Game Dynamics and Equilibrium Computation in the Population Protocol Model

Dan Alistarh
IST Austria
Austria
dan.alistarh@ist.ac.at

Krishnendu Chatterjee
IST Austria
Austria
krishnendu.chatterjee@ist.ac.at

Mehrdad Karrabi
IST Austria
Austria
mehrdad.karrabi@ist.ac.at

John Lazarsfeld
Yale University
United States of America
john.lazarsfeld@yale.edu

ABSTRACT

We initiate the study of game dynamics in the *population protocol* model: n agents each maintain a current local strategy and interact in pairs uniformly at random. Upon each interaction, the agents play a two-person game and receive a payoff from an underlying utility function, and they can subsequently update their strategies according to a fixed local algorithm. In this setting, we ask how the distribution over agent strategies evolves over a sequence of interactions, and we introduce a new *distributional equilibrium* concept to quantify the quality of such distributions. As an initial example, we study a class of *repeated prisoner's dilemma* games, and we consider a family of simple local update algorithms that yield non-trivial dynamics over the distribution of agent strategies. We show that these dynamics are related to a new class of *high-dimensional Ehrenfest random walks*, and we derive exact characterizations of their stationary distributions, bounds on their mixing times, and prove their convergence to approximate distributional equilibria. Our results highlight trade-offs between the local state space of each agent, and the convergence rate and approximation factor of the underlying dynamics. Our approach opens the door towards the further characterization of equilibrium computation for other classes of games and dynamics in the population setting.

CCS CONCEPTS

• **Theory of computation** → *Solution concepts in game theory*; *Distributed algorithms.*

KEYWORDS

Multi-agent equilibrium computation, population protocols

ACM Reference Format:
Dan Alistarh, Krishnendu Chatterjee, Mehrdad Karrabi, and John Lazarsfeld. 2024. Game Dynamics and Equilibrium Computation in the Population Protocol Model. In *ACM Symposium on Principles of Distributed Computing (PODC '24), June 17–21, 2024, Nantes, France.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3662158.3662768

PODC '24, June 17–21, 2024, Nantes, France

ACM ISBN 979-8-4007-0668-4/24/06.
https://doi.org/10.1145/3662158.3662768

1 INTRODUCTION

The emergence of complex global behavior from the interactions of simple, computationally-limited agents is a key topic of interest in distributed computing. A standard setting is the *population protocol* model, in which a set of n agents, modeled as simple, anonymous state machines, interact randomly in pairs with the goal of joint computation over the system's state. Since its introduction by Angluin et al. [9], this model has been used to characterize the evolution of several families of dynamics for solving fundamental tasks such as *majority* [2, 5, 19, 36, 38, 45, 70] and *leader election* [4, 20, 37, 43, 44]. A key feature of this model is allowing to characterize fine-grained notions of protocol convergence with respect to population size, total number of pairwise interactions (time), and available per-agent memory (space), leading to interesting trade-offs between the space and time complexity of simple local dynamics and their complex global convergence behavior [1, 20, 36, 37].

In this paper, we study population protocol dynamics in a model where, upon each randomly scheduled pairwise interaction, the two agents play a game, receive a payoff, and may subsequently update their strategies. Various multi-agent game settings have been studied over the past decades, notably in they game theory and evolutionary games literature, leading to work on proving fast convergence guarantees, hardness results, and different notions of global equilibrium for the system [26, 28, 30, 31, 47, 51, 64, 66, 73]. Yet, most prior multi-agent game settings assume a sequence of synchronous rounds, where the utility received by each individual agent depends on the actions of *all other agents in the population* in the most recent round [24, 26]. In many real-world settings, interactions are pair-wise, so an agent's utility may be *local* and depend only on its most recent interaction with a single other agent [23]. These types of settings are well-captured by the population protocol model, where agents interact randomly in pairs. However, studying the time and space complexity of computing a global equilibrium in this model has yet to be considered.

We initiate such a study of equilibrium computation in population protocols: we introduce a new *distributional equilibrium (DE)* concept that captures a notion of stability encoded by the distribution of individual agent's pure strategies in the population. For a class of *repeated prisoner's dilemma* games, we then present a simple family of local update dynamics that we prove converges to a stationary distribution corresponding to an approximate DE. In particular, these convergence results quantify the non-trivial time,

space, and approximation tradeoffs that can arise in this multi-agent equilibrium computation setting. For concreteness, we begin by introducing the general problem setting in more detail.

1.1 Problem Setting

1.1.1 Multi-Agent Distributional Games. We consider populations of n agents, each of which maintains a (pure) strategy S from a common finite set $\mathcal{S}$. At each time step t, two agents a and b are sampled from the population uniformly at random to interact and play a two-person game. Letting $S_1 \in \mathcal{S}$ and $S_2 \in \mathcal{S}$ denote the current strategies of a and b, respectively, agent a receives a payoff of $u_1(S_1, S_2)$, and agent b receives a payoff of $u_2(S_1, S_2)$, where u_1 and u_2 are fixed utility functions for the underlying two-player game.

In this setting, a *dynamics* $\mathcal{P}$ is a local update rule applied by agents a and b (following an interaction) to determine their (possibly) new strategies, and this update rule may depend on each agent's previous strategy and the payoffs received from the interaction. Under the randomness of the agent interactions, the dynamics induces a sequence of distributions $\{\boldsymbol{\mu}^t\}$, where the coordinates of $\boldsymbol{\mu}^t \in \Delta(\mathcal{S})$ specify the *fraction of agents*[1] maintaining each strategy after the t'th total interaction.

Given the set of utility functions and a *dynamics* $\mathcal{P}$, it is natural to ask whether (a) the distributions $\{\boldsymbol{\mu}^t\}$ induced by $\mathcal{P}$ *converge* (for some appropriate notion of convergence) to some distribution $\boldsymbol{\mu} \in \Delta(\mathcal{S})$ (and how quickly?), and (b) if so, whether $\boldsymbol{\mu}$ can be characterized as some global equilibrium (with respect to the utility functions u_1 and u_2). For this, we define the following *distributional equilibrium* concept for a distribution $\boldsymbol{\mu} \in \Delta(\mathcal{S})$:

Definition 1.1. Let $\boldsymbol{\mu}$ be a distribution over $\mathcal{S}$. Then for $\epsilon > 0$, $\boldsymbol{\mu}$ is an *ϵ-approximate distributional equilibrium (DE)* if it satisfies the following:

$$\mathop{\mathbf{E}}_{S_1\sim\boldsymbol{\mu},\, S_2\sim\boldsymbol{\mu}}\left[u_1(S_1,S_2)\right] \;\geq\; \max_{S'\in\mathcal{S}} \mathop{\mathbf{E}}_{S_2\sim\boldsymbol{\mu}}\left[u_1(S',S_2)\right] - \epsilon \quad (1)$$

$$\text{and} \quad \mathop{\mathbf{E}}_{S_1\sim\boldsymbol{\mu},\, S_2\sim\boldsymbol{\mu}}\left[u_2(S_1,S_2)\right] \;\geq\; \max_{S'\in\mathcal{S}} \mathop{\mathbf{E}}_{S_1\sim\boldsymbol{\mu}}\left[u_2(S_1,S')\right] - \epsilon\,. \quad (2)$$

An ϵ-approximate DE can be viewed as an approximate (symmetric) *mixed Nash equilibrium* (i.e., from classical game theory [64]), but where the "mixed strategy" $\boldsymbol{\mu} \in \Delta(\mathcal{S})$ is *induced* by the fraction of agents in the population maintaining each pure strategy $S \in \mathcal{S}$. In this view, the quantities $\mathbf{E}_{S_1\sim\boldsymbol{\mu},\, S_2\sim\boldsymbol{\mu}}[u_1(S_1,S_2)]$ and $\mathbf{E}_{S_1\sim\boldsymbol{\mu},\, S_2\sim\boldsymbol{\mu}}[u_2(S_1,S_2)]$ are the *expected payoffs* of two agents selected uniformly at random (and whose current strategies are random variables drawn independently from $\boldsymbol{\mu}$), and thus represent the agent payoffs from the "average interaction" in the population. Then if $\boldsymbol{\mu}$ is an ϵ-approximate DE, the inequalities (1) and (2) imply that for either agent, a unilateral deviation to any strategy $S' \in \mathcal{S}$ (while the other agent is still selected uniformly at random with strategy $S \sim \boldsymbol{\mu}$) can improve its expected payoff by at most ϵ.

While the problem setting and distributional equilibrium concept described above can be stated for any general set of strategies $\mathcal{S}$ and utilities u_1 and u_2, as a starting point, we investigate this concept for the special class of *repeated prisoner's dilemma* games, and in particular, the subclass of *repeated donation games*. Repeated games have been studied for decades in a variety of settings [69], leading to both well-known "folk theorems" [68], as well as to the study of evolutionary game dynamics for simple strategies [66]. Moreover, repeated prisoner's dilemma is among the most classic repeated games, and it is frequently studied to model the evolution of cooperation in populations [13, 71]. Thus to initiate the study of equilibrium computation in this population setting, we restrict our focus to this class of games, which we proceed to introduce in more detail:

1.1.2 Repeated Prisoner Dilemma Games and Population Structure. We consider *repeated donation (RD) games* (an important subclass of repeated prisoner's dilemma games) and populations where agents have one of three strategy types. In particular, we introduce the game's reward structure and strategy types as follows:

- **Reward Structure**: In RD games, two players begin by playing a single round of *prisoner's dilemma* (PD). At the end of each round, an *additional* round may be played with (independent) probability δ; otherwise, the game terminates. We call δ the *continuation* or *restart probability*. In a single round of PD, each player simultaneously chooses to *cooperate* (C) or *defect* (D), and the eponymous *dilemma* is that each player's payoff-maximizing decision is to defect, despite the fact that mutual cooperation leads to a higher payoff. In particular, we consider payoffs whose reward structure are *donation games*, which are the most important class of PD rewards [49, 59, 75]. These rewards are captured via a *reward vector* $\boldsymbol{v} := [b-c, -c, b, 0]^\top$ over the four game states $\mathcal{A} := \{CC, CD, DC, DD\}$, which are are defined by the ordered actions of the first and second players. Here, the entries in $\boldsymbol{v}$ satisfy $b > c \geq 0$ and specify the reward of the first player. Each player's *total reward* is then the sum of its payoffs over the individual rounds of the game.
- **Strategy Types and (α, β, γ) Populations**: We assume that each of the n agents in the population belongs to one of three subpopulations that correspond to distinct RD strategy types: an α fraction have strategy type AC (Always-Cooperate), a β fraction have strategy type AD (Always-Defect), and a γ fraction have strategy type GTFT (Generous-Tit-For-Tat), where $\alpha + \beta + \gamma = 1$ (and thus we refer to such populations as (α, β, γ) populations). An agent's strategy determines its action (C or D) in each round of the repeated game and may depend on its opponent's actions from prior rounds. Specifically:
  - AC *strategy*: play C at each round.
  - AD *strategy*: play D at each round.
  - GTFT *strategy*: given a *generosity parameter* $g \in [0, 1]$, play C with initial cooperation probability $s_1 \in [0, 1]$, and D with probability (w.p.) $1 - s_1$. In round $r + 1$, play the opponent's action from round r w.p. $(1 - g)$, and play C w.p. g.

  Moreover, we assume the strategy of AC and AD agents *always remains fixed*, and thus we consider update dynamics $\mathcal{P}$ that are only followed by agents with strategy GTFT. In particular, we assume each GTFT agent maintains a strategy from the set $\mathcal{G} = \{g_1, \ldots, g_k\}$ of $k \geq 2$ generosity parameter values, each of which are defined as follows: given a *maximum generosity parameter*[2]

[1] We write $\Delta(\mathcal{S})$ to denote the probability simplex over the finite set $\mathcal{S}$.

[2] Assuming such bounds on the generosity probability in the GTFT strategy are standard in RPD settings [66].

$\widehat{g} \leq 1$, for each $i \in [k]$ (and by a slight abuse of notation), the strategy g_i is the GTFT strategy with generosity parameter $g_i := \widehat{g} \cdot (\frac{i-1}{k-1})$. Here, $\mathcal{G}$ can be viewed as a discretization of the continuous space $[0, \widehat{g}]$ into k equidistant generosity parameter values.

- **Expected Payoff Functions**: Given the single RD game reward structure and the strategy types of the (α, β, γ) population, we can define an agent's utility function in the context of the more general multi-agent distributional game setting from Section 1.1.1. Specifically, we assume that when agents a and b interact, each agent's utility is its *expected payoff* in an RD game (over the randomness in each agent's strategy and the repeated rounds). For a pair of strategies $S_1, S_2 \in \mathcal{S} := \{AC, AD, g_1, \ldots, g_k\}$, we let $f(S_1, S_2)$ denote this expected payoff for an agent with strategy S_1 against an opponent with strategy S_2. By the symmetry of the single-round RD rewards, it follows that $f(S_2, S_1)$ is the expected payoff for the agent with strategy S_2.

Discussion. We remark that the focus on the three strategy types above is two-fold: first, a classical strategy in RPD is tit-for-tat (TFT) [13], which always repeats the opponent's previous action in the next round. It can be shown that the TFT strategy leads to the emergence of cooperation under suitable parameter values [13, 66], however its main drawback is lack of robustness: even in the two-player sequential setting, in the presence of noise or errors where a cooperative action may be replaced by defection, a single error makes two TFT players alternate between C and D, and after two errors both players will choose to defect forever. The key mechanism to deal with such errors is the introduction of *generosity* [60, 62, 67, 75], which motivates the GTFT class of strategies defined above and is the focus of this paper. While the analysis of other reactive strategies in RD games extends beyond GTFT [66], it is natural to begin the study of equilibrium computation with the simpler strategy sets described above. Moreover, several recent works on population protocols have considered populations containing a subset of agents whose states always remain fixed [1, 6, 8, 18, 39]. An (α, β, γ) population can be viewed as an example of this setting, where the α and β fractions of AC and AD agents remain invariant under any update dynamics.

To that end, given the RD and (α, β, γ) population structure described here, we can now more precisely describe the distributional equilibrium concept and the main algorithmic questions of interests for this setting.

1.1.3 Distributional Equilibrium in RD games for (α, β, γ) Populations. In this setting, we consider update dynamics for agents with GTFT strategies, and we are interested in the resulting distributions $\{\boldsymbol{\mu}^t\}$ over $\mathcal{G}$ that specify the fraction of GTFT agents with strategy parameter g_i for each $i \in [k]$.

For a fixed $k \geq 2$ and some distribution $\boldsymbol{\mu} \in \Delta(\mathcal{G})$, and given the (α, β, γ) population parameters, $\boldsymbol{\mu}$ also induces the $(k+2)$-part distribution $\widehat{\boldsymbol{\mu}} \in \Delta(\mathcal{S})$ over the full strategy set $\mathcal{S}$, where $\mathcal{S} := \{AC, AD, g_1, \ldots, g_k\}$. Specifically, given $\boldsymbol{\mu}$, define:

$$\widehat{\boldsymbol{\mu}}(AC) = \alpha, \quad \widehat{\boldsymbol{\mu}}(AD) = \beta, \quad \text{and} \quad \widehat{\boldsymbol{\mu}}(i) = \gamma \cdot \boldsymbol{\mu}(i) \quad \text{for } i \in [k]. \tag{3}$$

Then the ϵ-approximate distributional equilibrium concept of Definition 1.1 extends to the (α, β, γ) population setting as follows:

Definition 1.2. In the RD setting with an (α, β, γ) population, and for $\epsilon > 0$, we call $(\boldsymbol{\pi}, \widehat{\boldsymbol{\pi}}) \in \Delta(\mathcal{G}) \times \Delta(\mathcal{S})$ an *ϵ-approximate distributional equilibrium (DE)* if it satisfies the following property:

$$\mathop{\mathbf{E}}_{g \sim \boldsymbol{\pi},\, S \sim \widehat{\boldsymbol{\pi}}} \left[f(g, S) \right] \geq \max_{g' \in \mathcal{G}} \mathop{\mathbf{E}}_{S \sim \widehat{\boldsymbol{\pi}}} \left[f(g', S) \right] - \epsilon \,. \tag{4}$$

Given the one-to-one correspondence between $\boldsymbol{\pi}$ and $\widehat{\boldsymbol{\pi}}$, if expression (4) holds, we will simply say that $\boldsymbol{\pi}$ is an ϵ-approximate distributional equilibrium.

Here, the equilibrium definition can be viewed as being restricted to agent interactions where at least one agent has a GTFT strategy (this is motivated by the fact that AD and AC agents never change their strategies). In other words, $\mathbf{E}_{g \sim \boldsymbol{\pi},\, S \sim \widehat{\boldsymbol{\pi}}}[f(g, S)]$ is the expected RD game payoff of a randomly selected GTFT agent (under the distribution $\boldsymbol{\mu}$ over $\mathcal{G}$) playing against a randomly selected opponent with any strategy (under the distribution $\widehat{\boldsymbol{\mu}}$ over $\mathcal{S}$ that is induced by $\boldsymbol{\mu}$) If $\boldsymbol{\mu}$ satisfies expression (4), then the first randomly selected GTFT agent can improve its expected payoff by no more than ϵ when unilaterally deviating to a different GTFT strategy parameter (while still playing against an opponent whose strategy is drawn from $\widehat{\boldsymbol{\mu}}$). By symmetry of the expected RD payoff functions f, expression (4) also captures this same property if the second agent in the interaction is conditioned to be a GTFT agent.

With this ϵ-approximate DE definition in hand, we can more precisely state two natural questions of interest in this setting:

- *(Q1): Is there a local dynamics $\mathcal{P}$ for updating* GTFT *agents' strategies that converges to a distribution $\boldsymbol{\mu} \in \Delta(\mathcal{G})$ that is an ϵ-approximate DE, and for what approximation parameter ϵ?*
- *(Q2): If so, how many interactions are needed for $\mathcal{P}$ to converge to $\boldsymbol{\mu}$?*

1.2 Our Contributions

We introduce a family of update dynamics for the setting above that converges to an approximate DE, and we obtain quantitative bounds on its convergence rate and approximation factor:

- **k-IGT Dynamics** (Definition 2.1): We define a family of natural *incremental generosity tuning (IGT)* dynamics for updating the generosity parameters of GTFT agents among the set $\mathcal{G} = \{g_1, \ldots, g_k\}$ for each $k \geq 2$. In each instantiation, after a GTFT agent interacts with a second agent from the population, it increments or decrements its generosity parameter to the next highest or smallest parameter value in $\mathcal{G}$, depending on the strategy type of its opponent. We show this update rule can be viewed as an *introspection dynamics with local search* from evolutionary games [66], and we abbreviate the k-th dynamics by k-IGT.
- **Stationary properties of k-IGT** (Theorem 2.7): We derive exact characterizations of the stationary distribution of the k-IGT dynamics over the set of generosity parameters $\mathcal{G}$, and we derive bounds on its mixing time:
  (i) the stationary distribution of the k-IGT dynamics is *multinomial* with parameters $\gamma \cdot n$ and $(p_1, \ldots, p_k)$, where each $p_j \propto (1/\beta - 1)^{j-1}$.

(ii) when $\beta \neq 1/2$, the mixing time of the dynamics is at most $O(kn \log n)$ total interactions, and when $\beta = 1/2$, the dependence on k changes to k^2. Moreover, the mixing time is lower bounded by at least $\Omega(kn)$ total interactions.

- **Convergence to an ϵ-approximate DE** (Theorem 2.9): We prove that under suitable population and game setting regimes (when the maximum generosity parameter $\widehat{g}$ is bounded with respect to b, c, δ and the population parameters (α, β, γ)), the *mean* of the stationary distribution $\boldsymbol{\mu}$ for the k-IGT dynamics is an ϵ-approximate DE (Definition 1.2) for $\epsilon = O(1/k)$.

Our results highlight interesting tradeoffs in the multi-agent computation of an ϵ-approximate DE in this setting: for increasing k, meaning larger generosity parameter spaces $\mathcal{G}$ and larger local state requirements for each GTFT agent, the time (number of total interactions) to converge to the stationary distribution of the k-IGT dynamics grows linearly with k, but it results in a better equilibrium approximation factor ϵ decaying at a rate of $1/k$. In general, these are the first such results that quantify the convergence properties (both in mixing time, and in equilibrium approximation) of game dynamics in multi-agent settings with random pairwise interactions.

The main technical tool to achieve these results is to relate the evolution of the k-IGT dynamics to a new family of *high-dimensional, weighted Ehrenfest random walks*, which generalize the classic two-dimensional Ehrenfest urn process [41]. The stationary and mixing results we prove for these new Ehrenfest processes may be of independent interest. Our results also leave open the more general question of designing dynamics for other classes of games, and to quantify their resulting convergence properties to approximate distributional equilibria.

Structure of Paper. The remainder of the paper is structured as follows: in Section 1.3, we discuss several lines of related work in more detail. In Section 2, we introduce some additional preliminaries and provide a technical overview of our families of dynamics and convergence results. The proofs of most results can be found in the full version of the paper [3].

1.3 Related Work

Population protocol dynamics. The population model was originally introduced by Angluin et al. [9] to model computation in populations of passively mobile agents (such as sensor networks or animal populations), and has since found several other applications, from chemical reaction networks [25, 27, 35] to computing via synthetic DNA strands [29]. On the theoretical side, an impressive amount of effort has been invested in understanding the computational power of the model, e.g. [9, 11], on analyzing fundamental dynamics such as rumor spreading and averaging [15, 46], and on the complexity of core algorithmic tasks such as majority (consensus) [2, 5, 16, 17, 19, 36, 38, 45, 70] and leader election [4, 20, 37, 43, 44] in this model. The latter direction has recently lead to tight bounds on the space and task complexity of these tasks [2, 20, 36]. In this context, our contributions are to design and analyze a novel class of repeated game dynamics that lead to interesting time and space tradeoffs.

Evolutionary game dynamics. There is a huge literature on evolutionary game dynamics, and we briefly mention some key results and the relationship to our work. The first approach is to consider evolutionary dynamics in an infinite population with the aid of differential equations (aka, the replicator dynamics) [66, 74], and the goal is to study the existence and stability of equilibrium points. The second approach is to consider evolutionary dynamics in finite populations with a specific class of strategies e.g., reactive strategies or memory-1 strategies. In these approaches, the class of strategies is uncountable and simulation results suggest which strategies successfully evolve in the simulation of evolutionary dynamics [65, 67]. The third approach is to consider evolutionary dynamics on networks but with only two strategy types (AC and AD) [7, 57]. In contrast to the present paper, none of these works focus on quantitative aspects related to the mixing time (or convergence time) to the stationary behavior.

Complexity of equilibrium computation. Additionally, in both classical and evolutionary game settings, much attention has been given to understanding the computational complexity of computing an equilibrium: for example, it is known in general that computing a Nash equilibrium in a general-sum two-player game is PPAD-hard [28, 30]. Moreover, even determining the existence of evolutionary stable strategies in two-player strategic form games is NP-hard [42], and computing the fixation probability in evolutionary games with two strategies on graphs is PSPACE-complete [52]. In light of this, much progress has also been made on designing dynamics that provably converge for either (a) special subclasses of games (such as zero-sum games [63]) or (b) to approximate or weaker equilibrium concepts (such as correlated or coarse correlated equilibria [12, 47, 72]).

Other multi-player game settings. As mentioned earlier in the introduction, there is a large body of literature on other multi-player game settings. In the most basic, multi-agent normal-form game setting [26, 64], n players simultaneously choose actions at each round, and each receives a payoff according to a reward function that depends on the actions of *all other players*. In this setting and its closely related variants, extensive work has been devoted to designing local strategies that provably converge to an equilibrium, and to determine their corresponding rates of convergence (e.g., [24, 31, 47, 51, 58]). This is in contrast to the setting of the present work, where only a *single random pair* of agents interacts at each step, and thus these results are not directly comparable.

An orthogonal line of work to ours previously investigated game theoretic aspects of population protocols [21, 22], but the focus in these works is on understanding the *computational power* of interaction rules that correspond to symmetric games. Additionally, some prior work [40] studied certain random walk processes that model the evolution of cooperation in repeated prisoner's dilemma games in populations of interacting agents. However, the focus in that work is restricted to bounding the stabilization time of the process, and not on its characterization within the framework of equilibrium computation.

2 TECHNICAL OVERVIEW

2.1 Preliminaries

Notation. We use the shorthand notation $[k] = \{1, \dots, k\}$ for any $k \geq 0$, and we define the set $\Delta_k^m := \left\{(x_1, \dots, x_k) \in \mathbb{N}^k : \sum_{j\in[k]} x_j = m\right\}$. For non-negative integers x, a, and b, we write $[x]_a^b$ to denote $[x]$ truncated to the range $[a, b]$. For readability, when a and b are clear from context, we will (by slight abuse of notation) simply write $[x]$.

Markov chains. We consider discrete-time Markov chains $\{\boldsymbol{x}^t\}$ over a discrete space Ω, with transition matrix $\boldsymbol{P} : \Omega \times \Omega \to [0, 1]$. Recall that $\boldsymbol{\pi} : \Omega \to [0, 1]$ is a *stationary distribution* of $\{\boldsymbol{x}^t\}$ if $\boldsymbol{\pi}\boldsymbol{P} = \boldsymbol{\pi}$ (where we interpret the probability mass function (PMF) of $\boldsymbol{\pi}$ as a row vector). Recall also that any distribution $\boldsymbol{\nu} : \Omega \to [0, 1]$ satisfying the *detailed balance equations* $\boldsymbol{\nu}(\boldsymbol{x})\boldsymbol{P}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{\nu}(\boldsymbol{y})\boldsymbol{P}(\boldsymbol{y}, \boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$ is a stationary distribution for the process. Starting from $\boldsymbol{x}^1 := \boldsymbol{x}$ for any $\boldsymbol{x} \in \Omega$, we let $\boldsymbol{P}^t(\boldsymbol{x})$ denote the the distribution of $\boldsymbol{x}^t$ (i.e., of the process after t steps), and we write $d(t) := \max_{\boldsymbol{x}\in\Omega} \|\boldsymbol{P}^t(\boldsymbol{x}) - \boldsymbol{\pi}\|_{\mathrm{TV}}$ to denote the *distance to stationarity* (in total variation) of the process after t steps, maximized over all initial states. Then we define the *mixing time* t_{mix} of $\{\boldsymbol{x}^t\}$ as $t_{\mathrm{mix}} := \min\{t \geq 0 \ : \ d(t) \leq 1/4\}$. We refer the reader to the text of Levin and Peres [56] for more background and preliminaries on Markov chains and mixing times.

Multinomial distributions. We recall basic facts about multinomial distributions. For $m \geq 1$, $k \geq 2$, and a sequence $(p_1, \dots, p_k)$ such that $p_1 + \cdots + p_k = 1$, a distribution $\boldsymbol{\nu}$ is *multinomial with parameters m and* $(p_1, \dots p_k)$ if the PMF of $\boldsymbol{\nu}$ is given by $\boldsymbol{\nu}(\boldsymbol{x}) = p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k} \cdot \binom{m}{x_1,\dots,x_k}$ for all $\boldsymbol{x} = (x_1, \dots, x_k) \in \Delta_k^m$, where the multinomial coefficient is defined as $\binom{m}{x_1,\dots,x_k} = \frac{m!}{x_1!\cdot x_2!\dots x_k!}$. Writing $\boldsymbol{\nu} = (\nu_1, \dots, \nu_k)$, it is known that $\mathrm{E}[\nu_j] = m \cdot p_j$ for all $j \in [k]$. When $k = 2$, then $\boldsymbol{\nu}$ is a *binomial distribution*, and we can simply say that $\boldsymbol{\nu}$ is *binomial with parameters m and* p_1.

2.2 k-IGT Dynamics

We begin by formally introducing our family of local update dynamics for the problem setting described in Section 1.1. Fixing $k \geq 2$, recall that we assume each GTFT agent maintains a generosity parameter $g \in \mathcal{G} = \{g_1, \dots, g_k\}$ at each time step. The k-IGT dynamics then follows two transition types: (a) after a GTFT agent u interacts with an AC agent or a second GTFT agent, u increases its generosity parameter to the next largest value in $\mathcal{G}$, and (b) after a GTFT agent u interacts with an AD agent, u decreases its generosity parameter to the next smallest value in $\mathcal{G}$. Defined formally:

Definition 2.1 (Incremental-Generosity-Tuning (IGT) Dynamics)**.** Consider an (α, β, γ) population and an RD game setting with maximum generosity parameter $\widehat{g}$. For any $k \geq 2$, define the set of k generosity parameters $\mathcal{G} = \{g_1, \dots, g_k\}$ where each $g_j = \widehat{g} \cdot \left(\frac{j-1}{k-1}\right)$. Randomly initialize the parameter of every GTFT agent to some $g \in \mathcal{G}$. Then the k-IGT dynamics is the population protocol that evolves for all $j \in \{1, \dots, k\}$ according to the following transitions over the strategy types of interacting nodes:

$$\begin{aligned}
&\text{(i)} && g_j \ + \ \mathsf{AC} \longrightarrow \mathrm{Inc}(g_j) \ + \ \mathsf{AC} \\
&\text{(ii)} && g_j \ + \ g_i \longrightarrow \mathrm{Inc}(g_j) \ + \ g_i \ , \qquad \text{for all } i \in [k] \\
&\text{(iii)} && g_j \ + \ \mathsf{AD} \longrightarrow \mathrm{Dec}(g_j) \ + \ \mathsf{AD} \ ,
\end{aligned}$$

where $\mathrm{Inc}(g_j) := g_{\min\{j+1,k\}}$ and $\mathrm{Dec}(g_j) := g_{\max\{j-1,1\}}$.

Figure 1 shows an example of how the parameter value of a GTFT agent is updated depending on the strategy type of its interaction partner.[3] Note that while the transition rules of the k-IGT dynamics are defined with respect to the *strategies* of the two agents, these transitions could alternatively be defined with respect to the *observed game actions*. We remark that for sufficiently large δ the resulting transition rules will be essentially the same as in Definition 2.1, as in this case each agent can infer the strategy type of its opponent with high probability. It follows that the resulting dynamics will be essentially the same as to those induced by Definition 2.1, up to some small approximation error. Thus for simplicity, the transitions of the k-IGT dynamics are defined with respect to the strategy types of the interacting agents.

Bridging k-IGT and introspection dynamics. Under mild constraints on the reward vector $\boldsymbol{v}$ and maximum generosity parameter $\widehat{g}$, we show that the transition rules of the k-IGT dynamics are *locally optimal* in the following sense: under any transition rule of Definition 2.1, the expected payoff $f(g, \mathrm{S})$ will *never decrease* had the GTFT agent used the *updated* generosity parameter value (specified by the transition rules) against its previous opponent with strategy S. In particular, this bridges the relationship between the k-IGT dynamics and the classic concept of *introspection dynamics with local search* from evolutionary games [50], where an agent explores the local neighborhood of its strategy space to adopt a new strategy that would have performed better. Formally, in Appendix B of the full version, we prove the following proposition, which relies on (i) deriving exact expressions for the expected payoffs $f(g, \cdot)$, and (ii) differentiating these payoff functions with respect to g:

Proposition 2.2. *Consider an RD game setting consisting of (a) an initial cooperation probability $s_1 \in [0, 1)$ (b) a restart probability $\delta > \frac{c}{b}$ and (c) a maximum generosity parameter $\widehat{g} < 1 - \frac{c}{\delta b}$. Then for all $g, g' \in [0, \widehat{g}]$ such that $g < g'$, the following three statements hold:*

$$\begin{aligned}
&\textit{(i)} && f(g, g'') \ < \ f(g', g'') \quad \textit{for all } g'' \in [0, \widehat{g}] \\
&\textit{(ii)} && f(g, \mathsf{AC}) \ \leq \ f(g', \mathsf{AC}) \\
&\textit{(iii)} && f(g, \mathsf{AD}) \ > \ f(g', \mathsf{AD}).
\end{aligned}$$

2.2.1 Analysis setup. Given these k-IGT dynamics, our first main result (Theorem 2.7) characterizes the stationary and mixing properties of the distribution over GTFT strategies induced by the dynamics. For this, let $m := \gamma \cdot n$ denote the number of GTFT nodes in the population, and fix $k \geq 2$. We define $\boldsymbol{z}^t := (z_1^t, \dots, z_k^t) \in \Delta_k^m$ as the *count vector* specifying the number of agents with strategy g_i after the t'th step, and we study the Markov chain $\{\boldsymbol{z}^t\}$. To this end, we begin by specifying how the transitions of the k-IGT dynamics

[3] We refer to the first agent in such interactions as the *initiator*, and in this setting we assume only that the initiator ever updates its strategy following an interaction. This type of *one-way* protocol is a standard modeling assumption in the population protocol literature, e.g. [9–11, 16, 17].

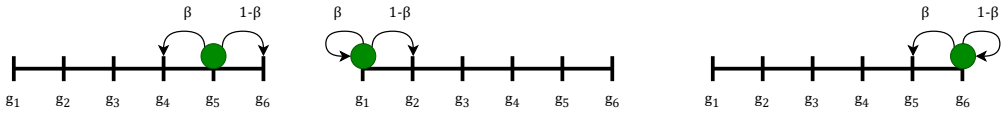

Figure 1: When $k = 6$, three examples showing how the parameter value of a GTFT agent is probabilistically updated under the k-IGT dynamics. Note that conditioned on a GTFT agent u being sampled as the first agent, u increments its parameter value with probability $(1 - \beta)$ and decrements its parameter value with probability β (where in both cases, the values are truncated to the range $[0, \widehat{g}]$).

map to the transitions $z^t \to z^{t+1}$: recall from Definition 2.1 that following any (non-null) interaction, exactly one GTFT agent updates its parameter. Then conditioned on an interaction at step t whose initiator has strategy g_j for some $j \in [k]$, then the coordinates of z^{t+1} can be specified by one of the following cases, depending on the strategy of the sampled interaction partner:

(a) If the second agent has strategy AC or GTFT, then for each $j \in [k]$:

$$z_i^{t+1} = \begin{cases} z_i^t - 1 & \text{if } i = j \text{ and } j < k \\ z_i^t + 1 & \text{if } i = j+1 \text{ and } j < k \\ z_i^t & \text{otherwise.} \end{cases}$$

(b) If the second agent has strategy AD, then for each coordinate $i \in [k]$:

$$z_i^{t+1} = \begin{cases} z_i^t - 1 & \text{if } i = j \text{ and } j > 1 \\ z_i^t + 1 & \text{if } i = j-1 \text{ and } j > 1 \\ z_i^t & \text{otherwise.} \end{cases}$$

Given that the pair of interacting agents are sampled uniformly at random at each time step, this implies that the update in (a) occurs with (unconditional) probability $(z_j^t/n) \cdot (1 - \beta)$, and the update in (b) occurs with probability $(z_j^t/n) \cdot \beta$. Then given z^t, we can summarize all transitions $z^t \to z^{t+1}$ (for $z^{t+1} \neq z^t$) that occur with non-zero probability as follows: for all $j \in [k-1]$,

$$\begin{aligned} z^{t+1} &= (z_1^t, \ldots, z_j^t - 1, z_{j+1}^t + 1, \ldots, z_k^t) \\ &\quad \text{w.p. } \tfrac{z_j^t}{m} \cdot (1 - \alpha - \beta)(1 - \beta), \text{ and} \\ z^{t+1} &= (z_1^t, \ldots, z_j^t + 1, z_{j+1}^t - 1, \ldots, z_k^t) \\ &\quad \text{w.p. } \tfrac{z_{j+1}^t}{m} \cdot (1 - \alpha - \beta)\beta \ . \end{aligned} \tag{5}$$

Observe that transition probabilities in (5) are normalized by m (using the definition $n = m/(1-\alpha-\beta)$), and that the coefficients $(1-\alpha-\beta)(1-\beta)$ and $(1-\alpha-\beta)(\beta)$ are absolute constants with respect to the coordinates of z^t. Thus we can view the process as a special case of a more general class of Markov chains $\{\boldsymbol{x}^t\}$ over Δ_k^m, whose transition probabilities (up to the absolute constant coefficients) are of the form in expression (5). We proceed to define and analyze this more general set of processes, from which characterizing the stationary and mixing properties of $\{z^t\}$ will follow.

2.3 High-Dimensional, Weighted Ehrenfest Processes

We introduce and analyze a more general class of random walks on Δ_k^m, which we refer to as *high-dimensional, weighted Ehrenfest processes*. Defined formally:

Definition 2.3 ((k, a, b, m)-Ehrenfest Process)**.** Fix $k \geq 2$, and $a, b > 0$ such that $a + b \leq 1$. Let $\{\boldsymbol{x}^t\}$ be the Markov chain on Δ_k^m with transition matrix $\boldsymbol{P} : \Delta_k^m \to \Delta_k^m$, where for all $j \in [k-1]$ and $\boldsymbol{x} = (x_1, \ldots, x_k) \in \Delta_k^m$:

$$\begin{aligned} \boldsymbol{P}\big(\boldsymbol{x}, (x_1, \ldots, x_j - 1, x_{j+1} + 1, \ldots, x_k)\big) &= p_{\boldsymbol{x}}^{j,j+1} := a \cdot \frac{x_j}{m} \\ \boldsymbol{P}\big(\boldsymbol{x}, (x_1, \ldots, x_j + 1, x_{j+1} - 1, \ldots, x_k)\big) &= p_{\boldsymbol{x}}^{j+1,j} := b \cdot \frac{x_{j+1}}{m} \\ \boldsymbol{P}(\boldsymbol{x}, \boldsymbol{x}) &= p_{\boldsymbol{x}}^{\perp} := 1 - \left(\textstyle\sum_{j=1}^{k-1} p_{\boldsymbol{x}}^{j,j+1} + p_{\boldsymbol{x}}^{j+1,j}\right), \end{aligned}$$

and $\boldsymbol{P}(\boldsymbol{x}, \boldsymbol{y}) = 0$ for all other $\boldsymbol{y} \in \Delta_k^m$. Then we call $\{\boldsymbol{x}^t\}$ the (k, a, b, m)-Ehrenfest process.

Relationship to the two-urn Ehrenfest Process. When $k = 2$ and $a = b = \frac{1}{2}$, the process reduces to the classical *Ehrenfest Urn Process* [41, 48, 76] from statistical physics. Here, m balls are distributed in two urns. At each step, an urn is sampled proportionally to its load, and with probability half, a ball from the sampled urn is placed into the other urn. The (k, a, b, m)-Ehrenfest process generalizes this original setting to a weighted, high-dimensional regime: we consider m balls distributed over a sequence of k urns, and after sampling the j'th urn proportionally to its load, a ball from urn j is placed into urn $[j+1]$ with probability a, and into urn $[j-1]$ with probability b. While the stationary and mixing behavior of two-urn process (including several weighted variants) is well-studied [32, 33, 53–55, 61], we give the first such analyses for the weighted, high-dimensional analogs from Definition 2.3.

2.3.1 Deriving the stationary distributions. We exactly characterize the stationary distributions of (k, a, b, m)-Ehrenfest processes: we show these distributions are *multinomial* with parameters $(p_1, \ldots, p_k)$, and m, where each $p_j \propto (a/b)^{j-1}$. For $k = 2$ and 3, this is obtained by viewing the process as a weighted random walk on a graph with vertex set Δ_k^m, and by solving the recurrences stemming from the detailed balance equations (these calculations are derived formally in Appendix A of the full version).

For higher dimensions (i.e., general k), we use the form of the stationary PMFs for $k = 2$ and 3 as an Ansatz for the specifying and verifying (via the detailed balance equations) the stationary PMF. Stated formally, we prove the following result, the proof of which is given in Appendix A of the full version.

Theorem 2.4. *Fix $a, b > 0$ with $a + b \leq 1$, and let $\lambda := a/b$. For any $k, m \geq 2$, let $\{\boldsymbol{x}^t\}$ be the (k, a, b, m)-Ehrenfest process, and let $\boldsymbol{\pi} : \Delta_k^m \to [0, 1]$ be its stationary distribution. Then $\boldsymbol{\pi}$ is multinomial with parameters m and $(p_1, \ldots, p_k)$, where $p_j := \frac{\lambda^{j-1}}{\sum_{i=1}^k \lambda^{i-1}}$ for all $j \in [k]$.*

2.3.2 Bounds on mixing times. Let t_{mix} and $d(t)$ denote the mixing time and distance to stationarity (as defined in Section 2.1) of the (k, a, b, m)-Ehrenfest process. We prove the following upper and lower bounds on t_{mix}:

Theorem 2.5. *Fix $a, b > 0$ with $a + b \leq 1$, and $k, m \geq 2$. Let t_{mix} be the mixing time of the (k, a, b, m)-Ehrenfest process. Then*

$$t_{mix} = \begin{cases} O\big(\min\{\frac{k}{|a-b|}, k^2\} \cdot m \log m\big) & \text{when } a \neq b \\ O\big(k^2 \cdot m \log m\big) & \text{when } a = b \end{cases}.$$

Moreover, $t_{mix} = \Omega(km)$.

Here, observe that the case distinction in the upper bound quantifies the impact of more *biased* a and b parameters in speeding up convergence to the stationary distribution, while the lower bound establishes a linear dependence on k that is uniform over all a and b. The full proof of the theorem is developed in Appendix A of the full version, but we provide a high-level proof sketch here:

Proof sketch of Theorem 2.5. To derive an upper bound on t_{mix}, we introduce the following coupling: first, let $\{X_t\}$ and $\{Y_t\}$ be random walks over $\Omega = \{1, \ldots, k\}^m$. At time t, we sample a coordinate $i \in [m]$ uniformly at random, and simultaneously increment or decrement the i'th coordinate of both X_t and Y_t (with values truncated to $[k]$) with probability a and b, respectively.

It is straightforward to see that the vector of *counts* of each value $j \in [k]$ in both X_t and Y_t evolve as (k, a, b, m)-Ehrenfest processes. Then using the standard relationship between the *coupling time* of $\{(X_t, Y_t)\}$ (the first t when $X_t = Y_t$) and mixing times [56], it suffices to probabilistically upper bound the coupling time of the joint process. We achieve this by estimating the time to coalesce each of the m coordinates of the process, and this reduces to bounding the expected absorption times of m independent (possibly) biased random walks on $\{-k, \ldots, k\}$ (which necessitates the case distinction between $a \neq b$ and $a = b$).

For the $\Omega(km)$ lower bound on t_{mix}, we use a standard diameter lower bound approach [56], which says that $t_{\text{mix}} \geq \Omega(D)$, where D is the diameter of the graph induced by the probability transition matrix of the process. For this, we can derive an straightforward estimate of $D \geq \Omega(km)$ using the structure of the transition probabilities from Definition 2.3.

Remark 2.6. The results we establish for the (k, a, b, m)-Ehrenfest process are general and may be of independent and broader interest. Moreover, our mixing time bounds of Theorem 2.5 open up several interesting questions. First, notice that our lower bounds on t_{mix} leave open at least a $O(\log m)$ gap relative to the upper bound. On the other hand, $\Omega(m \log m)$ is a known lower bound on t_{mix} for the original, unweighted $k = 2$ process [34, 56], and thus we conjecture that our upper bounds on t_{mix} for the general (k, a, b, m) process are asymptotically optimal. We leave establishing lower bounds for t_{mix} of this same order as future work. Additionally, the original two-urn process is known to exhibit a *cut-off* phenomenon [14], in which the distance to stationarity of the process sharply decays precisely at around $\frac{1}{2} \cdot m \log m$ steps [56]. Investigating this phenomenon for the general (k, a, b, m) process (and obtaining such exact cutoff constants in terms of a and b) is an interesting line of future work.

2.4 Stationary Properties of k-IGT Dynamics

By combining the arguments of Sections 2.2.1 and 2.3, we can use the stationary and mixing time analysis of the k-dimensional, weighted Ehrenfest process to analyze the evolution of the k-IGT dynamics and formally state our main results. Specifically, based on the transition probabilities in (5), for any $k \geq 2$, the sequence $\{\boldsymbol{z}_t\}$ induced by the k-IGT dynamics is exactly a (k, a, b, m)-Ehrenfest process, where $a := \gamma(1-\beta)$, $b := \gamma\beta$, and $m = \gamma n$. Thus the stationary distribution and mixing time bounds developed in Theorems 2.4 and 2.5 exactly apply to the sequence $\{\boldsymbol{z}_t\}$ induced by the k-IGT dynamics.

Before stating this main result, for convenience and readability, we first summarize in Table 1 our core notation and the components of RD games in the (α, β, γ) populations:

Symbol	Definition
b, c	donation game reward parameters
α	fraction of AC agents
δ	continuation probability
β	fraction of AD agents
s_1	initial cooperation probability
γ	fraction of GTFT agents
$\widehat{g}$	maximum generosity parameter
k	number of GTFT parameter values

Table 1: Summary of Notation in RD Games and (α, β, γ) populations

Then the stationary and convergence properties of the k-IGT dynamics are as follows:

Theorem 2.7. *Fix $k \geq 2$, and consider the sequence $\{\boldsymbol{z}^t\}$ induced by the k-IGT dynamics on an (α, β, γ) population and an RD game setting with maximum generosity parameter $\widehat{g}$. Then $\{\boldsymbol{z}^t\}$ converges to a multinomial stationary distribution $\boldsymbol{\pi}$ with parameters m and $(p_1, \ldots, p_k)$, where each $p_j = \frac{(1/\beta-1)^{(j-1)}}{\sum_{i=1}^{k}(1/\beta-1)^{(i-1)}}$ for $j \in [k]$. Moreover, letting t_{mix} denote the mixing time of $\{\boldsymbol{z}^t\}$ to $\boldsymbol{\pi}$:*

$$t_{mix} \leq \begin{cases} O\big(\min\{\frac{k}{|1-2\beta|}, k^2\} \cdot n \log n\big) & \text{when } \beta \neq \frac{1}{2} \\ O\big(k^2 \cdot n \log n\big) & \text{when } \beta = \frac{1}{2} \end{cases},$$

and $t_{mix} \geq \Omega(kn)$.

Observe that the mixing time of the process speeds up in regimes where β is bounded away from half (i.e., the number of AD agents is sufficiently small or sufficiently large). Similarly, the mean of the stationary distribution $\mathbf{E}[\boldsymbol{\pi}] = (\mathbf{E}[\pi_1], \ldots, \mathbf{E}[\pi_k])$ grows increasingly *less uniform* over $\mathcal{G}$ in this regime of β. In particular, given that $\boldsymbol{\pi}$ is a multinomial distribution, it follows that $\mathbf{E}[\pi_j] = m \cdot p_j = \gamma n \cdot p_j$ for each $j \in [k]$. Thus for $\beta < 1/2$, we expect the largest generosity parameter $g_k \in \mathcal{G}$ to have the greatest adoption among GTFT agents after t_{mix} many steps, and this mass increases as β grows smaller, and as the size k of the parameter space $\mathcal{G}$ increases.

Average stationary generosity. Given the set of generosity values $\mathcal{G} = \{g_1, \ldots, g_k\}$ and any $\boldsymbol{z} = (z_1, \ldots, z_k) \in \Delta_k^m$, we define the *average generosity value* specified by $\boldsymbol{z}$ as $\frac{1}{m}\sum_{j\in[k]} g_j \cdot z_j$. Then the stationary distribution $\boldsymbol{\pi}$ from Theorem 2.7 allows us to derive an *average stationary generosity* value $\widetilde{g}$ for the k-IGT dynamics, which we define as the average generosity value with respect to $\mathbf{E}[\boldsymbol{\pi}]$. We derive this value for all $k \geq 2$ in the following proposition:

Proposition 2.8. *Fix $k \geq 2$, and let $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_k)$ denote the stationary distribution of the k-IGT dynamics from Theorem 2.7 on an (α, β, γ) population with maximum generosity parameter $\widehat{g}$. Let $\mathcal{G} = \{g_1, \ldots, g_k\}$ be the set of generosity parameter values from Definition 2.1. Let $\widetilde{g}$ be the average stationary generosity of the dynamics, where $\widetilde{g}_k := \frac{1}{m}\sum_{j=1}^{k} g_j \cdot \mathbf{E}[\pi_j]$, and let $\lambda := (1-\beta)/\beta$. Then*

$$\widetilde{g} \;=\; \widehat{g}\cdot\left(\frac{\lambda^k}{\lambda^k-1} - \left(\frac{1}{k-1}\right)\left(\frac{\lambda}{\lambda-1}\right)\left(\frac{\lambda^{k-1}-1}{\lambda^k-1}\right)\right) \;\; \text{for } \beta \neq 1/2\,,$$

and $\widetilde{g} = \widehat{g}/2$ for $\beta = 1/2$.

Roughly speaking, Proposition 2.8 shows that for the k-IGT dynamics, $\widetilde{g} \approx \widehat{g}\cdot\left(1 - \frac{\beta}{(1-2\beta)k}\right)$ when β is bounded below $1/2$, and $\widetilde{g} \approx \widehat{g}\cdot\left(\frac{1-\beta}{(2\beta-1)k}\right)$ when β is bounded above $1/2$. Thus when the fraction of AD agents is sufficiently small, the average stationary generosity approaches the maximum generosity parameter $\widehat{g}$ at a rate of $O(1/k)$, and it approaches 0 at this same rate otherwise. This again highlights the tradeoffs between the size k of the parameter space, and the resulting levels of generosity induced by the dynamics. The proof of the proposition is given in Appendix C of the full version.

2.5 Convergence of k-IGT Dynamics to an ϵ-Approximate Distributional Equilibrium

Theorem 2.7 shows that the k-IGT converges to a stationary distribution $\boldsymbol{\pi}$ and gives a bound on its mixing time. Our next main result shows that this stationary distribution corresponds to an ϵ-approximate distributional equilibrium. Specifically, we show under suitable regimes of the RD game parameters that the (normalized) *mean* of the stationary distribution $\boldsymbol{\mu} = \frac{1}{m}\cdot\mathbf{E}[\boldsymbol{\pi}]$ of the k-IGT dynamics is an ϵ-DE for $\epsilon = O(1/k)$. Formally, we show the following:

Theorem 2.9. *Let $\boldsymbol{\pi}$ be the stationary distribution and $\boldsymbol{\mu} := \frac{1}{m}\cdot\mathbf{E}[\boldsymbol{\pi}]$ be the normalized mean stationary distribution of the k-IGT dynamics for $\lambda := \frac{1-\beta}{\beta} \geq 2$. Consider RD game settings with $s_1 \in [0,1)$, $\frac{b}{c} > 1 + \frac{\beta c}{\gamma(1-s_1)}$ and $\delta < \sqrt{1 - \frac{\beta c}{\gamma(b-c)(1-s_1)}}$, and assume $\widehat{g} < 1 - \frac{1}{\delta}\left(\frac{\beta c}{\gamma(b-c)(1-\delta)(1-s_1)} - 1\right)$. Then $\boldsymbol{\mu}$ is an ϵ-approximate distributional equilibrium (Definition 1.2) for $\epsilon = O(1/k)$.*

Combined with Theorem 2.7, the result of Theorem 2.9 gives formal, quantitative answers to the main algorithmic questions *(Q1)* and *(Q2)* stated in Section 1.1.3. Specifically, for the setting where $\lambda = (1-\beta)/\beta \geq 2$, and under suitable game parameter regimes:

- Theorem 2.9 gives a formal answer to *(Q1)* by showing that the mean stationary distribution $\boldsymbol{\mu}$ of the k-IGT dynamics is an ϵ-approximate Distributional Equilibrium, for $\epsilon = O(1/k)$.
- Theorem 2.7 gives a formal answer to *(Q2)* by showing that the k-IGT dynamics converges to its stationary distribution (and thus this ϵ-DE) within $O(kn\log n)$ total interactions.

Thus as k increases, the k-IGT dynamics converges to an approximate DE with a tighter approximation factor, but at the expense of (i) a linear increase in the convergence rate and (ii) a linear growth in the local memory required by each GTFT agent (which is needed to store the set of k generosity parameter values $\mathcal{G} = \{g_1, \ldots, g_k\}$).

Additionally, note in Theorem 2.9 that the constraints on the game setting parameters require the reward ratio b/c be sufficiently large, and that the restart probability δ and maximum generosity parameter $\widehat{g}$ both be bounded away from 1. Moreover, the result is stated only for $\lambda := (1-\beta)/\beta \geq 2$. We remark that a similar statement to Theorem 2.9 can be shown for values of λ less than 1 (e.g., $\lambda < 1/2$) that rely on assuming an upper bound on the ratio b/c, and a lower bound on δ and $\widehat{g}$, but for simplicity we focus on the case when $(1-\beta)/\beta$ is greater than 1.[4]

The full proof of Theorem 2.9 is developed in Appendix D of the full version, but we provide a high-level sketch of the proof here:

Proof Sketch of Theorem 2.9. To prove that $\boldsymbol{\mu}$ is an ϵ-approximate DE, we use the following high-level strategy: first, recall from the Definition 1.2 that $\boldsymbol{\mu}$ must satisfy

$$\mathop{\mathbf{E}}_{g\sim\boldsymbol{\mu},\, \mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g,\mathrm{S})\right] \;\geq\; \max_{g'\in\mathcal{G}} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g',\mathrm{S})\right] - \epsilon\,, \tag{6}$$

where $\widehat{\boldsymbol{\mu}} \in \Delta(\mathcal{S})$ is the distribution induced by $\boldsymbol{\mu} \in \Delta(\mathcal{G})$ in the (α,β,γ) population. Given that $\mathcal{G} = \{g_1, \ldots, g_k\}$ and $|\mathcal{G}| = k$, the condition in expression (6) can be equivalently written as

$$\mathop{\mathbf{E}}_{g\sim\boldsymbol{\mu},\, \mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g,\mathrm{S})\right] \;\geq\; \max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S})\right] - \epsilon\,. \tag{7}$$

Rearranging terms, the distribution $\boldsymbol{\mu}$ must satisfy

$$\max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S})\right] - \mathop{\mathbf{E}}_{g\sim\boldsymbol{\mu},\, \mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g,\mathrm{S})\right] \;\leq\; \epsilon\,. \tag{8}$$

For convenience, let Ψ denote the left hand side of expression (8), meaning our goal in proving Theorem 2.9 is to show that $\Psi \leq O(1/k)$. For this, we show via a first-order Taylor approximation argument that, so long as the magnitude of the second derivatives of $f(g,\mathrm{S})$ (with respect to g) are uniformly bounded by a constant $L > 0$:

$$\mathop{\mathbf{E}}_{g\sim\boldsymbol{\mu},\, \mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g,\mathrm{S})\right] \;\geq\; \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(\widetilde{g},\mathrm{S})\right] \;-\; L\cdot\mathop{\mathbf{Var}}_{g\sim\boldsymbol{\mu}}[g]\,, \tag{9}$$

[4] Observe that the result of Theorem 2.9 does not necessarily hold for $1/2 \leq \lambda \leq 2$. For such λ (multiplicatively) close to 1, the fraction β of ALLD agents in the population is close to $1/2$. Thus in this regime of λ (and depending on the game parameter settings), the mean of the stationary distribution of the corresponding Ehrenfest random walk may be far from the $g' \in \mathcal{G}$ that maximizes $\mathbf{E}_{\mathrm{S}\sim\boldsymbol{\mu}}[f(g',\mathrm{S})]$, and therefore the convergence to a DE may not hold. In this sense, the result of Theorem 2.9 requires enough "signal" from λ (i.e., needing this ratio to be bounded away from 1) in order to hold.

where $\widetilde{g} := \mathbf{E}_{g\sim\boldsymbol{\mu}}[g]$ is the average stationary generosity value from Proposition 2.8. Then it follows that we can further write

$$\begin{aligned}\Psi &:= \max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S})\right] - \mathop{\mathbf{E}}_{g\sim\boldsymbol{\mu},\,\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g,\mathrm{S})\right] \\ &\le \max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S})\right] - \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(\widetilde{g},\mathrm{S})\right] + L\cdot\mathop{\mathbf{Var}}_{g\sim\boldsymbol{\mu}}[g] \\ &= \max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S}) - f(\widetilde{g},\mathrm{S})\right] + L\cdot\mathop{\mathbf{Var}}_{g\sim\boldsymbol{\mu}}[g]\ . \qquad (10)\end{aligned}$$

Under the parameter assumptions in the statement of the theorem, we can then bound the two terms in expression (10) separately as follows:

$$\text{Propositions D.2 and D.3}: \quad L\cdot\mathop{\mathbf{Var}}_{g\sim\boldsymbol{\mu}}[g] = O(1/k^2)$$

$$\text{Proposition D.4}: \quad \max_{i\in[k]} \mathop{\mathbf{E}}_{\mathrm{S}\sim\widehat{\boldsymbol{\mu}}}\left[f(g_i,\mathrm{S}) - f(\widetilde{g},\mathrm{S})\right] = O(1/k)\ .$$

The proof of the theorem then follows by combining these two results (i.e., into expression (10)), and we develop the full proof in Appendix D of the full version.

3 DISCUSSION

In this work, we initiated the study of game dynamics and equilibrium computation in the population protocol model. We introduced a simple family of k-IGT dynamics for a class of repeated prisoner's dilemma games, and we quantified the convergence of these dynamics to an approximate distributional equilibrium. In particular, by linking the k-IGT dynamics to a new class of high-dimensional Ehrenfest processes, we obtain convergence results that highlight the time, space, and approximation factor tradeoffs for computing such global equilibria in this setting. Our work also opens the door for several future directions: first, at a broad level, it would be interesting to study game dynamics in this population setting for other classes of games (both from classical game theory and evolutionary games). In particular, it remains open to quantify the tradeoffs involved in computing a distributional equilibrium (from Definition 1.1) in this setting more generally. At a more technical level, our work also leads to several interesting questions on (k, a, b, m)-Ehrenfest processes that are discussed in Remark 2.6, and these directions are left for future work.

ACKNOWLEDGMENTS

This work was supported in part by the ERC-2020-CoG 863818 (FoRM-SMArt) grant. We thank James Aspnes and Thomas Sauerwald for several helpful discussions on Ehrenfest random walks.

REFERENCES

[1] Dan Alistarh, James Aspnes, David Eisenstat, Rati Gelashvili, and Ronald L Rivest. 2017. Time-space trade-offs in population protocols. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2560–2579.

[2] Dan Alistarh, James Aspnes, and Rati Gelashvili. 2018. Space-optimal majority in population protocols. In *Proceedings of the 29th ACM-SIAM Symposium on Discrete Algorithms ((SODA))*. 2221–2239.

[3] Dan Alistarh, Krishnendu Chatterjee, Mehrdad Karrabi, and John Lazarsfeld. 2024. Game Dynamics and Equilibrium Computation in the Population Protocol Model. arXiv:2307.07297 [cs.DC]

[4] Dan Alistarh and Rati Gelashvili. 2015. Polylogarithmic-time leader election in population protocols. In *Automata, Languages, and Programming*. Springer, 479–491.

[5] Dan Alistarh, Rati Gelashvili, and Milan Vojnović. 2015. Fast and exact majority in population protocols. In *Proceedings of the 2015 ACM Symposium on Principles of Distributed Computing*. 47–56.

[6] Dan Alistarh, Martin Töpfer, and Przemysław Uznański. 2021. Comparison dynamics in population protocols. In *Proceedings of the 2021 ACM Symposium on Principles of Distributed Computing*. 55–65.

[7] Benjamin Allen, Gabor Lippner, Yu-Ting Chen, Babak Fotouhi, Naghmeh Momeni, Shing-Tung Yau, and Martin A Nowak. 2017. Evolutionary dynamics on any population structure. *Nature* 544, 7649 (2017), 227–230.

[8] Talley Amir, James Aspnes, and John Lazarsfeld. 2020. Approximate majority with catalytic inputs. *arXiv preprint arXiv:2009.08847* (2020).

[9] Dana Angluin, James Aspnes, Zoë Diamadi, Michael J. Fischer, and René Peralta. 2006. Computation in networks of passively mobile finite-state sensors. *Distributed Comput.* 18, 4 (2006), 235–253. https://doi.org/10.1007/s00446-005-0138-3

[10] Dana Angluin, James Aspnes, and David Eisenstat. 2008. Fast computation by population protocols with a leader. *Distributed Computing* 21, 3 (Sept. 2008), 183–199.

[11] Dana Angluin, James Aspnes, and David Eisenstat. 2008. A simple population protocol for fast robust approximate majority. *Distributed Computing* 21, 2 (2008), 87–102.

[12] Robert J Aumann. 1987. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica: Journal of the Econometric Society* (1987), 1–18.

[13] Robert Axelrod and William D Hamilton. 1981. The evolution of cooperation. *Science* 211, 4489 (1981), 1390–1396.

[14] Riddhipratim Basu, Jonathan Hermon, and Yuval Peres. 2014. Characterization of cutoff for reversible Markov chains. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 1774–1791.

[15] Luca Becchetti, Andrea Clementi, Pasin Manurangsi, Emanuele Natale, Francesco Pasquale, Prasad Raghavendra, and Luca Trevisan. 2017. Average whenever you meet: Opportunistic protocols for community detection. *arXiv preprint arXiv:1703.05045* (2017).

[16] Luca Becchetti, Andrea Clementi, and Emanuele Natale. 2020. Consensus dynamics: An overview. *ACM SIGACT News* 51, 1 (2020), 58–104.

[17] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Riccardo Silvestri. 2014. Plurality consensus in the gossip model. In *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 371–390.

[18] Petra Berenbrink, Felix Biermeier, Christopher Hahn, and Dominik Kaaser. 2021. Loosely-stabilizing phase clocks and the adaptive majority problem. *arXiv preprint arXiv:2106.13002* (2021).

[19] Petra Berenbrink, Tom Friedetzky, Peter Kling, Frederik Mallmann-Trenn, and Chris Wastell. 2016. Plurality Consensus via Shuffling: Lessons Learned from Load Balancing. *arXiv preprint arXiv:1602.01342* (2016).

[20] Petra Berenbrink, George Giakkoupis, and Peter Kling. 2020. Optimal time and space leader election in population protocols. In *Proc. 52nd Annual ACM SIGACT Symposium on Theory of Computing (STOC 2020)*. 119–129.

[21] Olivier Bournez, Jérémie Chalopin, Johanne Cohen, and Xavier Koegler. 2009. Playing with population protocols. *arXiv preprint arXiv:0906.3256* (2009).

[22] Olivier Bournez, Jérémie Chalopin, Johanne Cohen, Xavier Koegler, and Mikael Rabie. 2013. Population Protocols that Correspond to Symmetric Games. *Int. J. Unconv. Comput.* 9, 1-2 (2013), 5–36.

[23] Yann Bramoullé, Rachel Kranton, and Martin D'amours. 2014. Strategic interaction and networks. *American Economic Review* 104, 3 (2014), 898–930.

[24] Yang Cai and Constantinos Daskalakis. 2011. On minmax theorems for multiplayer games. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 217–234.

[25] Luca Cardelli, Marta Kwiatkowska, and Luca Laurenti. 2016. Programming discrete distributions with chemical reaction networks. In *Proceedings of the 22nd International Conference on DNA Computing and Molecular Programming (DNA22)*. Springer, 35–51.

[26] Nicolo Cesa-Bianchi and Gábor Lugosi. 2006. *Prediction, learning, and games*. Cambridge university press.

[27] Ho-Lin Chen, Rachel Cummings, David Doty, and David Soloveichik. 2017. Speed faults in computation by chemical reaction networks. *Distributed Computing* 30, 5 (2017), 373–390.

[28] Xi Chen, Xiaotie Deng, and Shang-Hua Teng. 2009. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM (JACM)* 56, 3 (2009), 1–57.

[29] Yuan-Jyue Chen, Neil Dalchau, Niranjan Srnivas, Andrew Phillips, Luca Cardelli, David Soloveichik, and Georg Seelig. 2013. Programmable chemical controllers made from DNA. *Nature Nanotechnology* 8, 10 (2013), 755–762.

[30] Constantinos Daskalakis, Paul W Goldberg, and Christos H Papadimitriou. 2009. The complexity of computing a Nash equilibrium. *Commun. ACM* 52, 2 (2009), 89–97.

[31] Constantinos Daskalakis and Christos H Papadimitriou. 2015. Approximate Nash equilibria in anonymous games. *Journal of Economic Theory* 156 (2015), 207–245.

[32] Holger Dette. 1994. On a generalization of the Ehrenfest urn model. *Journal of applied probability* 31, 4 (1994), 930–939.

[33] Persi Diaconis. 2011. The mathematics of mixing things up. *Journal of Statistical Physics* 144, 3 (2011), 445–458.

[34] Persi Diaconis, Ronald L. Graham, and John A. Morrison. 1990. Asymptotic analysis of a random walk on a hypercube with many dimensions. *Random Structures & Algorithms* 1, 1 (1990), 51–72.

[35] David Doty. 2014. Timing in chemical reaction networks. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '14)*. SIAM, 772–784.

[36] David Doty, Mahsa Eftekhari, Leszek Gąsieniec, Eric Severson, Przemyslaw Uznański, and Grzegorz Stachowiak. 2022. A time and space optimal stable population protocol solving exact majority. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 1044–1055.

[37] David Doty and David Soloveichik. 2018. Stable leader election in population protocols requires linear time. *Distributed Computing* 31, 4 (2018), 257–271.

[38] Moez Draief and Milan Vojnovic. 2012. Convergence speed of binary interval consensus. *SIAM Journal on Control and Optimization* 50, 3 (2012), 1087–1109.

[39] Bartłomiej Dudek and Adrian Kosowski. 2018. Universal protocols for information dissemination using emergent signals. In *STOC*. 87–99. https://doi.org/10.1145/3188745.3188818

[40] Martin Dyer, Leslie Ann Goldberg, Catherine Greenhill, Gabriel Istrate, and Mark Jerrum. 2002. Convergence of the iterated prisoner's dilemma game. *Combinatorics, Probability and Computing* 11, 2 (2002), 135–147.

[41] Paul Ehrenfest and Tatjana Ehrenfest-Afanassjewa. 1907. *Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem*. Hirzel.

[42] Kousha Etessami and Andreas Lochbihler. 2008. The computational complexity of evolutionarily stable strategies. *International Journal of Game Theory* 37 (2008), 93–113.

[43] Michael Fischer and Hong Jiang. 2006. Self-stabilizing leader election in networks of finite-state anonymous agents. In *Principles of Distributed Systems*. Springer, 395–409.

[44] Leszek Gąsieniec and Grzegorz Stachowiak. 2018. Fast Space Optimal Leader Election in Population Protocols. In *Proceedings of the 29th ACM-SIAM Symposium on Discrete Algorithms ((SODA))*. 2653–2667.

[45] Mohsen Ghaffari and Merav Parter. 2016. A polylogarithmic gossip algorithm for plurality consensus. In *Proceedings of the 35th ACM Symposium on Principles of Distributed Computing ((PODC))*. 117–126.

[46] George Giakkoupis, Yasamin Nazari, and Philipp Woelfel. 2016. How asynchrony affects rumor spreading time. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*. 185–194.

[47] Noah Golowich, Sarath Pattathil, and Constantinos Daskalakis. 2020. Tight last-iterate convergence rates for no-regret learning in multi-player games. *Advances in neural information processing systems* 33 (2020), 20766–20778.

[48] Geoffrey Grimmett and David Stirzaker. 2020. *Probability and random processes*. Oxford university press.

[49] Christian Hilbe, Martin A Nowak, and Karl Sigmund. 2013. Evolution of extortion in iterated prisoner's dilemma games. *Proceedings of the National Academy of Sciences* 110, 17 (2013), 6913–6918.

[50] Josef Hofbauer and Karl Sigmund. 1998. *Evolutionary games and population dynamics*. Cambridge university press.

[51] Shuyue Hu, Harold Soh, and Georgios Piliouras. 2023. The Best of Both Worlds in Network Population Games: Reaching Consensus and Convergence to Equilibrium. In *Thirty-seventh Conference on Neural Information Processing Systems*.

[52] Rasmus Ibsen-Jensen, Krishnendu Chatterjee, and Martin A Nowak. 2015. Computational complexity of ecological and evolutionary spatial dynamics. *Proceedings of the National Academy of Sciences* 112, 51 (2015), 15636–15641.

[53] Mark Kac. 1947. Random walk and the theory of Brownian motion. *The American Mathematical Monthly* 54, 7P1 (1947), 369–391.

[54] Samuel Karlin and James McGregor. 1965. Ehrenfest urn models. *Journal of Applied Probability* 2, 2 (1965), 352–376.

[55] Olaf Krafft and Martin Schaefer. 1993. Mean passage times for tridiagonal transition matrices and a two-parameter Ehrenfest urn model. *Journal of Applied Probability* 30, 4 (1993), 964–970.

[56] David A Levin and Yuval Peres. 2017. *Markov chains and mixing times*. Vol. 107. American Mathematical Soc.

[57] Erez Lieberman, Christoph Hauert, and Martin A Nowak. 2005. Evolutionary dynamics on graphs. *Nature* 433, 7023 (2005), 312–316.

[58] Tianyi Lin, Zhengyuan Zhou, Panayotis Mertikopoulos, and Michael Jordan. 2020. Finite-time last-iterate convergence for multi-agent learning in games. In *International Conference on Machine Learning*. PMLR, 6161–6171.

[59] James AR Marshall. 2009. The donation game with roles played between relatives. *Journal of Theoretical Biology* 260, 3 (2009), 386–391.

[60] Luis A Martinez-Vaquero, Jose A Cuesta, and Angel Sanchez. 2012. Generosity pays in the presence of direct reciprocity: A comprehensive study of 2× 2 repeated games. *PLoS One* 7, 4 (2012), e35135.

[61] Michael Mitzenmacher and Eli Upfal. 2017. *Probability and computing: Randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press.

[62] Per Molander. 1985. The optimal level of generosity in a selfish, uncertain environment. *Journal of Conflict Resolution* 29, 4 (1985), 611–618.

[63] John Nash. 1951. Non-cooperative games. *Annals of mathematics* (1951), 286–295.

[64] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani. 2007. Algorithmic game theory, 2007. *Book available for free online* (2007).

[65] Martin Nowak and Karl Sigmund. 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature* 364, 6432 (1993), 56–58.

[66] Martin A Nowak. 2006. *Evolutionary dynamics: exploring the equations of life*. Harvard university press.

[67] Martin A Nowak and Karl Sigmund. 1992. Tit for tat in heterogeneous populations. *Nature* 355, 6357 (1992), 250–253.

[68] Martin J Osborne and Ariel Rubinstein. 1994. *A course in game theory*. MIT press.

[69] Guillermo Owen. 2013. *Game theory*. Emerald Group Publishing.

[70] Etienne Perron, Dinkar Vasudevan, and Milan Vojnovic. 2009. Using three states for binary consensus on complete graphs. In *INFOCOM 2009, IEEE*. IEEE, 2527–2535.

[71] Axelrod Robert. 1984. The evolution of cooperation. *Basic Books* (1984).

[72] Tim Roughgarden. 2016. *Twenty lectures on algorithmic game theory*. Cambridge University Press.

[73] William H Sandholm. 2010. *Population games and evolutionary dynamics*. MIT press.

[74] John Maynard Smith. 1982. *Evolution and the Theory of Games*. Cambridge University Press.

[75] Alexander J Stewart and Joshua B Plotkin. 2013. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences* 110, 38 (2013), 15348–15353.

[76] Lajos Takács. 1979. On an urn problem of Paul and Tatiana Ehrenfest. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 86. Cambridge University Press, 127–130.