Summary of scientific achievements

1 Name and surname: Piotr Więcek

2 Scientific degrees:

- 2000, M.Sc. in Mathematics, Institute of Mathematics, Faculty of Fundamental Problems of Technology, Wrocław University of Technology. The title of the dissertation: *Strategic Market Games.* Supervisor: Prof. Dr. Andrzej Nowak
- 2004, Ph.D. in Mathematics, Institute of Mathematics, Faculty of Fundamental Problems of Technology, Wrocław University of Technology. The title of the dissertation: *Simple Equilibria in Dynamic Strategic Market Games.* Supervisor: Prof. Dr. Tadeusz Radzik

3 Information on previous employment in scientific institutions:

- 2004–2006 *assistant* at Institute of Mathematics, Wrocław University of Technology
- 2006*–2007 *assistant* at Institute of Mathematics and Computer Science, Wrocław University of Technology
- 2007–2014 *assistant professor* at Institute of Mathematics and Computer Science, Wrocław University of Technology
- 2014*–2015 *assistant professor* at Department of Mathematics, Faculty of Fundamental Problems of Technology, Wrocław University of Technology
- 2015*–2016 *assistant professor* at Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology
- 2016*–today *assistant professor* at Department of Applied Mathematics, Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology

4 The indication of the scientific achievement.

(a) The title of the scientific achievement:
Discrete-time mean-field games with undiscounted rewards – theory and applications

(b) The list of papers constituting the scientific achievement:

[H1] P. Więcek, E. Altman, Y. Hayel, *Stochastic State Dependent Population Games in Wireless Communication.* IEEE Transactions on Automatic Control, vol. 56, no. 3 (2011), 492–505

[H2] P. Więcek, E. Altman, *Stationary Anonymous Sequential Games with Undiscounted Rewards.* Journal of Optimization Theory and Applications, 166(2) (2015), 686–710

[H3] P. Więcek, E. Altman, A. Ghosh, *Mean-field Game Approach to Admission Control of an $M/M/\infty$ Queue with Shared Service Cost.* Dynamic Games and Applications, 6 (2016), 538–566

*The changes were caused by the organizational changes at the University, i.e. the transformation of the Institute of Mathematics into the Institute of Mathematics and Computer Science, and then into the Department of Mathematics at the Faculty of Fundamental Problems of Technology, finally the foundation of the Faculty of Pure and Applied Mathematics and the Department of Applied Mathematics within this faculty.

[H4] P. Więcek, *Total Reward Semi-Markov Mean-Field Games with Complementarity Properties.* Dynamic Games and Applications 7(3) (2017), 507–529

[H5] P. Więcek, *Discrete-Time Ergodic Mean-Field Games with Average Reward on Compact Spaces.* Dynamic Games and Applications (2019), https://doi.org/10.1007/s13235-019-00296-1

(c) A discussion of the above-mentioned papers and the obtained results, together with a discussion of their possible use:

4.1 Introduction

Games with a continuum of players have been introduced to game theory by Wardrop [64] and Schmeidler [55] to approximate situations where the number of players is so large, that the impact of a single player becomes insignificant. Similar dynamic game models have appeared in literature in a paper by Jovanovic and Rosenthal [34] and its generalizations [9, 10] (discrete-time games) on one hand, and articles of Lasry and Lions [38] and Huang, Caines and Malhamé [33] (differential games) on the other. In both cases the authors considered games with a specific structure, where each of infinitely many identical players controls a private process of his own states, with each decision that he makes being a function of his own state and the distribution of the states of the other players only. It allowed to reduce the problem of finding an equilibrium in an n-person game, which for a large number of players becomes untractable, to a much simpler single-agent control problem. Differential games with a continuum of players, known as mean-field games, have been studied extensively over the last decade (see books [8, 14] or a review [26] for more details). Two basic types of problems considered in the literature were the existence of solutions to games of this type and the quality of approximation of solutions of n-person games by the solutions obtained for respective mean-field games. Similar questions were asked in the case of discrete-time games (usually called *anonymous sequential games* or *discrete-time mean-field games*). In [34, 9, 10, 16, 60, 1, 19, 54] the existence of equilibrium in discounted games of this type has been proved. In [28, 29, 32, 53, 54] the conditions for which the equilibria in games with infinitely many players are approximate equilibria in games with a large finite number of players were discussed. Also in this case, only the discounted rewards were considered. It is worth mentioning here that the lack of theory for discrete-time mean-field games with other payoff criteria resulted in the absence of applications of games of this type outside economics (where only discounted rewards are used), even though many other fields of science regularly make use of game-theoretic tools (notably, in engineering dynamic games without discounting have been used in many applications, see [7, 39]).

Three out of five papers constituting my scientific achievement (articles [H1], [H2] and [H5]) deal with discrete-time mean-field games. The main aim of my work on games of this type, was to generalize the existing results to two types of undiscounted rewards: long-time average reward and total reward, computed from the moment the player joins the game (his "birth") to his disappearance from the game (his "death"). The results concerning the existence of equilibria in such games together with some further ones about the quality of approximation of n-person stochastic games by their mean-field counterparts, published in [H2] and [H5], will be presented in section 4.2. An example of application of games of this type in wireless telecommunications (article [H1]) will be given in section 4.4.1.

The remaining results which constitute my scientific achievement (papers [H3] and [H4]) concern mean-field game models linking some features of continuous-time and dicrete-time mean-field games. Models of this type were first studied in the seminal paper by Gomes, Mohr and Souza [25] and later developed in [6, 15, 21]. In this type of games, the moments when the decisions are made are discrete, but follow separate controlled continuous time Markov chains, each controlled by a different player. As a result, these moments are discrete for each of the players, but the global state is following an ordinary differential equation. My results about the games of this type are both theoretic and applied. The theorems about the existence of equilibrium in games of this type and the conditions under which the solutions of the mean-field games are approximate solutions of similar games with a large finite number of players, published in [H4], are presented in section 4.3. The application of such games in modeling of an M/M/∞ queue with service cost divided among the users (article [H3]) is discussed in section 4.4.2.

4.2 Discrete-time mean-field games (papers [H2,H5])

4.2.1 The model

A discrete-time mean-field game is described by the following objects:

- The game is played by an infinite number (continuum) of players. Each player has a *private state* s_t^α (where $\alpha \in [0,1]$ denotes the index of the player, while $t \in \{1,2,\ldots\}$ is the stage of the game), changing over time. The set of individual states S is the same for each player, and it does not change over time. We assume that it is a nonempty compact metric space.
- A probability distribution μ_t over Borel sets[1] of S is called a *global state* of the game at stage t. It describes the proportion of the population which is in each of the individual states at time t. We assume that at every stage of the game, each player knows both his private state and the global state, and that his knowledge about individual states of his opponents is limited to the global state.
- The set of *actions* available to any player in state (s,μ) is given by $\mathcal{A}(s,\mu)$, where $\mathcal{A} : S\times\Delta(S) \to A$ is a non-empty valued correspondence and its set of values A is a compact metric space.

  At any stage of the game $\tau_t \in \Delta(S\times A)$ denotes the global distribution of the state-action pairs among the players
- Individual's α *immediate reward* is given by a bounded measurable function $r : S\times A\times\Delta(S\times A) \to \mathbb{R}$. $r(s_t^\alpha, a_t^\alpha, \tau_t)$ gives the reward of a player at any stage of the game when his private state is s_t^α, his action is a_t^α and the distribution of state-action pairs among the entire player population is τ_t.
- The sequence of the private states of player α, $(s_0^\alpha, s_1^\alpha,\ldots)$ is a Markov chain whose transitions are defined with a transition kernel Q as follows:

  $$P\{s_{t+1}^\alpha \in B | s_t^\alpha\} = Q(B|s_t^\alpha, a_t^\alpha, \tau_t) \quad \text{for } B \in \mathcal{B}(S).$$

  We assume that $Q : S\times A\times\Delta(S\times A) \to \Delta(S)$ is the same for each player and that $Q(B|\cdot,\cdot,\tau)$ is product-Borel measurable for any $B\in\mathcal{B}(S)$ and any $\tau \in \Delta(S\times A)$.

  For any $t\in\{0,1,\ldots\}$, the global state at time $t+1$ is given by the aggregation of individual transitions of the players,

  $$\mu_{t+1} = \Phi(\cdot|\tau^t) = \int_{S\times A} Q(\cdot|s,a,\tau_t)\tau_t(ds\times da).$$

  The evolution of the global state is thus deterministic.

The game is played as follows: At each stage $t = 0,1\ldots$, each player chooses his action from the set of actions available to him at that moment independently from the others, basing his decision on the knowledge of his current private state and current global state of the game. Based on the actions

[1] Here and in the sequel the Borel σ-algebra on a given set X is denoted by $\mathcal{B}(X)$, while the set of probability distributions on $(X,\mathcal{B}(X))$ is denoted by $\Delta(X)$.

chosen by the players, they are given their immediate rewards and Markov chains of their private states move to the next states. After that, the new private states are aggregated into the new global state, which is handed over to the players. We assume that at each stage of the game the players make their decisions using *stationary strategies*[2] defined as follows: A function $f : S \times \Delta(S) \to \Delta(A)$, such that $f(B|\cdot, \mu)$ is Borel measurable for any $B \in \mathcal{B}(A)$ and any $\mu \in \Delta(S)$, satisfying $f(\mathcal{A}(s, \mu)|s, \mu) = 1$ for every $s \in S$ and $\mu \in \Delta(S)$ is called a *stationary strategy*. A stationary strategy is applied as follows: each time the private state of the player is s while the global state of the game is μ, the player chooses his action according to the distribution $f(\cdot|s, \mu)$. The set of all stationary strategies is denoted by $\mathcal{F}$. In some situations we allow the use of strategies being discrete distributions over the set of stationary strategies. In that case, similarly as in the case of mixed strategies applied in extensive form games, the randomization is done only once, at the beginning of the game. The set of strategies of this type will be denoted by $\mathcal{F}^*$.

The aim of each player is, as in Markov decision processes or stochastic games, the maximization of some aggregated reward based on immediate rewards received over the entire course of the game. Two such aggregations are defined below.

Suppose player α uses a stationary strategy f against a stationary strategy g applied by the others. By the Ionescu-Tulcea theorem (see Chap. 7 in [11]), for any initial private state distribution of player α, μ_0^α and any initial distribution of private states of other players μ_0, there exists a unique probability measure $\mathbb{P}^{\mu_0^\alpha, \mu_0, Q, f, g}$ defined on the set of all infinite histories of the process of private states of player α, $H = (S \times A)^\infty$ endowed with Borel σ-algebra, such that for any $B \in \mathcal{B}(S)$, $D \in \mathcal{B}(A)$ and any partial history $h_t^\alpha = (s_0^\alpha, a_0^\alpha, \ldots, s_{t-1}^\alpha, a_{t-1}^\alpha, s_t^\alpha) \in (S \times A)^t \times S =: H_t$, $t \in \mathbb{N}$,

$$\mathbb{P}^{\mu_0^\alpha, \mu_0, Q, f, g}(h \in H : s_0^\alpha \in B) = \mu_0^\alpha(B),$$

$$\mathbb{P}^{\mu_0^\alpha, \mu_0, Q, f, g}(h \in H : a_t^\alpha \in D|h_t^\alpha) = f(D|s_t^\alpha),$$

$$\mathbb{P}^{\mu_0^\alpha, \mu_0, Q, f, g}(h \in H : s_{t+1}^\alpha \in B|(h_t^\alpha, a_t^\alpha)) = Q(B|s_t^\alpha, a_t^\alpha, \tau_t),$$

with subsequent state-action distributions τ_t defined for any $E \in \mathcal{B}(S \times A)$ recursively by the formula:

$$\tau_0(E) = \int_E g(da|s)\mu_0(ds), \quad \tau_{t+1}(E) = \int_E g(da|s)\Phi(ds|\tau_t), \quad t = 1, 2, \ldots. \tag{1}$$

Definition 1 Long-time average reward *of player α using strategy $f \in \mathcal{F}$ against strategy $g \in \mathcal{F}$ applied by all the other players, when initial distributions of private states of player α and of his rivals are μ_0^α and μ_0, is defined as follows:*

$$J^\alpha(\mu_0^\alpha, \mu_0, f, g) = \liminf_{T \to \infty} \frac{1}{T+1} \mathbb{E}^{\mu_0^\alpha, \mu_0, Q, f, g} \sum_{t=0}^{T} r(s_t, a_t, \tau_t),$$

with τ_t, $t = 0, 1, \ldots$ defined by (1).

To define the total reward, we add an artificial state s^* (denoting the "death" of a player) s^* to the set S. We then assume that in state s^* only one action a^* is available (not available in any other state), independently from the global state of the game.

[2] In case of stochastic games it is well known that limiting the strategies used by the players to stationary strategies in games with either discounted or average rewards is not problematic, as the best response against stationary strategies used by the opponents is to use a stationary strategy as well. This is also true for mean-field games. In case of the total reward used in some of our results similar result is true, as it can be interpreted as a discounted reward with state-dependent discount factor. It is known that limiting the set of strategies to the stationary ones can also be justified in that case.

Definition 2 Total reward *for player α using strategy $f \in \mathcal{F}$ against strategy $g \in \mathcal{F}$ applied by all the other players, when the distribution of private states of player α when he joins the game is μ_0^α, while the distribution of private states of all the players at the start of the game is μ_0, is defined as follows:*

$$\overline{J}^\alpha(\mu_0^\alpha, \mu_0, f, g) = \mathbb{E}^{\mu_0^\alpha, \mu_0, Q, f, g} \sum_{t=0}^{\mathcal{T}^\alpha - 1} r(s_t, a_t, \tau_t),$$

where $\mathcal{T}$ is the moment of the first arrival of the process of private states of player α to s_0 while the distributions τ_t, $t = 0, 1, \ldots$ are defined with (1).

The reward defined above is interpreted as the sum of immediate rewards of a player from the moment of his "birth" (that is, the moment when his private state moved from s^* to any $s \in S$) to that of his death (return of his private state to s^*).

The solution we are looking for differs significantly from that of *Nash equilibrium* used as a default solution in the non-cooperative game theory.

Definition 3 *A stationary strategy f and a global state $\mu \in \Delta(S)$ form a* stationary mean-field equilibrium *in the discrete-time mean-field game with a long-time average reward, if for every other stationary strategy $g \in \mathcal{F}$*

$$J(\mu, \mu, f, f) \geq J(\mu, \mu, g, f)$$

and τ_n defined for $n = 0, 1, \ldots$ by (1) with $\mu_0 = \mu$ and $g = f$ satisfies $(\tau_n)_S = \mu$ for any $n \in \{0, 1, \ldots\}$.

It means that the individual optimality condition used in the definition of the Nash equilibrium is somewhat weakened. In case of stationary equilibrium it only needs to be satisfied if the global state of the game is constant in time and equal to μ.

In case of the total reward mean-field games the definition is slightly modified:

Definition 4 *A stationary strategy f and a global state $\mu \in \Delta(S)$ form a* stationary mean-field equilibrium *in a discrete-time mean-field game with* total reward*, if*

(a) for any other stationary strategy $g \in \mathcal{F}$,

$$\overline{J}(\rho, \mu, f, f) \geq \overline{J}(\rho, \mu, g, f)$$

with $\rho = Q(\cdot|s^*, a^*, \tau(f, \mu))$ and $\tau(f, \mu)(E) = \int_E f(da|s)\mu(ds)$ for $E \in \mathcal{B}(S \times A)$, and

(b) τ_n defined for $n = 0, 1, \ldots$ by (1) with $\mu_0 = \mu$ and $g = f$ satisfies $(\tau_n)_S = \mu$ for any $n \in \{0, 1, \ldots\}$.

The definition is interpreted as follows: Any player joining the game at any stage maximizes his reward from birth to death, with the assumption that the distribution of his private state at his birth is the distribution of the private states of a player moving from s^* according to the transition probability Q. It is worth mentioning here that the mean-field games with total reward defined as we did above can be treated as counterparts of overlapping genarations models known from the stochastic game literature, see e.g. [44, 3, 43].

4.2.2 The existence of a stationary mean-field equilibrium in games with average reward

The existence of a stationary mean-field equilibrium in discrete-time mean-field games with long-time average reward, similarly as in the case of n-person stochastic games with this payoff criterion, depends crucially on assuming some conditions implying asymptotic regularity of average rewards of the players. A classical example of Gillette's *Big Match* [23] shows that without them even 2-person zero-sum stochastic game may have no stationary strategy equilibrium. While the notion of stationary mean-field equilibrium is weaker than that of stationary strategy Nash equilibrium, it turns out that an average-reward game in which the sets of transient and recurrent states in the Markov chain of private states of a player using some given stationary strategy are different for different distributions of state-action pairs in the game may have no stationary mean-field equilibrium, as shown in the example below.

Example 1 (Example 3.1 in [H2]) *Let us consider an average reward discrete-time mean-field game with*

$$S = \{1,2,3\}, \quad \mathcal{A}(s,\mu) = \begin{cases} \{0,1\}, & \text{if } s = 1 \\ \{0\}, & \text{if } s \neq 1 \end{cases}.$$

As the decision is only made by players in state $s = 1$, for simplicity we will denote this only decision by a. The immediate rewards are given by

$$r(s) = 3 - s,$$

while the transition matrix of the Markov chain of private states of each player is

$$\mathbb{Q}(a,\tau) = \begin{bmatrix} 1 - \frac{a+3p^*}{4} & \frac{a}{4} & \frac{3p^*}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{p^*}{2} & 0 & 1 - \frac{p^*}{2} \end{bmatrix}, \quad \text{where } p^* = \max\{0, 1 - 4\tau_{11}\}.$$

We will show, that such a game has no stationary mean-field equilibrium.

Suppose that (f, μ^*) is such an equilibrium and let τ be the state-action distribution corresponding to μ^* and f. We will consider two cases based on the value of τ_{11}:

(a) $\tau_{11} \geq \frac{1}{4}$: Then $p^* = 0$, and so, if a player uses action 1 with probability β, the stationary state of the chain of his states when his initial state's distribution is μ^* is $\left[\frac{2(\mu_1^*+\mu_2^*)}{2+\beta}, \frac{\beta(\mu_1^*+\mu_2^*)}{2+\beta}, \mu_3^*\right]$ and his long-time average reward is

$$\frac{(4+\beta)(\mu_1^*+\mu_2^*)}{2+\beta} = \left(1 + \frac{2}{2+\beta}\right)(\mu_1^*+\mu_2^*),$$

which is a strictly decreasing function of β. Thus his best response to f is the policy which assigns probability 1 to action $a = 0$ in state 1. But if all the players use such policy, $\tau_{11} = 0$, which contradicts our assumption that it is no less than $\frac{1}{4}$.

(b) $\tau_{11} < \frac{1}{4}$: Then the stationary state of any player's chain when he uses action 1 with probability $\beta \in [0,1]$ is independent of the initial distribution of his state μ^* and equal to $\left[\frac{2}{5+\beta}, \frac{\beta}{5+\beta}, \frac{3}{5+\beta}\right]$, which gives him the average reward of

$$\frac{4+\beta}{5+\beta} = 1 - \frac{1}{5+\beta}.$$

Clearly, it is a strictly increasing function of β. Thus the best response to f is to play action $a = 1$ with probability 1, which, if applied by all the players, results in stationary state $\mu^* = \left[\frac{1}{3}, \frac{1}{6}, \frac{1}{2}\right]$ and consequently $\tau_{11} = \frac{1}{3}$, contradicting the assumption that it is less than $\frac{1}{4}$.

Thus this game cannot have a stationary equilibrium.

Below we give two sets of assumptions which guarantee the existence of a stationary mean-field equilibrium in the average-reward discrete-time mean-field game. The first one has been considered in [H2].

(A1) The sets S and A are finite.

(A2) $r(s, a, \cdot)$ is continuous for any fixed values of $s \in S$ and $a \in A$.

(A3) $Q(\cdot|s, a, \tau)$ is a continuous function of $\tau \in \Delta(S \times A)$ for any fixed $s \in S$ and $a \in A$.

(A4) The set of individual states of any player S can be partitioned into two sets S_0 and S_1 such that for any $\tau \in \Delta(S \times A)$:

  (a) any private state $s \in S_0$ is transient in the Markov chain with the transition probability defined by
$$p_{ss'} = Q(s'|s, f(s, \tau_S), \tau)$$
  for any $f \in \mathcal{F}$,

  (b) there exists a strategy $\overline{f} \in \mathcal{F}$, such that any two states $s, s' \in S_1$ communicate in the Markov chain with the transition probability defined by
$$\overline{p}_{ss'} = Q(s'|s, \overline{f}(s, \tau_S), \tau).$$

(A5) The correspondence $\mathcal{A}(s, \cdot)$ is upper semi-continuous for any fixed $s \in S$.

In the second set of assumptions we deal with games where S and A are any compact metric spaces. Since the assumptions and the results make use of two types of convergence of probability measures, we introduce the following notation: $\mu_n \Rightarrow \mu$ for weak convergence[3] and $\mu_n \to \mu$ for strong convergence[4].

(B1) r is a continuous function.

(B2) For any sequence $\{s_n, a_n, \tau_n\} \subset S \times A \times \Delta(S \times A)$ such that $s_n \to s^*$, $a_n \to a^*$ and $\tau_n \Rightarrow \tau^*$, $Q(\cdot|s_n, a_n, \tau_n) \Rightarrow Q(\cdot|s^*, a^*, \tau^*)$. Moreover, for any fixed s and any sequence $\{a_n, \tau_n\} \subset A \times \Delta(S \times A)$ such that $a_n \to a^*$ and $\tau_n \Rightarrow \tau^*$, $Q(\cdot|s, a_n, \tau_n) \to Q(\cdot|s, a^*, \tau^*)$.

(B3) There exists a constant $\gamma > 0$ and a probability measure $P \in \Delta(S)$ such that
$$Q(D|s, a, \tau) \geq \gamma P(D)$$
for every $s \in S$, $a \in A$, $\tau \in \Delta(S \times A)$ and any Borel set $D \subset S$.

(B4) The correspondence $\mathcal{A}$ is continuous[5].

The following results are true:

Theorem 1 (Theorem 3.1 in [H2]) *Any discrete-time mean-field game with long-time average reward satisfying (A1–A5) has a stationary mean-field equilibrium.*

[3]We say that a sequence of probability measures on $(X, \mathcal{D})$, μ_n weakly converges to μ, iff $\int_X v(x)\mu_n(dx) \to \int_X v(x)\mu(dx)$ for any bounded continuous function $v : X \to \mathbb{R}$. It is known that for a compact metric set X, $\Delta(X)$ endowed with weak convergence topology is compact and metrizable (see e.g. Prop. 7.22 in [11]). There are several metrics consistent with weak convergence topology. As in our considerations we sometimes relate directly to the metric defining it, we have chosen one specific metric that will be used in our considerations (see Theorem 11.3.3 in [18]):
$$\rho_X(\mu_1, \mu_2) = \sup\left\{\left|\int_X v(x)(\mu_1 - \mu_2)(dx)\right|, \|v\|_{BL} \leq 1\right\},$$
where $\mu_1, \mu_2 \in \Delta(X)$ and $\|\cdot\|_{BL}$ is the metric on the set of bounded Lipschitz continuous functions from X to $\mathbb{R}$ defined by the formula
$$\|f\|_{BL} = \|f\|_\infty + \|f\|_L \text{ with } \|f\|_L = \sup_{x \neq y} \frac{|f(x) - f(y)|}{d_X(x, y)}.$$

[4]We say that a sequence of probability measures on $(X, \mathcal{D})$, μ_n strongly converges to μ, iff $\mu_n(D) \to \mu(D)$ for any $D \in \mathcal{D}$.

[5]With the source space $\Delta(S)$ endowed with the weak convergence topology.

In the proof we consider a Markov decision process $\mathcal{M}(\tau)$ of a single player maximizing his long-time average reward under assumption, that the state-action distribution in the game does not change over time and equals τ. By assumption (A4), the optimal reward in the MDP $\mathcal{M}(\tau)$ is independent from the initial state of the process (that is, the initial private state of the player). Let us denote this optimal reward by $G(\tau)$. It is elementary to show that G is a continuous function of τ. Next, we can use G to define a correspondence Ψ on $\Delta(S \times A)$. Let:

$$B(\tau) := \left\{ \rho \in \Delta(S \times A) : \sum_{s \in S} \sum_{a \in A} \rho_{sa} r(s, a, \tau) = G(\tau) \right\},$$

$$C(\tau) = \left\{ \rho \in \Delta(S \times A) : \sum_{a \in A} \rho_{sa} = \sum_{s' \in S} \sum_{b \in A} Q(s|s', b, \tau) \rho_{s'b} \right\},$$

$$\Psi(\tau) := B(\tau) \cap C(\tau).$$

$\Delta(S \times A)$ is obviously a compact convex subset of a Hausdorff linear topological space. The correspondence Ψ has nonempty compact values (the occupation measure corresponding to any optimal strategy in the MDP $\mathcal{M}(\tau)$ is clearly an element of $\Psi(\tau)$, convexity is immediate). The graph of Ψ is closed by the continuity of G. Glicksberg's fixed point theorem (see [24]) implies that there exists a $\tau^* \in \Psi(\tau^*)$. After the disintegration of the measure τ^* we obtain a stationary strategy f^* and a global state $\mu^* \in \Delta(S)$ which form a stationary mean-field equilibrium in the game.

In the second theorem we consider the case with general compact metric state and action spaces.

Theorem 2 (Theorem 1 in [H5]) *Any discrete-time mean-field game with long-time average reward satisfying (B1–B4) has a stationary mean-field equilibrium.*

As in the case of Theorem 1, the proof is based on an application of Glicksberg's fixed point theorem for a properly defined correspondence of the argument $\tau \in \Delta(S \times A)$. Let:

$$\Theta(\tau) := \left\{ \rho \in \Delta(S \times A) : \rho_S(\cdot) = \int_{S \times A} Q(\cdot|s, a, \tau) \rho(ds \times da) \text{ and } \int_{\mathrm{Gr}(\mathcal{A}(\cdot, \tau_S))} \rho(ds \times da) = 1 \right\},$$

$$\Psi(\tau) := \left\{ \rho \in \Theta(\tau) : \int_{S \times A} r(s, a, \tau) \rho(ds \times da) \geq \int_{S \times A} r(s, a, \tau) \sigma(ds \times da) \text{ for any } \sigma \in \Theta(\tau) \right\}.$$

Similarly as in the proof of the previous theorem, we need to show that the values of Ψ are nonempty and convex, and that the graph of Ψ is closed in the weak convergence topology. Proving nonemptiness and convexity of the values is elementary, so is the closedness of the graph of Θ. The main challenge in the proof of Theorem 2 is to prove that the graph of Ψ is closed. We are doing it in several steps:

(a) We notice that for any $\varepsilon > 0$ there exists a finite set of measurable functions $\alpha_i^\mu : S \to A$, $i = 1, \ldots, K_\varepsilon^\mu$, such that for any $s \in S$, $\mu \in \Delta(S)$ the set of values of these functions at point s, $\{\alpha_i^\mu(s), i = 1, \ldots, K_\varepsilon^\mu\}$, is an ε-net of $\mathcal{A}(s, \mu)$.

(b) We then prove that for any sequence η_n of elements of $\Delta(S \times A)$ converging weakly to $\eta \in \Delta(S \times A)$ and any function[6] $f : S \to \Delta(A)$ such that for any $s \in S$, $f(s) \in \mathcal{A}(s, \eta_S)$, a sequence of strategies $f_n : S \to \Delta(S)$ can be constructed in such a way that:

  - For each n and $s \in S$, $f_n(s) \in \mathcal{A}(s, (\eta_n)_S)$.

[6] We will call it a strategy in the remainder of the proof. Formally a strategy in a discrete-time mean-field game should describe the behaviour of its user at any global state, not only at η_S.

- For each n, the strategy f_n assigns positive probability only to the graphs of the functions $\alpha_i^{(\eta_n)_S}$, $i = 1, \ldots, K_\varepsilon^\mu$, with $\varepsilon = \frac{1}{n}$.
- For any $s \in S$, $f_n(\cdot|s) \Rightarrow f(\cdot|s)$.
- The invariant distribution on S of the Markov chain of private states of a player using f_n when the state-action distribution in the game is η_n, p_{f_n,η_n} strongly converges to $p_{f,\eta}$, the invariant distribution corresponding to f and the distribution η.

(c) We next use this property to prove that the graph of Ψ is closed as follows: suppose it is not; then there exist $\tau_n, \eta_n \in \Delta(S \times A)$ such that $\eta_n \Rightarrow \eta$, $\tau_n \Rightarrow \tau$ and $\eta_n \in \Psi(\tau_n)$, but $\eta \notin \Psi(\tau)$. Since the graph of Θ is closed, $\eta \in \Theta(\tau)$, whence there exist $\sigma \in \Theta(\tau)$ and $\varepsilon > 0$ such that

$$\int_{S\times A} r(s, a, \tau)\sigma(ds \times da) > \int_{S\times A} r(s, a, \tau)\eta(ds \times da) + \varepsilon. \tag{2}$$

σ can be disintegrated into a strategy f_σ and the invariant measure $p_{f_\sigma,\tau}$ corresponding to this strategy and the state-action distribution τ. Using (b) we may approximate them by a sequence of strategies f_σ^n such that $f_\sigma^n(\cdot|s) \in \mathcal{A}(s, (\tau_n)_S)$ for $s \in S$ and a sequence of corresponding invariant measures $p_{f_\sigma^n,\tau_n}$. This however implies that the measure $\sigma_n \in \Delta(S \times A)$ defined for any $D \in \mathcal{B}(S \times A)$ by the formula $\sigma_n(D) := \int_D f_\sigma^n(da|s)p_{f_\sigma^n,\tau_n}(ds)$ is an element of $\Theta(\tau_n)$, whence

$$\int_{S\times A} r(s, a, \tau_n)\sigma_n(ds \times da) \leq \int_{S\times A} r(s, a, \tau_n)\eta_n(ds \times da).$$

Passing to the limit, we obtain the inequality contradicting (2).

After applying the Glicksberg theorem to Ψ we obtain a fixed point τ^* which can be disintegrated into a strategy f^* and a global state μ^* which form a stationary mean-field equilibrium in the game.

4.2.3 The existence of a stationary mean-field equilibrium in games with total reward

In case of the total reward games, we only considered finite sets of private states and actions. In that situation, most of the assumptions used for the average reward games were applied again. The additional assumption (A6) was necessary to guarantee that the total reward of any player is always finite:

(A6) There exists a $p_0 > 0$, such that for any global state-action distribution τ and any stationary strategy f, the probability of getting from any state $s \in S \setminus \{s^*\}$ to s^* in $|S| - 1$ steps in the Markov chain defined by the transition probability

$$p_{ss'} = Q(s'|s, f(s, \tau_S), \tau)$$

is not smaller than p_0.

Theorem 3 (Theorem 4.1 in [H2]) *Any discrete-time mean-field game with total reward satisfying (A1–A3), (A5–A6) has a stationary mean-field equilibrium.*

Similarly as in the proof of Theorem 1, we consider a Markov decision process $\overline{\mathcal{M}}(\tau)$ of a single player. In this case he maximizes his total reward under the assumption that the global state-action distribution is fixed over time and equal to τ. In addition we assume that once the state s^* is reached for the first time, the process is absorbed. The assumption (A6) guarantees that the reward function considered in this model, $J^\tau(f, \mu_0)$, is continuous in τ, initial state distribution μ_0 and the stationary strategy of the player f. Let $\overline{G}(\tau)$ denote the optimal reward in the MDP $\overline{\mathcal{M}}(\tau)$ under the assumption that the initial state distribution for the player is $Q(\cdot|s^*, a^*, \tau)$. Similarly as in the proof of Theorem 1, we are using it to define a correspondence $\overline{\Psi} : \Delta(S \times A) \to \Delta(S \times A)$ whose fixed point should correspond to an equilibrium in the game:

$$\begin{aligned}\overline{B}(\tau) : &= \left\{ \eta \in \Delta(S \times A) : \exists f^\eta \in \mathcal{F} \forall s \in S, \left(\sum_{a\in A} \eta_{sa} > 0 \Rightarrow f(s, \tau_S) = \frac{\eta_{sa}}{\sum_{a\in A}\eta_{sa}} \right) \right. \\ &\qquad \left. \text{and } \overline{J}^\tau(f^\eta, Q(\cdot|s^*, a^*, \tau)) = \overline{G}(\tau) \right\},\end{aligned}$$

$$\overline{C}(\tau) := \left\{ \eta \in \Delta(S \times A) : \sum_{a\in A} \eta_{sa} = \sum_{s'\in S}\sum_{b\in A} Q(s|s', b, \tau)\eta_{s'b} \right\},$$

$$\overline{\Psi}(\tau) := \overline{B}(\tau) \cap \overline{C}(\tau).$$

Nonemptiness of values of $\overline{\Psi}$ can be justified using similar arguments as in the proof of Theorem 1. The fact that the graph of $\overline{\Psi}$ is closed is proved using the continuity of J^τ. Finally, the elementary renewal theorem (see Theorem 3.3.4 in [52]) is applied to transfer the condition defining $\overline{B}(\tau)$ into an equivalent linear form, which immediately gives us the convexity of the values of $\overline{\Psi}$. By the Glicksberg theorem, the correspondence $\overline{\Psi}$ has a fixed point τ^*. Disintegrating τ^*, we obtain a stationary strategy f^* and a global state μ^* which form a stationary mean-field equilibrium in the total reward game.

4.2.4 The relation with the n-person games

As we have already mentioned in the introduction, one of the main questions posed in the mean-field game literature concerns the relation between the equilibria in mean-field games with approximate equilibria in their n-person counterparts for large n. Below, we present a set of results (including one example) which try to answer this question as completely as possible. To formulate them formally we need to specify, what n-person stochastic games can be seen as counterparts of mean-field games presented above.

Definition 5 *The n-person stochastic game is called* the n-person counterpart *of the discrete-time mean-field game, if it is defined by the following objects:*

- *The state space is S^n and the action space for each player is A. The set of actions available to player i in state $\overline{s} = (s_1, \ldots, s_n)$ is given by $A^i_n(\overline{s}) := \mathcal{A}\left(s_i, \frac{1}{n}\sum_{j=1}^n \delta_{s_j}\right)$.*
- *Individual immediate reward of player i, $r^i_n : S^n \times A^n \to \mathbb{R}$, $i = 1, \ldots, n$ is defined for any $\overline{s} = (s_1, \ldots, s_n)$ and $\overline{a} = (a_1, \ldots, a_n)$ by*

$$r^i_n(\overline{s}, \overline{a}) := r\left(s_i, a_i, \frac{1}{n}\sum_{j=1}^n \delta_{(s_j,a_j)}\right).$$

- *The transition probability $Q_n : S^n \times A^n \to \Delta(S^n)$ can be defined for any $\overline{s} \in S^n$ and $\overline{a} \in A^n$ by the formula (for the clarity of exposition we write it only for Borel rectangles, which obviously defines the product measure on S^n):*

$$\begin{aligned}&Q_n(B_1 \times \ldots \times B_n|\overline{s}, \overline{a}) \\ := \ &Q\left(B_1|s_1, a_1, \frac{1}{n}\sum_{j=1}^n \delta_{(s_j,a_j)}\right) \ldots Q\left(B_n|s_n, a_n, \frac{1}{n}\sum_{j=1}^n \delta_{(s_j,a_j)}\right).\end{aligned}$$

- *The set of stationary strategies of player i is denoted by $\mathcal{F}_n^i$. We shall also use the notation $\mathcal{F}_n = \mathcal{F}_n^1 \times \ldots \times \mathcal{F}_n^n$. As in the case of mean-field games, in specific situations we will allow the use of strategies being discrete probability distributions over the sets $\mathcal{F}_n^i$. In this case, we shall use the notation $\mathcal{F}_n^{i*}$ (the set of player i's strategies of this type) and $\mathcal{F}_n^*$ (the set of vectors of such strategies).*
- *As in the case of mean-field games, the functional maximized by each player is either his long-time average reward or his total reward. They are defined for any initial state $\overline{s_0} = (s_0^1, \ldots, s_0^n) \in S^n$ and any profile of stationary strategies $\overline{f} = (f_1, \ldots, f_n) \in \mathcal{F}_n$ by the formulas:*

$$J_n^i(\overline{s}_0, \overline{f}) := \liminf_{T\to\infty} \frac{1}{T+1} \mathbb{E}^{\overline{s}_0, Q_n, \overline{f}} \sum_{t=0}^{T} r_n^i(\overline{s_t}, \overline{a_t}),$$

$$\overline{J}_n^i(\overline{s}_0, \overline{f}) = \mathbb{E}^{\overline{s}_0, Q_n, \overline{f}} \sum_{t=0}^{\mathcal{T}^i - 1} r_n^i(\overline{s_t}, \overline{a_t}),$$

where $\mathcal{T}^i$ denotes the time of first return of player i to state s^*.

Our aim is to prove, that the strategies in the stationary mean-field equilibrium in a discrete-time mean-field game can be used to construct strategies which are in an approximate equilibrium in its n-person counterparts (for n large enough). Below we formalize, what kind of equilibria we shall be looking for in these stochastic games.

Definition 6 *We say that a vector of strategies $\overline{f} \in \mathcal{F}_n$ is in a* ε-Nash equilibrium in stationary strategies *in an n-person stochastic game with long-time average reward, if it satisfies the inequalities*[7]

$$J_n^i(\overline{s}, \overline{f}) \geq J_n^i(\overline{s}, [\overline{f}_{-i}, g]) - \varepsilon$$

for any $\overline{s} \in S^n$, $g \in \mathcal{F}_n^i$, and $i \in \{1, \ldots, n\}$.

If the above inequalities are only true for vectors of strategies from a specific class $\mathcal{G}_n \subset \mathcal{F}_n$ (or $\mathcal{F}_n^*$), we say that $\overline{f}$ is in an ε-Nash equilibrium in the class $\mathcal{G}_n$.

Definition 7 *We say that a vector of strategies $\overline{f} \in \mathcal{F}_n$ and the probability distribution $\mu^* \in \Delta(S)$ are in a* weak stationary ε-equilibrium *in an n-person stochastic game with total reward, if it satisfies the inequalities*

$$\mathbb{E}\overline{J}_n^i(\overline{s}, \overline{f}) \geq \mathbb{E}\overline{J}_n^i(\overline{s}, [\overline{f}_{-i}, g]) - \varepsilon$$

for any $g \in \mathcal{F}_n^i$ and $i \in \{1, \ldots, n\}$, if initial private states s^j, $j \neq i$ are random variables drawn from the probability distribution μ^* and s^i is a random variable drawn from the probability distribution $Q(\cdot|s^*, a^*, \tau_n^*)$, while τ_n^* is an empirical state-action distribution when states s_j $j \neq i$ are drawn from the distribution μ^*, $s^i = s^*$, and actions of the players are chosen according to the stationary strategies given by $\overline{f}$.

Weak stationary ε-equilibrium in n-person stochastic games with long-time average reward is defined similarly, but in this case all the initial private states s^j are random variables with distribution μ^*.

In the first two theorems we concern the case where S and A are finite. They will make use of the following additional assumptions:

(A7) $Q(\cdot|s, a, \tau) = \widetilde{Q}(\cdot|s, a)$ for each $s \in S$, $a \in A$ and $\tau \in \Delta(S \times A)$. Moreover, $\mathcal{A}(\cdot, \mu) = \widetilde{\mathcal{A}}(\cdot)$ for each $\mu \in \Delta(S)$.

[7] $[\overline{f}_{-i}, g]$ denotes the vector $\overline{f}$ with its i-th coordinate replaced by g.

(A8) For any strategy $f \in \mathcal{F}$ and any fixed $\tau \in \Delta(S \times A)$, the Markov chain defined by the transition probability

$$p_{ss'} = Q(s'|s, f(s, \tau_S), \tau)$$

is aperiodic.

Theorem 4 (Theorem 5.1 in [H2]) *Suppose (f, μ) is a stationary mean-field equilibrium in either an average reward discrete-time mean-field game satisfying (A1–A5) and (A7) or a total reward discrete-time mean-field game satisfying (A1–A3) and (A5–A7). Then for every $\varepsilon > 0$ there exists an $\overline{n}_\varepsilon \in \mathbb{N}$ such that for every $n \geq \overline{n}_\varepsilon$, $(\overline{f}, \mu)$, where $\overline{f} = (f, \dots, f)$, is a weak stationary ε-equilibrium in n-person counterpart of this game.*[8]

The proof of this theorem is based on the following observation: Since in the case of games satisfying assumption (A7), the evolution of the private state of any given player is independent from the state-action distribution in the entire population, the long-time average reward of player i using strategy $g \in \mathcal{F}$ against $f \in \mathcal{F}$ (the strategy in the stationary mean-field equilibrium) applied by all the other players in n-person counterparts of the mean-field game can be represented as the sum:

$$\sum_{s \in S} \sum_{a \in A} \sum_{\tau \in \Delta^n(S \times A)} \sigma_{sa}(g) m^n_\tau(\overline{f}_{-i}, g) r(s, a, \tau), \tag{3}$$

where $\sigma(g)$ denotes the occupation measure on $S \times A$ corresponding to the individual decision process of a player using strategy g, $\Delta^n(S \times A)$ – a set of atomic distributions on $S \times A$ with atoms being multiples of $\frac{1}{n}$, while $m^n_\tau(\overline{f}_{-i}, g)$ is the measure of frequency of different values of the global state-action distribution τ_t during the game. The measure σ_f is independent from the number of the players, while $m^n_\tau(\overline{f}_{-i}, g)$ weakly converges to the global state-action distribution in the mean-field game, corresponding to the situation when all the players apply strategy f and the global state of the game is μ. This however means that (3) converges to the reward in the mean-field game corresponding to the situation when player i uses strategy g against f of the others, while the global state is μ. This is obviously enough to obtain the thesis for the average reward game. To prove the total reward case, we use the elementary renewal theorem to rewrite the reward in the n-person counterpart of the given total reward discrete-time mean-field game for a player using stationary strategy g against f of all the others, when initial distribution of states is μ, in a form similar to (3). Then we can use the arguments from the first part of the proof to obtain the thesis.

Theorem 5 (Theorem 5.2 in [H2]) *For every $\varepsilon > 0$ there exists an $n_\varepsilon \in \mathbb{N}$ such that for every $n \geq n_\varepsilon$ the n-person counterpart of the average-reward discrete-time mean-field game satisfying (A1–A5) and (A7–A8) has a symmetric ε-Nash equilibrium $(\pi^n, \dots, \pi^n) \in \mathcal{F}^*_n$ defined as follows: if (f, μ) is an equilibrium in the mean-field game, then π^n is of the form:*

$$\pi^n(s) = \sum_l \mu^*_l \delta[f^n_l(s)], \quad \textit{where} \quad f^n_l(s) = \begin{cases} \overline{f}(s), & \textit{if } s \notin S_l, \\ f(s), & \textit{if } s \in S_l, \end{cases}$$

$\overline{f}$ is the communicating policy introduced in part (b) of the assumption (A1)[9]*, S_l are ergodic classes of the private state process of a player when he applies strategy f*[10]*, and μ^* is the probability measure on the set of these ergodic classes corresponding to measure μ over S.*

[8] Formally, f is a function from $S \times \Delta(S)$ to A, while stationary strategies in n-person counterpart of a mean-field game are functions from S^n to A. It is easy to note though, that f applied to the empirical distribution of private states of the players in the n-person game (as its second argument) is in fact a function of the vector of private states of all the players. Obviously, if we wanted to be precise, we should formally define strategies of each of the players corresponding to f and then write that the vector of such strategies in in an ε-equilibrium in the n-person counterpart of the mean-field game. We shall simplify the notation in a similar manner in other theorems as well.

[9] Here it does not depend on τ by (A7).

[10] They also do not depend on τ by (A7).

The strategy π^n defined in the above theorem is designed in such a way that the stationary distribution on the set of private states of a player using this strategy in the n-person counterpart of the mean-field game is independent from the initial private state and equal to μ. This implies that the inequalities defining the weak stationary ε-equilibrium $(\overline{f}, \mu)$ obtained in Theorem 4 are for strategies f replaced by π^n true for any initial distribution of private states of a given player. As a consequence, for any $n \geq n_\varepsilon$, the vector $(\pi^n, \ldots, \pi^n)$ is an ε-Nash equilibrium in the n-person counterpart of the mean field game.

The next results correspond to the case when S and A are compact metric. They will make use of the following additional assumptions:

(B5) $Q(\cdot|s, a, \tau) = \widetilde{Q}(\cdot|s, a)$ for any $s \in S$, $a \in A$ and $\tau \in \Delta(S \times A)$. Moreover, $\mathcal{A}(\cdot, \mu) = \widetilde{\mathcal{A}}(\cdot)$ for any $\mu \in \Delta(S)$.

(B6) For any sequence $\{s_n, a_n, \tau_n\} \subset S \times A \times \Delta(S \times A)$, such that $s_n \to s$, $a_n \to a$ and $\tau_n \Rightarrow \tau$, $Q(\cdot|s_n, a_n, \tau_n) \Rightarrow Q(\cdot|s, a, \tau)$.

The next theorem can be treated as a counterpart of Theorem 4 for the case of compact S and A.

Theorem 6 (Theorem 2 in [H5]) *Suppose that (f^*, μ^*) is a stationary mean-field equilibrium in an average-reward discrete-time mean-field game satisfying (B1) and (B3–B6). Then for any $\varepsilon > 0$ there exists an $n_\varepsilon \in \mathbb{N}$, such that for $n \geq n_\varepsilon$ the vector of strategies $\overline{f} = (f, \ldots, f)$, where $f(\cdot|s, \mu) \equiv f^*(\cdot|s)$ is an ε-Nash equilibrium in the n-person counterpart of the mean-field game.*

In the case of mean-field games where the transitions of Markov chains of private states depend on the state-action distribution of the players, an equilibrium in the discrete-time mean-field game may fail to be an approximate equilibrium in its n-person counterparts for any value of n. This is shown by the following example[11]:

Example 2 (Example 2 in [H5]) *Consider an average-reward discrete-time mean-field game with $S = \{0, 1\} = A$ defined by:*

$$Q(\cdot|s, a, \mu) = \begin{cases} (2\mu_0 - 1)\delta_0 + 2\mu_1\delta_1 & \text{if } a = 0 \text{ and } \mu_0 \geq \frac{2}{3} \\ \frac{1}{3}\delta_0 + \frac{2}{3}\delta_1 & \text{if } a = 0 \text{ and } \mu_0 < \frac{2}{3} \\ \frac{2\mu_0 + 1}{3}\delta_0 + \frac{2\mu_1}{3}\delta_1 & \text{if } a = 1 \end{cases}$$

$$r(s, a, \mu) = \begin{cases} 6s & \text{if } a = 0 \\ 1 - s & \text{if } a = 1 \end{cases}$$

It can be shown that $f^* \equiv \delta_0$ and $\mu^* = \frac{1}{3}\delta_0 + \frac{2}{3}\delta_1$ form a stationary mean-field equilibrium in this game.

Now suppose all the players in the n-person counterpart of this game use strategy f^*. It is easy to see, that in such a situation all the private states become zeros after a finite number of stages, regardless of their initial values. Note however, that the state $\overline{s} = (0, \ldots, 0)$ is an absorbing state. Hence, the average reward corresponding to the profile consisting of strategies f^* in the n-person counterpart of the mean-field game is 0. Now suppose that one of the players changes his strategy to $g \equiv \delta_1$. Then the game is still absorbed at $\overline{s} = (0, \ldots, 0)$, but the ergodic reward of the player using strategy g is 1, so the profile of f^* is not an ε-Nash equilibrium in any of the n-person counterparts of the mean-field game for any $\varepsilon < 1$.

[11] It is worth mentioning here, that it is one of only two examples of this kind available in the mean-field game literature. The other example can be found in [13]. It is significantly different from ours.

Theorem 7 (Theorem 3 in [H5]) *For any $L > 0$, define*

$$\begin{aligned}\mathcal{F}^L &= \{f \in \mathcal{F} : f \text{ is weakly continuous and for any } s \in S,\\ &\quad f(\cdot|s,\cdot) \text{ is weakly Lipschitz continuous with constant } L\}.\end{aligned}$$

Let then (f^*, μ^*) be a stationary mean-field equilibrium in an average-reward discrete-time mean-field game satisfying (B1–B4). Assume further that:

(a) The stationary strategy f defined with the formula $f(\cdot|s,\mu) = f^*(\cdot|s)$ for any $s \in S$ and $\mu \in \Delta(S)$ is an element of $\mathcal{F}$. Moreover, it is weakly Lipschitz-continuous with constant β_f as a function of s.

(b) The transition kernel Q satisfies[12] *for any $s \in S$, $a_1, a_2 \in A$ and $\tau_1, \tau_2 \in \Delta(S \times A)$*

$$\|Q(\cdot|s, a_1, \tau_1) - Q(\cdot|s, a_2, \tau_2)\}\|_v \le \beta_Q(\max\{d_A(a_1, a_2), \rho_{S\times A}(\tau_1, \tau_2)\}).$$

(c) The constants β_f, β_Q satisfy $\beta_Q(1 + \beta_f) < \frac{\gamma}{2}$.

Then for any $\varepsilon > 0$ and $L > 0$ there exists an $n_{\varepsilon,L} \in \mathbb{N}$ such that for any $n \ge n_{\varepsilon,L}$ the profile of strategies where each player uses strategy f is an ε-Nash equilibrium in the class $\mathcal{F}^L_n$ in the n-person counterpart of the mean-field game.

The proofs of Theorems 6 and 7 are based on a result by Boissard (see Corollary 2.5 in [12]), giving a possibility of estimating the difference between the values of a given continuous function of a probability distribution in two cases: when its argument is some given distribution, and when we replace it by an empirical distribution from a k-element vector of random variables with this distribution. Using this result, we are able to prove that the rewards in the n-person counterparts of the average-reward discrete-time mean-field game converge to the rewards in the mean-field game under the following assumption: The marginals on S of the invariant distributions for the transition probabilities in the n-person counterparts of the mean-field game corresponding to the case when all players but one use the same stationary strategy weakly converge to the invariant distribution of the Markov chain of private states of a player in the mean-field game when all the players use this strategy[13]. This kind of convergence is enough to prove the theses of both theorems.

In the case of Theorem 6, the above-mentioned S-marginals of invariant distributions in the n-person model and the invariant distributions in the mean-field model are identical by (B5) (as the transitions in the Markov chains of private states of any given player are independent from the private states and the actions of the others, regardless of their number). In the case of Theorem 7, the crucial part of the proof of the convergence of the S-marginals of the invariant distributions in n-person games to the invariant distributions in the mean-field game, is showing the uniqueness of the invariant measure of the process of private states of a player using strategy $g \in \mathcal{F}^L$ against $f \in \mathcal{F}^L$ of all the others in the mean-field game (in general it may depend on the initial global state-action distribution – it is that way in the game from example 2). We do it by showing that the function $M_f : \Delta(S) \to \Delta(S)$ defined with the formula

$$M_f(\mu) := p_{f,\Pi(f,\mu)}, \quad \text{where } \Pi(f,\mu)(D) := \int_D f(da|s)\mu(ds) \text{ for } D \in \mathcal{B}(S \times A)$$

[12] d_A in the equality below is the metric on A. $\|\cdot\|_v$ denotes the norm defined on the set of all finite signed measures on $(S, \mathcal{B}(S))$ by the formula

$$\|\mu\|_v = \sup_{B\in\mathcal{B}(S)} \mu(B) + |\inf_{B\in\mathcal{B}(S)} \mu(B)|.$$

[13] Obviously, we need some additional assumptions concerning the continuity of the strategies used by the players.

satisfies the assumptions of the Banach fixed point theorem. The fixed point that we obtain, μ_{ff}, is the unique invariant measure of the process of private states of any given player when everyone uses strategy f in the mean-field game. The uniqueness of the invariant measure (denoted further as μ_{gf}) in the general case (i.e. when a player uses strategy g against f applied by other players) can be easily shown by combining the case with $f = g$ with geometric ergodicity of the Markov chain of private states of the player which follows from (B3).

The next step of the proof is noticing that any subsequence of the sequence of the S-marginals of the invariant measures in the n-person counterparts of the mean-field game when a player uses strategy g against f of the others has a convergent subsequence. For a game satisfying the assumptions of Theorem 7 it can be shown that this subsequence will converge to μ_{gf}. This however implies that the entire sequence converges to μ_{gf}, which is what we wanted to prove.

4.3 Semi-Markov mean-field games (paper [H4])

4.3.1 The model

The next model we are going to present is in many ways similar to the total reward model analyzed in previous sections. For that reason the description given below concentrates on the differences between the two formalisms. A semi-Markov mean-field game with total reward is described as follows:

- The sets of private states S and actions A are finite. Both, as in the case of discrete-time mean-field games with total reward, are complemented with elements s^* and a^* meaning "death" of a player and the only action available in state s^*.
- The global state of the game at time t is denoted by X_t to make a distinction between the discrete- and continuous-time models. As in the previous models, global state is a probability measure over S. It describes the mass of the population which is at time t in each of the individual states.
- We assume that the time is continuous ($t \geq 0$), but the individual state of player α can only change at specific times $T_0^\alpha, T_1^\alpha, \ldots$, where $T_0^\alpha = 0$. The time between successive transitions $\tau_k^\alpha = T_{k+1}^\alpha - T_k^\alpha$ is random exponentially distributed with intensity $\lambda(s_{T_{k-1}^\alpha}, X_{T_k^\alpha})$. Random variables τ_k^α are for different k and α independent. We assume that λ is a positive, Lipschitz continuous function of the global state of the game.
- The sets of actions available to a player at any time are, as in other models considered here, given by an upper semi-continuous correspondence $\mathcal{A} : S \times \Delta(S) \to A$.
- The transition in the process of private states of player α at time T_{k-1}^α is according to the transition function $Q : S \times A \times \Delta(S) \to \Delta(S)$ which is a Lipschitz continuous function of the global state.
- As before, we assume that all the players use stationary strategies. The set of all stationary strategies is denoted by $\mathcal{F}$. The set of deterministic stationary strategies is denoted by $\mathcal{F}_d$.
- As private states of different players change in different moments, the evolution of the global state is described by an ordinary differential equation:

$$\dot{X}_t^s = \sum_{s' \in S} \sum_{a \in A} X_t^{s'} \lambda(s', X_t) Q(s|s', X_t, a) \widehat{f}_a(s', X_t) - X_t^s \lambda(s, X_t), \quad s \in S \tag{4}$$

  with $X_0 \equiv x_0$, the initial global state and $\widehat{f}_a(s, X) := \int_0^1 \mathbb{1}\{f^\alpha(s, X) = a\} \, d\alpha$, where $f^\alpha \in \mathcal{F}_d$

denotes the stationary strategy of player α[14]. The ODE (4) is usually referred to as the Kurtz dynamics in the literature (see e.g. Theorem 5.3 in [56]).

- The total reward of player α using strategy $f \in \mathcal{F}$ against $g \in \mathcal{F}$ applied by all the others is computed according to the formula

$$\overline{J}^{\alpha}(f, g, \mu_0) = \mathbb{E}^{x_0,\mu_0,Q,f,g} \sum_{i=0}^{i_e^{\alpha}-1} \left(\widetilde{r}(s^{\alpha}_{T_i^{\alpha}}, X_{T_i^{\alpha}}(g), a^{\alpha}_{T_i^{\alpha}}) + \int_{T_i^{\alpha}}^{T_{i+1}^{\alpha}} r(s^{\alpha}_{T_i^{\alpha}}, X_t(g), a^{\alpha}_{T_i^{\alpha}})\, dt \right) P_{f,\mu_0},$$

where $T_{i_e^{\alpha}}$ is the moment of the first return of the private state of player α to s^*, μ_0 is the distribution of private states of new-born players, $r : S \times A \times \Delta(S) \to \mathbb{R}$ is the immediate reward function, while $\widetilde{r} : S \times A \times \Delta(S) \to \mathbb{R}$ is the reward received by a player upon the change of his state. We assume both r and $\widetilde{r}$ are the same for each player and continuous in the global state of the game.

- Stationary mean-field equilibrium in the game is defined similarly as in the case of discrete-time mean-field games with total reward.

4.3.2 The results

The existence of a stationary mean-field equilibrium in games of this type has been proved under some lattice-theoretic assumptions[15]:

(C1) There exists a $p_0 > 0$ such that for any global state X and any strategy $f \in \mathcal{F}$ the probability of getting from any private state $s \in S \setminus \{s^*\}$ to s^* in $|S| - 1$ steps in the Markov chain with transition probabilities
$$p_{ss'} = Q(s'|s, f(s, X), X)$$
is not smaller than p_0.

(C2) S and A are sublattices of $\mathbb{R}$ such that $s^* = \min\{S\}$ and $a^* = \min\{A\}$. Moreover, (a) for any $s \in S$ and $X \in \Delta(S)$, $\mathcal{A}(s, X)$ is a sublattice of A, (b) $\mathcal{A}(s, X)$ is a non-decreasing function of (s, X).

(C3) $r(s, a, X)$ and $\widetilde{r}(s, a, X)$ are non-negative, non-decreasing in s and supermodular in (s, a). Moreover, they have increasing differences in (s, a) and X.

(C4) $Q(\cdot|s, a, X)$ is stochastically supermodular in (s, a) and stochastically non-decreasing in s, a and X. Moreover, it has stochastically increasing differences in (s, a) and X.

(C5) $\lambda(s, X)$ does not depend on s and is a non-increasing function of X.

Assumptions of this type have been used in the game-theoretic literature for a long time, also in the case of dynamic games (see [4, 5, 17, 31, 42, 60, 63, 1]). They describe the situation when private

[14] The results proved in [H4] are all about the existence of a stationary mean-field equilibrium in which all the players use deterministic strategies. $\widehat{f}$ is defined here only in this case – the definition would become slightly more complicated, if we did not make such an assumption.

[15] The lattice-theoretic notions used there are briefly discussed below:

We say that a set X with partial order $\preceq_X$ is a *lattice*, if for any $x_1, x_2 \in X$ the set X contains $x_1 \vee x_2 := \sup\{x_1, x_2\}$ and $x_1 \wedge x_2 := \inf\{x_1, x_2\}$. X is a *complete lattice*, if any subset X has upper and lower bounds which are elements of X.

We say that a function $f : X \to \mathbb{R}$ is *supermodular*, if $f(x_1 \vee x_2) + f(x_1 \wedge x_2) \geq f(x_1) + f(x_2)$ for $x_1, x_2 \in X$. If in addition $(Y, \preceq_Y)$ is a lattice, we say that a function $g : X \times Y \to \mathbb{R}$ has increasing differences, if for $x_1, x_2 \in X$, $x_1 \preceq_X x_2$ and $y_1, y_2 \in Y$, $y_1 \preceq_Y y_2$, $g(x_2, y_2) - g(x_1, y_2) \geq g(x_2, y_1) - g(x_1, y_1)$ holds.

It is known that $(\mathbb{R}, \leq)$ is a lattice. Another well-known example of a lattice is the set $\Delta(\mathbb{R})$ with the (first-order) stochastic dominance ordering. Each time we refer to some lattice-theoretic properties defined with the help of this ordering, we add „stochastically" to the name of the property (e.g. we write about stochastic supermodularity, stochastically increasing differences etc.)

states and actions of the players are linked in the following way: If the private state of a player is big, it is more profitable to use big actions. Similarly, if the private states of the others are big, using bigger actions becomes more profitable. It turns out, that many practical applications of dynamic games (in both economics and engineering) can be modeled by games of this type.

Theorem 8 (Theorem 1 in [H4]) *A semi-Markov mean-field game with total reward satisfying (C1–C5) has a stationary mean-field equilibrium (f^*, X^*) with $f^* \in \mathcal{F}_d$, such that f^* is non-decreasing both in the private state of the player and in the global state of the game.*

The proof of the above theorem consists of several steps:

(a) We start by showing (using some standard techniques used in theory of supermodular games, see [62], and some standard tools used in dynamic programming, see e.g. [49]), that the optimal reward of a player maximizing his total reward in the mean-field game under the assumption that the global state does not change over time and equals X, V_X^*, preserves the properties of functions r and $\widetilde{r}$ given by (C3).

(b) Let $\mathcal{B}(X,s)$ denote the set of actions maximizing the RHS of the Bellman equation for the above-mentioned maximization problem. Moreover, let $\overline{B}(X,s) := \max \mathcal{B}(X,s)$, $\underline{B}(X,s) := \min \mathcal{B}(X,s)$. Using the Topkis theorem (see Theorem 2.8.3 in [62]) it can be shown that $\overline{B}$ and $\underline{B}$ are increasing in X and (for any fixed X) of s.

(c) Next, let

$$\mathcal{F}_0 := \{f \in \mathcal{F}_d : f(s,X) \text{ is nondecreasing in } X \text{ and for any fixed } X \text{ in } s\}.$$

Further, let $\overline{X}(f,X)$ and $\underline{X}(f,X)$ be the biggest and the smallest (in the stochastic ordering) stationary distributions in the Markov chain of private states of a player using strategy $f \in \mathcal{F}_0$ when the global state of the game is constant and equal to X. We then show that $\overline{X}$ and $\underline{X}$ are non-decreasing functions of f and X.

(d) Finally, we define $\overline{\Psi} : \Delta(S) \to \Delta(S)$ and $\underline{\Psi} : \Delta(S) \to \Delta(S)$ with the formulas

$$\overline{\Psi}(X) := \overline{X}(\overline{B}, X) \quad \text{and} \quad \underline{\Psi}(X) := \underline{X}(\underline{B}, X).$$

The properties that we have proved in (b) and (c) imply that they are both nondecreasing endomorphisms defined on a complete lattice $\Delta(S)$. By the Tarski fixed point theorem (see [61]) both functions have fixed points: respectively $\overline{X}^*$ and $\underline{X}^*$. Taking either $f^* = \overline{B}$, $\mu^* = \overline{X}^*$ or $f^* = \underline{B}$, $\mu^* = \underline{X}^*$, we obtain the thesis of the theorem.

The next theorem presents a simple distributed learning procedure. It allows the players to learn one of the strategies in stationary mean-field equilibrium. The proof is rather technical, so we have decided to skip it.

Theorem 9 (Theorem 2 in [H4]) *The algorithm as follows:*

For each time moment $t \geq 0$ repeat the following step:

(a) Every player making his move at time t chooses action $a_t = \underline{B}(X_t, s)$.

applied in a semi-Markov mean-field game with total reward satisfying (C1–C5) and

(C6) $\mathcal{A}(s,X)$ does not depend on X.

with $x_0 = \delta_{s^*}$ has the following properties:

(a) For any α, $a^\alpha_{T^\alpha_{i+1}} \geq a^\alpha_{T^\alpha_i}$, $i = 0, 1, \ldots, i^\alpha_e - 1$.

(b) X_t is an increasing function of t converging to some $\mathcal{X}$ as $t \to \infty$, such that $(\underline{B}, \mathcal{X})$ is a stationary mean-field equilibrium in the game.

We have also proved a theorem concerning the relation between the rewards obtained by the players in the semi-Markov mean-field games with total reward and those in their n-person counterparts, defined similarly as in the case of discrete-times games (we skip the details of the definition).

Theorem 10 (Theorem 3 in [H4]) *Let*

$$\mathcal{F}_c = \{f \in \mathcal{F} : f(\cdot|s, X) \text{ does not depend on } X\}.$$

Suppose the assumption (C1) holds. Then for any $\varepsilon > 0$ there exists an $N_\varepsilon \in \mathbb{N}$ such that for $n \geq N_\varepsilon$ the expected reward of player α from playing policy $g \in F_c$ against $f \in F_c$ played by all the other players in the n-person counterpart of the semi-Markov mean-field game with total reward differs from his expected reward when he plays g against f in the mean-field game by at most ε.

The proof is based on the application of the Kurtz theorem (Theorem 5.3 in [56]), which implies that for any closed interval the trajectories of the process of the global state in n-person counterparts of the mean-field game when all players but one use some stationary strategy f converge for $n \to \infty$ uniformly in probability to the solution of (4). Using (C1) together with continuity and positivity of λ, we can find a finite horizon T_ε in such a way that the expected payoff of any player after time T_ε from his birth is a sufficiently small fraction of ε. Applying the Kurtz theorem to the interval $[T^\alpha_0, T^\alpha_0 + T_\varepsilon]$ and using continuity of r and $\widetilde{r}$ we obtain the thesis of the theorem.

The following fact is a natural consequence of the above theorem:

Corollary 1 *Suppose a semi-Markov mean-field game with total reward satisfies (C1–C6) and take any $\varepsilon > 0$. Then for a sufficiently big n, $(\underline{B}(\underline{X}^*, \cdot), \underline{X}^*)$ and $(\overline{B}(\overline{X}^*, \cdot), \overline{X}^*)$ are weak stationary ε-equilibria in the class $\mathcal{F}_c$ in n-person counterparts of the mean-field game.*

4.4 The applications of mean-field game models in wireless telecommunication (papers [H1,H3])

In the next part of this summary we discuss two applications of the mean-field models discussed in previous sections. Both of them are related to some engineering problems arising in wireless telecommunications. The proofs of these results are based on a precise analysis of the properties of the reward functions (which in both cases can be written in a closed form), which makes them all rather technical. We have thus decided not to discuss them in this presentation.

4.4.1 Power control for mobile terminals (paper [H1])

The first problem can be described as follows: A large population of mobile phones fights for access to a base station. Each of them attempts transmission over a sequence of time slots, at each attempt making a decision on the transmission power. The choice it makes is important for three reasons:

- Bigger transmission power means higher throughput.
- Bigger transmission power implies bigger interference for other players, decreasing their throughputs.
- Bigger transmission power means that the mobile's battery gets empty after a shorter period of time.

In our model we assume that the transmission ends when the battery is emptied, hence, each player maximizes his his throughput minus the cost of the transmission over one lifetime of his battery.

As this is a game with a large population of symmetric players, approximating it using a discrete-time mean-field game seems a good idea. The players maximize their utility aggregated over one lifetime of the battery, so the total-reward mean-field game model described in section 4.2.3 should be the most appropriate. The private state s^i of a player i will be the energy available in his battery. We shall assume that $S = \{s_1, s_2, \ldots, s_M\} \cup s^*$ with $s_1 < s_2 < \ldots < s_M$. Actions of the players will be their transmission powers. We shall assume that the set of actions is $A = \{a_1, a_2, \ldots, a_K\}$, with $a_1 < a_2 < \ldots < a_K$. In addition, we assume that the set of actions available in private state $s \neq s^*$ is of the form

$$\mathcal{A}(s) = \{a_1, \ldots, a_{k_s}\} \subset A,$$

and such that for any $s < s'$, $k_s \leq k_{s'}$. The transition probabilities in the private Markov chains of the players will be defined as follows:

- The probability of staying in a private state $s \neq s^*$ by a player using action a is $p(a)$, where

  $$p(a) = 1 - \alpha a - \gamma,$$

  α and γ are some positive constants.
- The state decreases by 1 with the remaining probability.

State s^* corresponds to the situation when the battery of a player is empty. At any given time a mobile whose battery is empty may have it recharged (up to s_M) with probability p_{0N}.

The immediate reward of a player in a private state s using action a, if the global state-action distribution in the game equals τ can be computed using the formula

$$r(s, a, \mu) = \frac{a}{\sigma^2 + C \sum_{k=1}^{K} a_l \sum_{m=1}^{M} \tau_{mk}} - \beta a,$$

where C is the interference parameter, σ^2 is the noise power, while β is the unit energy cost. The first part of the utility function is the so-called *signal to interference ratio* (SINR). It is the standard measure used to evaluate the quality of a wireless transmission (see [48]). The second part is the cost of the energy used for the transmission.

It is rather easy to verify that the game presented above satisfies all the assumptions of both Theorem 3 and Theorem 4, so the existence of a stationary mean-field equilibrium, as well as the fact that such equilibrium corresponds to approximate equilibria in similar models with large finite number of mobiles, are guaranteed. Our main aim in paper [H1] was to prove something more though: firstly, we wanted to find a class of strategies with the simplest possible structure (which in practice means that they are easy to implement) which contains an equilibrium in this game; secondly, we wanted to find efficient algorithms that could allow computing these strategies. The types of strategies of particular interest are defined below:

Any strategies using only the smallest and the biggest powers available in each state:

$$\mathcal{F}_m = \left\{ f \in \mathcal{F} : \exists r_1, \ldots, r_M \in [0,1], f(s_m, \mu) = r_m \delta[a_1] + (1 - r_m)\delta[a_{k_{s_m}}], m = 1, \ldots, M, \mu \in \Delta(S) \right\}.$$

Threshold strategies using only the smallest and the biggest powers available:

$$\mathcal{F}_{mp} = \left\{ f \in \mathcal{F} : \exists s^0 \in S, \exists r \in [0,1], \forall \mu \in \Delta(S), f(s, \mu) = \left\{ \begin{array}{ll} \delta[a_1], & s < s^0 \\ r\delta[a_1] + (1-r)\delta[a_{k_s}], & s = s^0 \\ \delta[a_{k_s}], & s > s^0 \end{array} \right. \right\}.$$

Theorem 11 (Theorem 2 in [H1]) *The game under consideration always possesses a stationary mean-field equilibrium (f^*, μ^*) such that $f^* \in \mathcal{F}_{mp}$. Moreover,*

(i) It is unique in the set $\mathcal{F}_{mp}$.

(ii) If

$$\beta C \frac{\sum_{m=1}^{M} \frac{a_{k_{s_m}}}{\alpha a_{k_{s_m}}+\gamma}}{\sum_{m=1}^{M} \frac{1}{\alpha a_{k_{s_m}}+\gamma} + \frac{1}{p_{0N}}} \leq 1 - \beta\sigma^2, \tag{5}$$

then $f^* = f^+$, where $f^+(s, \mu) := \delta_{a_{k_s}}$, $s = s_1, \ldots, s_M$, $\mu \in \Delta(S)$.

(iii) If

$$\beta C a_1 > 1 - \beta\sigma^2 \quad \text{and} \quad p_{0N} \geq \frac{(\alpha a_1 + \gamma)(1 - \beta\sigma^2)}{M(\beta C a_1 - (1 - \beta\sigma^2))}, \tag{6}$$

then $f^* = f^-$, where $f^-(s, \mu) := \delta_{a_1}$, $s = s_1, \ldots, s_M$, $\mu \in \Delta(S)$.

Theorem 12 (Theorem 3 in [H1]) *Suppose neither f^+ nor f^- is a strategy in a stationary mean-field equilibrium in the game under consideration. Moreover, the set*

$$S_0 := \{s \in S : |\mathcal{A}(s)| > 1\}$$

has at least two elements. Then there is a continuum of stationary mean-field equilibria (f^*, μ^*) in the game such that $f^* \in \mathcal{F}_m$. Moreover, $f^* \in \mathcal{F}_m$ is a strategy in a mean-field equilibrium iff it satisfies the following equation:

$$\sum_{m=1}^{M} \overline{r}_m \left(\frac{1}{\alpha a_1 + \gamma} - \frac{1}{\alpha a_{k_{s_m}} + \gamma)} \right) = \frac{C\beta(M p_{0N} + \gamma}{p_{0N}(\alpha - \alpha\beta\sigma^2 + C\beta\gamma)} - \sum_{m=1}^{M} \frac{1}{\alpha a_{k_{s_m}} + \gamma} - \frac{1}{p_{0N}},$$

where

$$\overline{r}_m := \frac{r_m(\alpha a_1 + \gamma)}{r_m + (1 - r_m)(\alpha a_{k_{s_m}} + \gamma)}$$

with r_m satisfying $f^*(s_m, \mu) = r_m \delta[a_1] + (1 - r_m)\delta[a_{k_{s_m}}]$ for $m = 1, \ldots, M$ and $\mu \in \Delta(S)$.

The next two results give us some efficient methods of computing the equilibrium strategies discussed in the previous theorems:

Theorem 13 (Theorem 4 in [H1]) *define $\theta : S \times [0, 1] \to \mathbb{R}$*

$$\theta(s^0, r) := \sigma^2 + C \frac{\frac{(s^0-1)a_1}{\alpha a_1+\gamma} + \frac{r a_1 + (1-r) a_{k_{s^0}}}{r(\alpha a_1+\gamma)+(1-r)(\alpha a_{k_{s^0}}+\gamma)} + \sum_{s=s^0+1}^{s_M} \frac{a_{k_s}}{\alpha a_{k_s}+\gamma}}{\frac{s^0-1}{\alpha a_1+\gamma} + \frac{1}{r(\alpha a_1+\gamma)+(1-r)(\alpha a_{k_{s^0}}+\gamma)} + \sum_{s=s^0+1}^{s_M} \frac{1}{\alpha a_{k_s}+\gamma} + \frac{1}{p_{0N}}}$$

and $h : \mathcal{F}_{mp} \to [0, M]$

$$h(f) := M + 1 - ind(s^0) - r,$$

where $ind(s) = l \iff s = s_l$. A strategy in a stationary mean-field equilibrium $f^* \in \mathcal{F}_{mp}$ in the game under consideration can be computed by applying bisection to the function

$$\phi(x) = \theta(h^{-1}(x)) - \frac{1}{\beta}$$

on $[0, M]$. The approximate value of f^* will then be given by $h^{-1}(x^*)$, where x^* is the (approximate) zero of ϕ. If for any $x \in [0, M]$, $\phi(x) < 0$, f^+ is the strategy in equilibrium. If for any $x \in [0, M]$, $\phi(x) > 0$, f^- is the strategy in equilibrium.

Theorem 14 (Theorem 5 in [H1])

(i) Let u be any linear function from $\mathbb{R}^N$ to $\mathbb{R}$. The following procedure gives a strategy $f^* \in \mathcal{F}_m$ in stationary mean-field equilibrium in the game under consideration, such that the randomization occurs at most in one private state:

(a) Check whether the parameters of the model satisfy (5) or (6). If they satisfy the former, take $f^* := f^+$; if they satisfy the latter, take $f^* := f^-$. If they do not satisfy any of the two, pass to (b).

(b) Using the simplex method solve the LP:

$$\begin{aligned} &\textit{maximize} && u(\overline{r}_1, \dots, \overline{r}_M) \\ &\textit{subject to} && \sum_{m=1}^{M} \overline{r}_m \left(\frac{1}{\alpha a_1 + \gamma} - \frac{1}{\alpha a_{k_{s_m}} + \gamma} \right) = D - \sum_{m=1}^{M} \frac{1}{\alpha a_{k_{s_m}} + \gamma} \\ & && 0 \leq \overline{r}_m \leq 1, \quad m = 1, \dots, M, \end{aligned}$$

where

$$D = C\beta \frac{M p_{0N} + \gamma}{p_{0N}(\alpha - \alpha\beta\sigma^2 + C\beta\gamma)} - \frac{1}{p_{0N}}, \tag{7}$$

(c) For $m = 1, \dots, M$ compute $r_m := \frac{\overline{r}_m(\alpha a_{k_{s_m}} + \gamma)}{\overline{r}_m(\alpha a_{k_{s_m}} + \gamma) + (1 - \overline{r}_m)(\alpha a_1 + \gamma)}$ and $f^*(s_m, \mu) := r_m \delta[a_1] + (1 - r_s)\delta[a_{k_{s_m}}]$.

(ii) If we take

$$u(\overline{r}_1, \dots, \overline{r}_M) := \sum_{m=1}^{M} G^{M-m} \left(\frac{1}{\alpha a_1 + \gamma} - \frac{1}{\alpha a_{k_{s_m}} + \gamma} \right) \overline{r}_m, \tag{8}$$

where $G > 1$ is any fixed constant, Then f^* obtained is the unique strategy $\mathcal{F}_{mp}$ in stationary mean-field equilibrium in the game.

The next result gives us the method to compute a strategy from $\mathcal{F}_{mp}$ which maximizes the average reward of a player in the game under consideration over the set of all stationary strategies $\mathcal{F}$ (we may assume that this is the strategy that would be chosen, if the choice was made in a centralized manner).

Theorem 15 (Theorem 6 in [H1]) *There exists a strategy $\overline{f} \in \mathcal{F}_{mp}$ which maximizes the average reward of a player in the game under consideration over the set $\mathcal{F}$. It can be computed using the following procedure:*

(a) Check whether the parameters of the model satisfy (6). If they do, take $\overline{f} := f^-$ and terminate. If not, pass to (b).

(b) Find d maximizing

$$H(d) := \left(\frac{1}{\sigma^2 + C \left[\frac{M p_{0N} + \gamma}{\alpha p_{0N}(d + \frac{1}{p_{0N}})} - \frac{\gamma}{\alpha} \right]} - \beta \right) \left(\frac{M}{\alpha} - \frac{\gamma}{\alpha} d \right)$$

on the interval $\left[\sum_{m=1}^{M} \frac{1}{\alpha a_{k_{s_m}} + \gamma}, \frac{M}{\alpha a_1 + \gamma} \right]$, d^*.

(c) Perform steps (b) and (c) of the procedure from Theorem 14 with D replaced by d^* and u given by (8).

The article [H1] also contains some examples comparing the average rewards obtained by the players in cases when they use their equilibrium strategies and when they use strategies given by the above theorem. They show that it is possible that the two will be different. A formal analysis of the impact of independent decision-making on the social welfare in the game (by computing the so-called *price of anarchy*, see [37], which is a standard procedure in problems of this type) is tricky, as in our model the rewards of the players may both be positive and negative. In particular, the rewards in any stationary mean-field equilibrium in the game when neither f^+ nor f^- is an equilibrium strategy is always zero, as implied by the next theorem. It also gives some information about other important metrics applied in practice in this kind of problems.

Theorem 16 (Theorem 7 in [H1]) *Suppose the parameters of the model are such that neither f^+ nor f^- is in stationary mean-field equilibrium in the game under consideration. Then for any stationary mean-field equilibrium (f^*, μ^*) such that $f^* \in \mathcal{F}$:*

(i) The expected reward of any player at equilibrium is

$$\overline{J}(\delta_{s_M}, \mu^*, f^*, f^*) = 0.$$

(ii) The total throughput at any time t if all the players apply strategy f^* is

$$Th(f^*) = \frac{\beta(N - \gamma D)}{\alpha D}$$

with D defined by (7).

(iii) The average lifetime of a battery when all the players use strategy f^* is $T(f^*) = D$.

4.4.2 M/M/∞ queues with service cost shared among the users (paper [H3])

The second practical problem that we have considered which can be formulated as a mean-field game comes from the article [H3]. It can be described as follows: A group of mobile terminals tries to access some resource, whose quality depends on the lack of significant transmission delays. Examples of such resources are live sports broadcasts or some TV shows available on demand. The problem we are dealing with can be described with a help of an M/M/∞ queue. The players are the mobile terminals deciding to queue or not to queue based on the queue length which they observe. Their rewards are computed as the value of service (which is the same for any player and equal to γ) minus the cost of service, which is computed as the integral of a continuous non-increasing[16] function of the queue length $c : \mathbb{R}^+ \to \mathbb{R}^+$ from zero to the time the service is completed. The process of arrival of users and the time of service are modeled by an M/M/∞ queue with known incoming rate λ and service rate μ. As in such a model the number of players is large and all are identical, we may approximate it with a mean-field game. As any player takes part in the game only during his service time, which is an exponential random variable, it will be a semi-Markov mean-field game with total reward as described in section 4.3. Note however, that the only decision made by any player is whether to join the queue or not. Moreover, it is made exactly at the moment when he joins the game. This allows us to simplify the model presented in section 4.3 in the following manner:

- We do not consider private states and assume that the global state of the game at any time t is the length of the queue X_t.
- The set of players' actions is $A = \{E, N\}$, where E denotes 'entering the queue', while N goes for 'not entering the queue'.

[16]The assumption that the cost function is non-increasing is nonstandard. It follows from the interpretation of the model – in case of a live transmission to a number of mobile terminals we may assume that the cost is divided among all the terminals, which implies that the cost per terminal decreases with the number of users.

- The players make their decisions based on the length of the queue X_t, thus their stationary strategies are functions from $\mathbb{R}^+$ to $\Delta(A)$.
- The evolution of the global state when all but a finite number use strategy $f \in \mathcal{F}$ can be described by a counterpart of the ODE (4) which takes the form

$$\begin{cases} \dot{X}_t(f) = \mathbb{E}f(\{E\}|X_t)\lambda - \mu X_t(f), \forall t \geq 0 \\ X_0 = x_0, \end{cases}$$

- The total reward of a player joining the game at time T_0^α and using strategy $f \in \mathcal{F}$ against $g \in \mathcal{F}$ of all the others can be directly computed using the formula[17]:

$$\overline{J}^\alpha(X_{T_0^\alpha}, f, g) = \mathbb{E}f(\{E\}|X_{T_0^\alpha})\left[\gamma - \int_{T_0^\alpha}^{T_0^\alpha+\tau_0^\alpha} c(X_t(g))\,dt,\right]$$

  where τ_0^α is the service time of player α, that is, an exponential (with parameter μ) random variable.

Unlike in the model from [H4], here we try to find the stationary strategy Nash equilibria in the game. This obviously implies that the results presented in section 4.3 cannot be applied in here. The most important result in [H3] is the theorem characterizing all the stationary strategy Nash equilibria in the model. It makes use of the following definition:

Definition 8 *A stationary strategy $f \in \mathcal{F}$ in the game under consideration is a $[\Theta, q]$*-threshold strategy, *if it is of the form:*

$$f(\cdot|x) = \begin{cases} \delta_N(\cdot), & \text{if } x < \Theta \\ q\delta_E(\cdot) + (1-q)\delta_N(\cdot), & \text{if } x = \Theta \\ \delta_E(\cdot), & \text{if } x > \Theta \end{cases}$$

Theorem 17 (Theorem 1 in [H3]) *Let $\overline{\Theta}$ and $\underline{\Theta}$ be the unique solutions to the equations*

$$\frac{1}{\lambda - \overline{\Theta}\mu}\int_{\overline{\Theta}}^{\frac{\lambda}{\mu}} c(u)\,du = \gamma, \quad \frac{1}{\underline{\Theta}\mu}\int_0^{\underline{\Theta}} c(u)\,du = \gamma.$$

The game under consideration always has a symmetric equilibrium where each of the players uses the same $[\Theta, q]$-threshold strategy. Moreover:

(a) If $\gamma \in \left(0, \frac{1}{\mu}\lim_{u\to\infty} c(u)\right]$ then the equilibrium is unique, with $\Theta = \infty$, which means that the equilibrium policies prescribe every user never to enter the queue.

(b) If $\gamma \in \left(\frac{1}{\mu}\lim_{u\to\infty} c(u), \frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\,du\right)$ then there are infinitely many equilibria, whose forms depend on the relation between $\overline{\Theta}$ and $\frac{\lambda}{\mu}$:
(b1) If $\overline{\Theta} < \frac{\lambda}{\mu}$ then there are equilibria of five types: $\Theta = \overline{\Theta}$ and any $q > \frac{\overline{\Theta}\mu}{\lambda}$; $\Theta = \Theta^*$, with Θ^* satisfying $c(\Theta^*) = \mu\gamma$ and $q = \frac{\Theta^*\mu}{\lambda}$; $\Theta = \underline{\Theta}$ and any $q \in [0,1]$; any $\Theta \in \left[\overline{\Theta}, \frac{\lambda}{\mu}\right]$ and $q = 0$; any $\Theta \in \left[\overline{\Theta}, \frac{\lambda}{\mu}\right]$ and $q = 1$.
(b2) If $\overline{\Theta} = \frac{\lambda}{\mu}$ then either $\Theta = \overline{\Theta}$ and $q \in \{0,1\}$ or $\Theta = \underline{\Theta}$ and q is any number from $[0,1]$.
(b3) If $\overline{\Theta} > \frac{\lambda}{\mu}$ then either $\Theta = \underline{\Theta}$ and q is an arbitrary number from $[0,1]$ or $\Theta = \frac{\lambda}{\mu}$ and $q = 0$.

[17] As we do not use private states in the description of this model, here we also skip the first argument of $\overline{J}^\alpha$, which is the initial distribution of the private state of player α.

(c) *If* $\gamma \in \left[\frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\,du, \frac{1}{\mu}c(0)\right)$ *then there are infinitely many equilibria of three types: with* $\Theta \in [0, \underline{\Theta}]$ *and* $q = 0$*; with* $\Theta \in [0, \underline{\Theta}]$ *and* $q = 1$*; with* $\Theta = \Theta^*$ *satisfying* $c(\Theta^*) = \mu\gamma$ *and* $q = \frac{\Theta^*\mu}{\lambda}$.

(d) *If* $\gamma \geq \frac{1}{\mu}c(0)$ *then the equilibrium is unique, with* $\Theta = 0$ *and* $q = 1$*, which means that the equilibrium policies prescribe every user to always enter the queue.*

As in the case of the previous model, we are also interested with a strategy that would have been chosen, if it was chosen in a centralized manner in order to maximize the average reward of a player in the game. The next theorem states, how this strategy should be chosen and what average reward it would induce (depending on the parameters of the model).

Theorem 18 (Theorem 2 in [H3])

(a) *If* $c\left(\frac{\lambda}{\mu}\right) < \gamma\mu$ *then the biggest average reward in the game equals* $\frac{1}{\mu}\left(c\left(\frac{\lambda}{\mu}\right) - \gamma\mu\right)$ *and is attained for the strategy profile consisting of* $[0, 1]$*-threshold strategies of all the players, prescribing to always join the queue.*

(b) *If* $c\left(\frac{\lambda}{\mu}\right) = \gamma\mu$ *then the biggest average reward in the game equals 0 and is attained for any symmetric strategy profile consisting of* $[\Theta, q]$*-threshold strategies such that* $\Theta \neq \frac{q\lambda}{\mu}$.

(c) *If* $c\left(\frac{\lambda}{\mu}\right) > \gamma\mu$ *then the biggest average reward in the game equals 0 and is attained for the strategy profile consisting of* $[\infty, 0]$*-threshold strategies of all the players, prescribing never to join the queue.*

Comparison of average rewards at equilibria with their optimal values given by Theorem 18 leads to the following conclusions:

Corollary 2 (Theorem 3 in [H3]) *Two situations are possible:*

(a) *If* $\gamma \in \left(\frac{1}{\mu}c\left(\frac{\lambda}{\mu}\right), \frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\,du\right)$ *or* $\gamma \in \left[\frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\,du, \frac{1}{\mu}c(0)\right)$ *and* $x_0 \leq \underline{\Theta}$*, then the average player's reward at the worst Nash equilibrium in the game under consideration is zero, even though a positive average reward is possible.*

(b) *In any other situation the average reward at any Nash equilibrium in the game is the same as its optimal value.*

Corollary 3 (Theorem 4 in [H3]) *Two situations are possible:*

(a) *If* $\gamma \in \left(\frac{1}{\mu}c\left(\frac{\lambda}{\mu}\right), \frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\,du\right)$ *and* $x_0 < \overline{\Theta}$*, then the average player's reward at the best Nash equilibrium in the game under consideration is zero, even though a positive average reward is possible.*

(b) *In any other situation there exists a Nash equilibrium in the game the average reward at which is the same as its optimal value.*

Since there are situations when average players' rewards at any Nash equilibrium in the game are zero even though positive averages are possible, in the remainder of paper [H3] we have asked, if it is possible to give incentive to the players to play more efficiently by limiting the information available to them. More precisely, we have assumed that the players are not able to observe the exact value of the queue's length. Instead, the information they have access to is the answer to the question whether current value of X_t is greater or smaller than some predefined constant Ψ. It is a specific game of incomplete information where players do not have access to any statistical information about exact situation in the game. The solution used in this type of games is that of *robust Nash equilibrium*, see [2]. In case of our game it can be defined as follows:

Definition 9 *We say that a strategy $f \in \mathcal{F}$ of player α is a best robust response to the strategy $g \in \mathcal{F}$ applied by the others, if*

$$f(\cdot|x_1) = f(\cdot|x_2), \text{ if } x_1, x_2 < \Psi \text{ or } x_1, x_2 \geq \Psi. \tag{9}$$

Moreover, for any other strategy $h \in \mathcal{F}$ satisfying (9),

$$\inf_{x<\Psi} \overline{J}^{\alpha}(x, f, g) \geq \inf_{x<\Psi} \overline{J}^{\alpha}(x, h, g) \text{ and } \inf_{x\geq\Psi} \overline{J}^{\alpha}(x, f, g) \geq \inf_{x\geq\Psi} \overline{J}^{\alpha}(x, h, g)$$

We say that a profile of strategies where all the players use strategy f is a symmetric robust Nash equilibrium *in the game under consideration, if f is a best robust response to f applied by all other players.*

The next theorem describes how the above-defined solution depends on the parameters of the model and the value of Ψ. It uses the following convention: we say that a player uses strategy (a_1, a_2), $a_1, a_2 \in A$, if he chooses a_1, if the global state is smaller than Ψ and action a_2, if the global state is greater or equal Ψ.

Theorem 19 (Theorem 5 in [H3]) *Let:*

$$L_{EE}(\Psi) := \frac{1}{\lambda} \int_0^{\frac{\lambda}{\mu}} c(u)\, du, \quad L_{NE}(\Psi) := L_{NN}(\Psi) := \frac{1}{\mu} c(0), \quad H_{EE}(\Psi) := \frac{1}{\lambda - \Psi\mu} \int_{\Psi}^{\frac{\lambda}{\mu}} c(u)\, du,$$

$$H_{NE}(\Psi) := \begin{cases} \frac{1}{\Psi\mu} \int_0^{\Psi} c(u)\, du, & \text{if } \Psi > \frac{\lambda}{\mu} \\ \frac{1}{\lambda - \Psi\mu} \int_{\Psi}^{\frac{\lambda}{\mu}} c(u)\, du, & \text{if } \Psi \leq \frac{\lambda}{\mu} \end{cases}, \quad H_{NN}(\Psi) := \frac{1}{\Psi\mu} \int_0^{\Psi} c(u)\, du.$$

For any $\Psi \geq 0$ the game with partial information under consideration has a pure-strategy symmetric robust Nash equilibrium. Moreover:

(a) When $\gamma > L_{NN}(\Psi)$ then all the players use policy EE in equilibrium;

(b) When $L_{NN}(\Psi) \geq \gamma \geq L_{EE}(\Psi)$ and $\gamma > H_{NN}(\Psi)$ then strategy profiles where all the players use policy EE and where all the players use policy NE are equilibria;

(c) When $H_{NN}(\Psi) \geq \gamma \geq \max\{L_{EE}(\Psi), H_{NE}(\Psi)\}$ then any strategy profile where all the players use the same policy is an equilibrium;

(d) When $H_{NE}(\Psi) > \gamma \geq L_{EE}(\Psi)$ then strategy profiles where all the players use policy EE and where all the players use policy NN are equilibria;

(e) When $L_{EE}(\Psi) > \gamma > H_{NN}(\Psi)$ then all the players use policy NE in equilibrium;

(f) When $\min\{L_{EE}(\Psi), H_{NN}(\Psi)\} \geq \gamma \geq H_{NE}(\Psi)$ then strategy profiles where all the players use policy NE and where all the players use policy NN are equilibria;

(g) When $\min\{L_{EE}(\Psi), H_{NE}(\Psi)\} > \gamma$ then all the players use policy NN in equilibrium.

Using the last theorem, we are able to choose Ψ in order to maximize the average of players' rewards at robust Nash equilibria:

Theorem 20 (Theorem 7 in [H3]) *If we want to maximize the average of players' rewards at robust Nash equilibrium in the game with partial information under consideration, we should choose:*

(a) *Any* Ψ, *if* $\gamma \leq \frac{1}{\mu}\lim_{u\to\infty} c(u)$ *– then all the players will use strategy* NN *in equilibrium.*

(b) *Any* $\Psi < \underline{\Theta}$, *if* $\gamma \in \left(\frac{1}{\mu}\lim_{u\to\infty} c(u), \frac{1}{\mu}c\left(\frac{\lambda}{\mu}\right)\right)$ *– then all the players will use strategy* NN *in equilibrium.*

(c) $\Psi = \overline{\Theta}$, *if* $\gamma \in \left(\frac{1}{\mu}c\left(\frac{\lambda}{\mu}\right), \frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\, du\right)$ *– then all the players will use strategy* NE *in equilibrium.*

(d) $\Psi = \underline{\Theta}$, *if* $\gamma \in \left[\frac{1}{\lambda}\int_0^{\frac{\lambda}{\mu}} c(u)\, du, \frac{1}{\mu}c(0)\right)$ *– then in the pessimistic scenario all the players will use strategy* NE *in equilibrium, while in the optimistic one all the players will use strategy* EE *in equilibrium.*

(e) $\Psi = 0$, *if* $\gamma = \frac{1}{\mu}c(0)$ *– then all the players will use strategy* EE *in equilibrium.*

(f) *Any* Ψ, *if* $\gamma > \frac{1}{\mu}c(0)$ *– then all the players will use strategy* EE *in equilibrium.*

Comparing the average rewards obtained in the above theorem to the optimal average rewards, we arrive at the following conclusion:

Corollary 4 (Theorem 8 in [H3]) *Average rewards in the best and the worst robust Nash equilibria in the game with partial information when* Ψ *is chosen in order to maximize these averages are the same as those in the best and the worst Nash equilibria in the game with complete information.*

The last problem that we have addressed in [H3] was the standard problem of the quality of approximation of the real-life models with a large finite numbers of players by our model with infinitely many players. It should be noted here that the proper rescaling of our mean-field game to n-person case is a game where the players decide whether to join or not to join an M/M/∞ queue with incoming rate $n\lambda$, service rate μ and the service cost $c^n(x) := c\left(\frac{x}{n}\right)$. We have formulated the following result:

Theorem 21 (Theorem 9 in [H3]) *Suppose that the initial normalized*[18] *state of the queue* $x_0 \in [0, x_{max}]$ *for some fixed* x_{max} *and that player* α *plays against* $[\Theta, q]$*-threshold strategies of all the others in the mean-field model with service cost* c*, incoming rate* λ *and service rate* μ*. Then for any* $\varepsilon > 0$ *there exists an* $N_\varepsilon \in \mathbb{N}$*, such that for any* $n \geq N_\varepsilon$ *his expected total reward if he joins the queue in the* n*-person counterpart of the mean-field game where all the other players use* $[n\Theta, q]$*-threshold strategies differs from his expected total reward in the mean-field game by at most* ε*.*

One of the crucial consequences of this theorem is that all the equilibria obtained in Theorems 17 and 19 can be treated as approximate equilibria in n-person models. It should be noted though, that formally for this to be true, the normalized state of the queue at any moment when any player decides whether to join the queue or not should not exceed x_{max}, which is not always satisfied. These statements (with all the assumptions necessary) have been formulated in [H3] as Corollary 4, Corollary 4 and Corollary 5.

5 Description of other scientific achievements

List of papers which have not been included into the scientific achievement:

[P1] P. Więcek, *On application of Schauder's fixed point theorem in discounted stochastic games.* Prace Naukowe Instytutu Matematyki Politechniki Wrocławskiej. Seria: Konferencje. (I Konferencja dla Młodych Matematyków - Karpacz 2000) 24 (2000), no. 3, 121–128.

[18] The normalized state of the queue in the n-person counterpart of the mean-field queuing game under consideration is $\overline{X}_t = \frac{X_t}{n}$.

[P2] P. Więcek, *Convex Stochastic Games of Capital Accumulation with Nondivisible Money Unit.* Scientiae Mathematicae Japonicae, 57 (2003), 397–411

[P3] P. Więcek, *Continuous Convex Stochastic Games of Capital Accumulation.* In Advances in Dynamic Games. Applications to Economics, Finance, Optimization and Stochastic Control (Annals of the International Society of Dynamic Games vol. 7), A.S. Nowak, K. Szajowski eds., Birkhäuser, Boston, 2005, 111–125

[P4] P. Więcek, T. Radzik, *On a continuous dynamic strategic market game.* In Game Theory and Applications vol. 11, L. Petrosjan, V. Mazalov eds., Nova Science Publishers, Commack, NY, 2007, 187–195

[P5] W. Połowczuk, P. Więcek, T. Radzik, *On the existence of almost-pure-strategy Nash equilibria in n-person finite games.* Mathematical Methods of Operations Research, 65 (2007), 141-152

[P6] A.S. Nowak, P. Więcek, *On Nikaido-Isoda type theorems for discounted stochastic games.* Journal of Mathematical Analysis and Applications, 332 (2007), 1109–1118

[P7] P. Więcek, *Pure equilibria in a simple dynamic model of strategic market game.* Mathematical Methods of Operations Research, 69 (2009), 59–79

[P8] P. Więcek, *n-person dynamic strategic market games.* Applied Mathematics & Optimization, vol. 65, no. 2 (2012), 147–173

[P9] W. Połowczuk, T. Radzik, P. Więcek, *Simple equilibria in semi-infinite games.* International Game Theory Review, 14(3) (2012) 1250017-1–1250017-19

[P10] M. Haddad, P. Więcek, E. Altman, H. Sidi, *An Automated Dynamic Offset for Network Selection in Heterogeneous Networks.* IEEE Transactions on Mobile Computing, 15(9) (2016), 2151–2164

[P11] M. Haddad, P. Więcek, O. Habachi, Y. Hayel, *On the Two-User Multi-Carrier Joint Channel Selection and Power Control Game.* IEEE Transactions on Communications, 64(9) (2016), 3759–3770

In the article [P1] I have presented the results from my Master's thesis, while the papers [P2,P3,P4,P7] covered the topics included in my Ph.D. dissertation. Below I briefly describe the results presented in all the other articles.

5.1 Stochastic games (paper [P6])

The first of the above-mentioned papers deals with two-person non-zero sum stochastic games with discounted reward. In games of this type the players jointly control a Markov chain of states of the game, choosing at each of infinitely many stages actions which then influence the distribution of the state at the next stage of the game. Also at each stage every player obtains an immediate reward depending on the current state as well as the actions chosen by the players. These immediate rewards at all the stages of the game are then aggregated into discounted rewards for each of the players. As this is a non-cooperative game, each player tries to maximize his own discounted reward, knowing that his opponent does the same (for a precise description of games of this type see [30], chapter 8). The game is thus described with a set of states S, sets of actions A_1 and A_2, correspondences describing the sets of actions available in each of the states $\mathcal{A}_i : S \to A_i$, $i = 1, 2$, immediate reward functions $r_i : D \to \mathbb{R}$, $i = 1, 2$, where $D := \{(s, a_1, a_2) : a_1 \in \mathcal{A}_1(s), a_2 \in \mathcal{A}_2(s)\}$, a transition probability $Q : D \to \Delta(S)$ and a discount factor $\beta \in (0, 1)$. In the classic paper of Federgruen [20] it is proved that each stochastic game with a countable state space, compact action sets and continuous immediate rewards and transition probability has a Nash equilibrium in stationary strategies. The main result of [P6] is a generalization of this result in case of two-person games:

Theorem 22 (Theorem 1 in [P6]) *Every two-person non-zero sum stochastic game with a finite set of states S and A_i, $\mathcal{A}_i(s)$, $i = 1, 2$, $s \in S$ being compact subsets of a metric space, satisfying in addition:*

(D1) For every $s \in S$, the function $r_1(s, \cdot, \cdot) + r_2(s, \cdot, \cdot)$ is continuous on $\mathcal{A}_1(s) \times \mathcal{A}_2(s)$;

(D2) For every $s \in S$ and $a_1 \in \mathcal{A}_1(s)$, $a_2 \in \mathcal{A}_2(s)$, the functions $r_i(s, \cdot, a_2)$ and $r_i(s, a_1, \cdot)$, $i = 1, 2$, are continuous on $\mathcal{A}_1(s)$ and $\mathcal{A}_2(s)$, respectively;

(D3) For every $s, s' \in S$, the function $Q(s'|s, \cdot, \cdot)$ is continuous on $\mathcal{A}_1(s) \times \mathcal{A}_2(s)$;

has a stationary Nash equilibrium in the class of all strategies of the players.

The proof of this theorem is based on an approximation of the stochastic game under consideration with games with finite sets of actions A_1^n, A_2^n constructed in such a way that $A_i^n \subset A_i^{n+1}$, $n = 1, 2, \ldots$ and $\bigcup_{n=1}^{\infty} A_i^n$ are dense in A_i ($i = 1, 2$). Each of such games has a stationary equilibrium by Federgruen's theorem. If we take the limit over a subsequence of the sequence of stationary strategies in equilibria in the approximating games (such a limit exists, as the sets of stationary strategies of the players are compact by Tikhonov's theorem), we obtain a pair of stationary strategies in the approximated game, (f^*, g^*). Using assumptions (D1) and (D2) we can then prove that the players' discounted rewards preserve the continuity properties of immediate rewards (as functions of stationary strategies of the players). Using these properties we then show that (f^*, g^*) is a Nash equilibrium in the stochastic game discussed in the theorem.

The second result proved in [P6] (Theorem 2) says that under additional assumptions about the concavity of immediate rewards and affinity of transition probability, the equilibrium obtained in Theorem 22 is a nonrandomized one. The proof of this fact is elementary. It is based on the observation that under these additional assumptions the function on the RHS of the Bellman equation of each player is concave.

5.2 Dynamic strategic market games (paper [P8])

The next paper was a continuation of articles [P4,P7] based on parts of my Ph.D. thesis. Its subject was an n-person discounted stochastic game based on the following simple economic model: n players – each holding an integral amount of money – compete for portions of some nondurable commodity. At each of infinitely many stages of the game, one (nondivisible) unit of the good is brought to the market, and players compete for it in an auction, bidding integral parts of their money. If a bid of a player is accepted, he consumes the good (which gives him some positive utility), but he also pays for it, which decreases his budget for the future consumption. We assume that there is also a constant inflow of money into the game with budgets of the players supplied with some random amounts at each stage of the game. The state in this game is the vector of amounts of money possessed by each of the players $s = (s_1, \ldots, s_n) \in \mathbb{N}^n$, while the actions are the offers of the players. Hence, for any i, $a_i \in \{0, , 1, \ldots, s_i\}$. The other objects defining this game (the immediate rewards and the transition probability) are given by the formulas:[19]

$$r_i(s, a_1, \ldots, a_n) = \begin{cases} \frac{1}{|\{j : a_j = a_i\}|} u_i(1) & \text{if } a_i = \max_j a_j \\ u_i(0) & \text{if } a_i < \max_j a_j \end{cases},$$

where u_i is the utility function of player i;

$$Q(\cdot|s, a_1, \ldots, a_n) = \frac{1}{n}\delta_s(\cdot) + \frac{1}{n}\sum_{k \neq i} \delta_{(s_{-i-k}, s_i - a_i, s_k + a_i)}(\cdot), \text{ where } i = \arg\max_j a_j.$$

Two variants of games of this type were considered in the literature before: my previous papers [P4,P7] and [59, 57, 58] concentrated on the 2-player case, while [35, 36, 22] analyzed similar models

[19] (s_{-i-k}, s_i', s_k') denotes here the vector s with its i-th coordinate replaced by s_i' and its k-th one replaced by s_k'.

with a continuum of players. The only results dealing with n-person games of this type appeared in [47] (there however the good was divisible and divided proportionally to the offers made by the players). Two main results of [P8] characterize the equilibria in the game and present some basic properties of the expected discounted rewards of the players when the equilibrium strategies are applied under the assumption that the discount factor is small enough. Below we present the first of these theorems.

Theorem 23 (Theorem 1 in [P8]) *Let us define the following classes of strategies in the dynamic strategic market game under consideration:*
A stationary strategy f_i of player i is called bold, *if $f_i(\cdot|s) = \delta_{s_i}$ for $s_i \leq \max_{j\neq i} s_j$ and $f_i(\{\max_{j\neq i} s_j + 1, \ldots, s_i\}|s)) = 1$ for $s_i > \max_{j\neq i} s_j$.*
A stationary strategy f_i of player i is called weakly bold, *if $f_i(\cdot|s) = \delta_{s_i}$ for $s_i \leq \max_{j\neq i} s_j$ and $f_i(\{\max_{j\neq i} s_j, \ldots, s_i\}|s)) = 1$ for $s_i > \max_{j\neq i} s_j$.*
The following statements are true:

(a) For every $n \geq 3$ and every $\beta \leq \frac{1}{3}$, the game possesses a symmetric stationary equilibrium $f = (f_1, f_2, \ldots, f_n)$, where f_i are bold strategies.

(b) For every $n \geq 3$ and every $\beta \leq 1 - \sqrt[3]{\frac{2n^2-6n+4}{n^3}}$, the game possesses a symmetric stationary equilibrium $f = (f_1, f_2, \ldots, f_n)$, where f_i are weakly bold strategies.

(c) For every $\beta \geq \frac{1}{2}$ there exists an n_β such that for every $n \geq n_\beta$, there is no strategy profile where the strategies of all the players are bold which is a stationary equilibrium in the game.

An important consequence of part (b) of the above theorem is stated by the corollary below:

Corollary 5 (Corollary 1 in [P6]) *For every $\beta \in (0,1)$ there exists an $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$ the n-person dynamic strategic market game possesses a symmetric stationary equilibrium $f = (f_1, f_2, \ldots, f_n)$, where f_i are weakly bold strategies.*

It means that the results obtained in [P6] are true for any discount factor, provided that the number of players n is big enough.

The proofs of Theorem 23 and Theorem 2 in [P6] are rather technical. Their main part is a detailed analysis of the properties of the RHS of the Bellman equation in the problem of maximization of discounted reward of a player when his opponents use strategies from the classes defined in Theorem 23. The application of the Banach fixed point theorem allows then to prove that the optimal discounted rewards in this case have the same properties. They are next used to show that the best responses to any vector of (weakly) bold strategies in the game is a strategy of the same type (under assumptions given in parts (a) and (b) of Theorem 23, respectively). In the last step of the proof we apply the Kakutani fixed point theorem to the best-response correspondence (against a vector of identical (weakly) bold strategies of other players). The fixed point obtained is the strategy used in the symmetric Nash equilibrium in the game.

5.3 Equilibria in one-stage games with a specific structure (papers [P5,P9])

In article [P5] we have considered n-person one-stage games with finite strategy sets and utility functions having some properties which can be identified as discrete counterparts of convexity or concavity. We define them below:

Definition 10 *Let $E_1 = \{1, \ldots, k_1\}, \ldots, E_n = \{1, \ldots, k_1\}$ and $E = E_1 \times \ldots \times E_n$. We say that a function $H : E \to \mathbb{R}$ is* convex (concave) *in its i-th variable, if for $j = 1, \ldots, n$ there exist strictly increasing sequences $(x^j_1, \ldots, x^j_{k_j})$ of elements of $[0,1]$ and a function $\overline{H} : [0,1]^n \to \mathbb{R}$ which is convex (concave) in its i-th variable, and satisfying for every $i_1 \in E_1, \ldots, i_n \in E_n$ the equality $H(i_1, \ldots, i_n) = \overline{H}(x^1_{i_1}, \ldots, x^n_{i_n})$.*

Two main results of [P5] are the following theorems:

Theorem 24 (Theorem 4.4 in [P5]) *Let $1 \leq s \leq n$. If in an n-person one-stage game with strategy sets $E_1,\ldots,E_n$ and utility functions $H_1,\ldots,H_n$, for each $i \in \{1,\ldots,s\}$, H_i is concave in its i-th variable, then the game has a mixed-strategy Nash equilibrium $(\mu_1^*,\ldots,\mu_n^*)$ where each of the strategies μ_i^*, $i = 1,\ldots,s$ is concentrated on two adjoining elements of E_i.*

Theorem 25 (Theorem 4.5 in [P5]) *Let $1 \leq s \leq n$. If in an n-person one-stage game with strategy sets $E_1,\ldots,E_n$ and utility functions $H_1,\ldots,H_n$, for each $i \in \{1,\ldots,s\}$, H_i is convex in its i-th variable, then the game has a mixed-strategy Nash equilibrium $(\mu_1^*,\ldots,\mu_n^*)$ where each of the strategies μ_i^*, $i = 1,\ldots,s$ is concentrated on the set $\{1, k_i\}$.*

The proof of the first theorem is based on the observation that for each $i \in \{1,\ldots,s\}$ the set of mixed strategies concentrated on adjoining elements of set E_i is homeomorphic with $X_i := [1, k_i]$. We then define a game with strategy sets $X_1,\ldots,X_s,\Delta(E_{s+1}),\ldots,\Delta(E_n)$ and utility functions computed as the expected utilities in the initial game corresponding to the strategies which are homeomorhic to respective strategies from X_i. Such a game satisfies the assumptions of the well-known generalization of the Nash theorem for games with quasi-concave utilities due to Glicksberg (see [24]). The pure strategy equilibrium in the modified game corresponds to the mixed-strategy equilibrium we are looking for in Theorem 24. The proof of Theorem 25 is also based on a modification of the game given there – this time we reduce the strategy sets in the game to $\{1, k_1\},\ldots,\{1, k_s\}$, $E_{s+1},\ldots,E_n$. This new game has a mixed-strategy Nash equilibrium by the Nash theorem. It is then immediate to show that the equilibrium in the modified game is exactly the equilibrium in the initial game we are looking for.

In article [P9] we have presented some generalizations of the above results and the results presented in [50, 51, 45, 46] to two-person games, where the set of pure strategies of one of the players is finite, while that of his opponent is infinite but countable (or, in case of some results, it is any compact metric set). Seven theorems presented there give conditions similar to those appearing in Theorems 24 and 25 for the existence of Nash equilibria (or ε-Nash equilibria in some cases) where the players use pure strategies or mixed strategies with two-point supports.

5.4 Network (channel) selection in wireless networks (papers [P10,P11])

The last two papers we shall discuss here concern some applications of noncooperative games in wireless telecommunication. In the first of the two, [P10], we consider a game between mobile phones deciding which of the two networks to connect to. In the case of the first network, the quality of connection is guaranteed by the operator. The quality of connection to the second network depends on the channel the player gets access to. It may therefore be either better or worse than that for the first network. In addition, we assume that the player has no access to the exact information about the quality of his channel. All that he knows is the information whether the channel quality indicator, denoted by h_i, is bigger or smaller than some predefined threshold Ψ_i plus the probability distribution of h_i (in real wireless networks this distribution is known to be exponential, and the parameter of this distribution can be estimated using some historical data). In addition, he knows that the quality of connection to this network also depends on the number of the phones connected there. This situation can be described as an n-person one-stage game of incomplete information. The main results of the article can be divided into two groups. The first group consists of characterizations of Bayes-Nash equilibria in this game in 2-person case (Proposition 1 in [P10]) and in the symmetric[20] n-person case (Proposition 5). The second group tries to answer the questions, how the thresholds Ψ_i should be chosen in order to obtain one of two practical goals:

(a) to maximize the throughput of average user;

[20]Symmetry here means that the distributions from which all the channel quality indicators h_i are drawn are the same, and that each threshold Ψ_i is the same. In the two-person case we have also considered the case when these parameters differ.

(b) to maximize the total throughput of the first network.

The paper gives answers to them in the 2-person case (Proposition 3) and the symmetric n-person case (Proposition 6). In addition we present the analysis of computational complexity of the algorithm given in Proposition 6 (in Proposition 8) and a comprehensive numerical study of the proposed solutions.

In [P11] we consider the situation when two mobile phones send some data using a wireless network with K channels, taking into account the quality of each channel, which may be different both for different phones and different channels (in this case we assume that the players have full information about the quality of all the channels). The choice they make is about the powers they use to transmit on different channels. In this case however they do not maximize their throughput, but their *energy efficiency*, which can be defined as the ratio of some measure of the quality of transmission and the energy used for this transmission (the formal definition of energy efficiency is given in [27]). The problem of this type has already been considered in [41], where the problem has been presented as a non-cooperative game and a heuristic algorithm allowing to compute a Nash equilibrium in this game has been given. The question that we have posed was the following: does introducing hierarchy in this model (which in game-theoretic terms means replacing Nash equilibrium with Stackelberg solution) affect the solution to this game and energy efficiency corresponding to this solution. It is worth noting that it is not a question without a practical meaning – in all real networks the choices made by the players are asynchronous. The main results of [P11] were two algorithms: the first one (given in Proposition 3) allowed to compute the Stackelberg solution in this model, if it exists. The second one (from Proposition 4) allowed to compute ε-Stackelberg solutions, if it does not. Further results (Proposition 7–Proposition 10) were meant to compare in several ways the quality of solutions obtained by our algorithms to that of Nash equilibria computed in [41]. We have also presented some numerical illustrations of our findings.

References

[1] S. Adlakha, R. Johari, *Mean field equilibrium in dynamic games with strategic complementarities.* Operations Research 61(4) (2013), 971–989

[2] M. Aghassi, D. Bertsimas, *Robust game theory.* Mathematical Programming 107 (2006), 231–273

[3] A. Alj, A. Haurie, *Dynamic equilibria in multigenerational stochastic games.* IEEE Trans. Automat. Control 28 (1983), 193–203

[4] R. Amir R, *Complementarity and Diagonal Dominance in Discounted Stochastic Games.* Annals of Operations Research 114 (2002), 39–56

[5] Ł. Balbus, K. Reffett, Ł. Woźny, *A Constructive Study of Markov Equilibria in Stochastic Games with Strategic Complementarities.* Journal of Economic Theory 150 (2014), 815–840

[6] R. Basna, A. Hilbert, V.N. Kolokoltsov, *An approximate Nash equilibrium for pure jump Markov games of mean-field-type on continuous state space.* Stochastics 89(6–7) (2017), 967–993

[7] D. Bauso, *Game Theory with Engineering Applications (Advances in Design and Control).* SIAM–Society for Industrial and Applied Mathematics (2016)

[8] A. Bensoussan, J. Frehse, P. Yam, *Mean field games and mean field type control theory.* Springer, New York (2013)

[9] J. Bergin, D. Bernhardt, *Anonymous sequential games with aggregate uncertainty.* Journal of Mathematical Economics 21 (1992), 543–562

[10] J. Bergin, D. Bernhardt, *Anonymous sequential games: existence and characterization of equilibria.* Economic Theory 5(3) (1995), 461–89

[11] D.P. Bertsekas, S.E. Shreve, *Stochastic optimal control: the discrete time case.* Academic Press, New York (1978)

[12] E. Boissard, *Simple bounds for convergence of empirical and occupation measures in 1-Wasserstein distance.* Electronic Journal of Probability 16 (2011), 2296–2333

[13] L. Campi, M. Fischer, *N-player games and mean-field games with absorption.* Ann. Appl. Probab. 28(4) (2018), 2188–2242

[14] R. Carmona, F. Delarue, *Probabilistic Theory of Mean Field Games with Applications.* Springer International Publishing (2018)

[15] A. Cecchin, M. Fischer, Probabilistic Approach to Finite State Mean Field Games. M. Appl Math Optim. (2018), https://doi.org/10.1007/s00245-018-9488-7

[16] S.K. Chakrabarti, *Pure strategy Markov equilibrium in stochastic games with a continuum of players.* J. Math. Econom. 39(7) (2003), 693–724

[17] L. Curtat, *Markov Equilibria in Stochastic Games with Complementarities.* Games and Economic Behavior 17 (1996), 177–199

[18] R.M. Dudley, *Real analysis and probability.* Cambridge University Press (2004)

[19] R. Elliot, X. Li, Y. Ni, *Discrete time mean-field stochastic linear-quadratic optimal control problems.* Automatica 49 (2013), 3222–3233

[20] A. Federgruen, *On N-Person Stochastic Games with Denumerable State Space.* Adv. Appl. Prob. 10 (1978), 452–471

[21] E. Ferreira, D.A. Gomes, *On the convergence of finite state mean-field games through Γ-convergence.* Journal of Mathematical Analysis and Applications 418(1) (2014), 211–230

[22] J. Geanakoplos, I. Karatzas, M. Shubik, W.D. Sudderth, *A strategic market game with active bankrupcy.* Journal of Mathematical Economics 34 (2000), 359–396

[23] D. Gillette, *Stochastic games with zero stop probabilities.* In Dresher M., Tucker A.W., Wolfe P. (eds.), Contributions to the Theory of Games, Vol. III, Annals of Mathematics Studies 39, Princeton University Press, Princeton, NJ (1957), pp. 179–187.

[24] I. Glicksberg, *A further generalization of the Kakutani's fixed point theorem with applications to Nash equilibrium points.* Proceedings of the American Mathematical Society 38 (1952), 170–174

[25] D.A. Gomes, J. Mohr, R.R. Souza, *Continuous Time Finite State Mean Field Games.* Appl Math Optim 68 (2013), 99–143

[26] D.A. Gomes, J. Saúde, *Mean field games models—a brief survey.* Dyn. Games Appl. 4(2) (2014), 110–154

[27] D. Goodman, N. Mandayam, *Power control for wireless data.* IEEE Pers. Commun. 7(2) (2000), 48–54

[28] E. Green, *Noncooperative Price Taking in Large Dynamic Markets.* Journal of Economic Theory 22 (1980), 155–181

[29] E. Green, *Continuum and Finite–Player Noncooperative Models of Competition.* Econometrica 52 (1984), 975–993

[30] A. Haurie, J.B. Krawczyk, G. Zaccour, *Games and Dynamic Games.* World Scientific (2012)

[31] U. Horst, *Stationary Equilibria in Discounted Stochastic Games with Weakly Interacting Players.* Games and Economic Behavior 51(1) (2005), 83–108

[32] D. Housman, *Infinite Player Noncooperative Games and the Continuity of the Nash Equilibrium Correspondence.* Mathematics of Operations Research 13 (1988), 488–496

[33] M. Huang, R.P. Malhamé, P.E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle.* Communications in Information Systems 6 (2006), 221-252

[34] B. Jovanovic, R.W. Rosenthal, *Anonymous sequential games.* Journal of Mathematical Economics 17 (1988), 77–87

[35] I. Karatzas, M. Shubik, W.D. Sudderth, *Construction of Stationary Markov Equilibria in a Strategic Market Game.* Math. Oper. Res. 19(4) (1992), 975–1006

[36] I. Karatzas, M. Shubik, W.D. Sudderth, *A Strategic Market with Secured Lending.* J. Math. Econ. 28 (1997), 207–247

[37] E. Koutsoupias, C. Papadimitriou, *Worst-case equilibria.* Computer Science Review 3(2) (1999), 65–69

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

[39] A.B. MacKenzie, L.A. DaSilva, *Game Theory for Wireless Engineers.* Morgan & Claypool Publishers (2006)

[40] M. Maschler, E. Solan, S. Zamir, *Game Theory.* Cambridge University Press (2013)

[41] F. Meshkati, M. Chiang, H.V. Poor, S.C. Schwartz, *A game-theoretic approach to energy-efficient power control in multicarrier CDMA systems.* IEEE J. Sel. Areas Commun. 24(6) (2006), 1115–1129

[42] A.S. Nowak, *On Stochastic Games in Economics.* Mathematical Methods of Operations Research 66(3) (2007), 513–530

[43] A.S. Nowak, *On a Noncooperative Stochastic Game Played by Internally Cooperating Generations.* Journal of Optimization Theory and Applications 144(1) (2010), 88–106

[44] E.S. Phelps, R.A. Pollak, *On second-best national savings and game equilibrium growth.* Rev. Econ. Stud. 35 (1968), 195–199

[45] W. Połowczuk, *Pure Nash equilibria in finite two-person non-zero-sum games.* International Journal of Game Theory 32 (2003), 229–240

[46] W. Połowczuk, *On almost-pure-strategy Nash equilibria in bimatrix games with convexity properties.* Applicationes Mathematicae 33 (2006), 71-84

[47] L. Pontiggia, *Topics in stochastic games.* Ph.D. dissertation, University of Minnesota, United States (2004)

[48] J.G. Proakis, *Communication Systems Engineering.* Prentice Hall International Editions (1994)

[49] M. Puterman, *Markov Decision Processes.* Wiley-Interscience, New York (1994)

[50] T. Radzik, *Pure-strategy ε-Nash equilibrium in two-person non-zero-sum games.* Games and Economic Behavior 3 (1991), 356–371

[51] T. Radzik, *Characterization of optimal strategies in matrix games with convexity properties.* International Journal of Game Theory 29 (2000), 211–227

[52] S.M. Ross, *Stochastic Processes.* 2nd Edition, John Wiley & Sons (1996)

[53] H. Sabourian, *Anonymous repeated games with a large number of players and random outcomes.* Journal of Economic Theory 51(1) (1990), 92–110

[54] N. Saldi, T. Başar, M. Raginsky, *Markov-Nash equilibria in mean-field games with discounted cost.* (2016), arXiv:1612.07878

[55] D. Schmeidler, *Equilibrium points of nonatomic games* J. Statist. Physics 17 (1973), 295–300

[56] A. Schwartz, A. Weiss, *Large Deviations for Performance Analysis.* Chapman & Hall, London (1995)

[57] P. Secchi, W.D. Sudderth, *A Two-person Strategic Market Game.* Quaderni di Dipartamento 83(6-98) (1998), Dipart. di Econ. Polit. e Met. Quant., Università di Pavia

[58] P. Secchi, W.D. Sudderth, *A Simple Two-Person Stochastic Game with Money.* In Nowak A.S., Szajowski K. (eds.) Advances in Dynamic Games. Applications to Economics, Finance, Optimization, and Stochastic Control. Ann. of ISDG vol. 7. Birkhäuser, Boston (2005), pp. 39–66

[59] M. Shubik, W. Whitt, *Fiat Money in an Economy with One Nondurable Good and No Credit. A Noncooperative Sequential Game.* In Blaquiere (ed.) Topics in Differential Games. North Holland, Amsterdam (1973), pp. 401–449

[60] C. Sleet, *Markov perfect equilibria in industries with complementarities.* Economic Theory 17(2) (2001), 371–397

[61] A. Tarski, *A lattice-theoretical fixpoint theorem and its applications.* Pacific Journal of Mathematics 5 (1955), 285–309

[62] D.M. Topkis, *Supermodularity and complementarity.* Princeton University Press, Princeton, New Jersey (1998)

[63] X. Vives, *Strategic Complementarity in Multi-Stage Games.* Economic Theory 40(1) (2009), 151–171

[64] J.G. Wardrop, *Some theoretical aspects of road traffic research.* Proc. Inst. Civ. Eng. 2 (1952), 325–378

P. Więcek