Sung, Sheung Chi Phillip Yam, and Siu-Pang Yung. 2016. Linear-quadratic mean field games. *Journal of Optimization Theory and Applications* 169, 2 (2016), 496–529.

[3] Kai Cui and Heinz Koeppl. 2021. Approximately solving mean field games via entropy-regularized deep reinforcement learning. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 1909–1917.

[4] Soham De, Michele J Gelfand, Dana Nau, and Patrick Roos. 2015. The inevitability of ethnocentrism revisited: Ethnocentrism diminishes as mobility increases. *Scientific reports* 5, 1 (2015), 1–7.

[5] Soham De, Dana S Nau, and Michele J Gelfand. 2017. Understanding norm change: An evolutionary game-theoretic approach. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. 1433–1441.

[6] Boualem Djehiche, Alain Tcheukam, and Hamidou Tembine. 2017. A mean-field game of evacuation in multilevel building. *IEEE Trans. Automat. Control* 62, 10 (2017), 5154–5169.

[7] Dmitriy Drusvyatskiy. 2017. The proximal point method revisited. *arXiv preprint arXiv:1712.06038* (2017).

[8] Richard Durrett and Simon Levin. 1994. The importance of being discrete (and spatial). *Theoretical population biology* 46, 3 (1994), 363–394.

[9] Karthik Elamvazhuthi and Spring Berman. 2019. Mean-field models in swarm robotics: A survey. *Bioinspiration & Biomimetics* 15, 1 (2019), 015001.

[10] Feng Fu, Martin A Nowak, and Christoph Hauert. 2010. Invasion and expansion of cooperators in lattice populations: Prisoner's dilemma vs. snowdrift games. *Journal of theoretical biology* 266, 3 (2010), 358–366.

[11] Herbert Gintis. 2009. *Game Theory Evolving, Second Edition*. Princeton University Press.

[12] Bruce Hajek and Michael Livesay. 2019. On non-unique solutions in mean field games. arXiv:1903.05788 [math.OC]

[13] Christoph Hauert and György Szabó. 2005. Game theory and physics. *American Journal of Physics* 73, 5 (2005), 405–414.

[14] Benjamin Herd, Simon Miles, Peter McBurney, and Michael Luck. 2013. Verification and validation of agent-based simulations using approximate model checking. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*. Springer, 53–70.

[15] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. 2006. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems* 6, 3 (2006), 221–252.

[16] Christian Kuehn. 2016. Moment closure—a brief review. In *Control of self-organizing nonlinear systems*. Springer, 253–271.

[17] Jean-Michel Lasry and Pierre-Louis Lions. 2007. Mean field games. *Japanese journal of mathematics* 2, 1 (2007), 229–260.

[18] Wonjun Lee, Siting Liu, Wuchen Li, and Stanley Osher. 2021. Mean field control problems for vaccine distribution. arXiv:2104.11887 [math.OC]

[19] Gianluca Manzo and Toby Matthews. 2014. Potentialities and limitations of agent-based simulations. *Revue française de sociologie* 55, 4 (2014), 653–688.

[20] Hirotsugu Matsuda, Naofumi Ogita, Akira Sasaki, and Kazunori Satō. 1992. Statistical mechanics of population: the lattice Lotka-Volterra model. *Progress of theoretical Physics* 88, 6 (1992), 1035–1049. https://academic.oup.com/ptp/article/88/6/1035/1854847

[21] Javier Morales, Michael Wooldridge, Juan A Rodríguez-Aguilar, and Maite López-Sánchez. 2018. Off-line synthesis of evolutionarily stable normative systems. *Autonomous agents and multi-agent systems* 32, 5 (2018), 635–671.

[22] Satoru Morita. 2008. Extended pair approximation of evolutionary game on complex networks. *Progress of theoretical physics* 119, 1 (2008), 29–38.

[23] François Mériaux, Vineeth Varma, and Samson Lasaulce. 2012. Mean field energy games in wireless networks. In *2012 Conference Record of the Forty Sixth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*. 671–675. https://doi.org/10.1109/ACSSC.2012.6489095

[24] Hisashi Ohtsuki, Christoph Hauert, Erez Lieberman, and Martin A Nowak. 2006. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* 441, 7092 (2006), 502–505. https://www.nature.com/articles/nature04605

[25] Christopher E Overton, Mark Broom, Christoforos Hadjichrysanthou, and Kieran J Sharkey. 2019. Methods for approximating stochastic evolutionary dynamics on graphs. *Journal of Theoretical Biology* 468 (2019), 45–59.

[26] Jorge Peña, Bin Wu, Jordi Arranz, and Arne Traulsen. 2016. Evolutionary games of multiplayer cooperation on graphs. *PLoS computational biology* 12, 8 (2016), e1005059.

[27] Steve Phelps, Peter McBurney, and Simon Parsons. 2010. Evolutionary mechanism design: a review. *Autonomous agents and multi-agent systems* 21, 2 (2010), 237–264.

[28] Marc Ponsen, Karl Tuyls, Michael Kaisers, and Jan Ramon. 2009. An evolutionary game-theoretic analysis of poker strategies. *Entertainment Computing* 1, 1 (2009), 39–45.

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

[30] P. Shakarian, P. Roos, and A. Johnson. 2012. A review of evolutionary graph theory with applications to game theory. *Biosystems* (2012).

[31] György Szabó and Gábor Fáth. 2007. Evolutionary games on graphs. *Physics Reports* 446, 4-6 (Jul 2007), 97–216. https://doi.org/10.1016/j.physrep.2007.04.004

[32] György Szabó and Csaba Tőke. 1998. Evolutionary prisoner's dilemma game on a square lattice. *Physical Review E* 58, 1 (1998), 69.

[33] Arne Traulsen and Christoph Hauert. 2009. Stochastic evolutionary game dynamics. *Reviews of nonlinear dynamics and complexity* 2 (2009), 25–61.

[34] Karl Tuyls and Simon Parsons. 2007. What evolutionary game theory tells us about multiagent learning. *Artificial Intelligence* 171, 7 (2007), 406–416.

[35] Sven Van Segbroeck, Jorge M Pacheco, Tom Lenaerts, and Francisco C Santos. 2012. Emergence of fairness in repeated group interactions. *Physical review letters* 108, 15 (2012), 158104.

[36] Hobart Peyton Young. 2001. *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton University Press.