

# Correlated equilibria and mean field games: a simple model

Luciano Campi\*      Markus Fischer†

April 13, 2020; last revision June 3, 2021

## Abstract

In the context of simple finite-state discrete time systems, we introduce a generalization of mean field game solution, called correlated solution, which can be seen as the mean field game analogue of a correlated equilibrium. Our notion of solution is justified in two ways: We prove that correlated solutions arise as limits of exchangeable correlated equilibria in restricted (Markov open-loop) strategies for the underlying  $N$ -player games, and we show how to construct approximate  $N$ -player correlated equilibria starting from a correlated solution to the mean field game.

**Keywords and phrases:** Nash equilibrium, correlated equilibrium, mean field game, weak convergence, restricted strategy, exchangeability.

**2020 AMS subject classifications:** 60B10, 91A06, 91A16, 93E20

## 1 Introduction

Correlated equilibria are generalisations of Nash equilibria that allow for correlation between players' strategies. In this paper, we consider correlated equilibria for a simple class of symmetric finite horizon  $N$ -player games and their natural mean field game counterpart as the number of players  $N$  goes to infinity.

*Mean field games* (MFGs, for short), introduced by Huang et al. [2006] and Lasry and Lions [2007], arise as limit systems for certain symmetric stochastic  $N$ -player games with mean field interaction as the number of players  $N$  tends to infinity. Each player interacts with her competitors only via the empirical distribution of their positions so that, when  $N \rightarrow \infty$ , one expects the empirical distribution to converge to the law of the

---

\*Department of Statistics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom.

†Department of Mathematics “Tullio Levi-Civita,” University of Padua, via Trieste 63, 35121 Padova, Italy.

“representative player” (Law of Large Numbers or Propagation of Chaos). In the limiting MFG, the “representative player” reacts optimally to the behavior of the population, which in turn should arise at equilibrium by aggregation of all identical players’ best responses. For a thorough treatment of MFG theory from a probabilistic perspective we refer to the two-volume book by Carmona and Delarue [2018].

A rigorous connection between MFGs and the underlying  $N$ -player games can be established in two directions: constructing approximate Nash equilibria for  $N$ -player games starting from a solution to the MFG (for instance, Huang et al. [2006], Carmona and Delarue [2013], Gomes et al. [2013], to name a few), or by showing convergence of approximate  $N$ -player Nash equilibria to solutions of the MFG, as  $N \rightarrow \infty$ . Crucial, especially in the second direction, is the choice of admissible strategies in the definition of  $N$ -player Nash equilibria. Particularly difficult is the question of convergence in the non-stationary case when Nash equilibria are considered in closed-loop strategies (Markov feedback strategies with full state information). A breakthrough in this direction was made in Cardaliaguet et al. [2019], where convergence of Nash equilibria is established through the so-called master equation provided the latter is well-posed, an assumption that implies uniqueness of MFG solutions. More recently, in Lacker [2020], a general convergence result was proved in the non-degenerate diffusion setting, but to weak solutions of the MFG. For weak MFG solutions, the limiting flow of measures can be stochastic even without common noise. An important question, converse to the convergence result, is whether all weak MFG solutions can be obtained as limits of convergent closed-loop  $N$ -player approximate Nash equilibria. The analysis performed in Lacker [2020, Sect. 7] seems to suggest that this is not always possible. We believe that a way to have a characterization of all MFG solutions as limits of approximate Nash equilibria in  $N$ -player games is to consider a more general concept of solution, such as correlated equilibria.

*Correlated equilibria* were introduced in two seminal papers by Robert Aumann [Aumann, 1974, 1987] for many-player games. Aumann’s main idea can be explained as follows: a *mediator* or *correlation device* randomly selects a strategy profile according to some publicly known distribution, then recommends to each player *in private* a strategy according to the profile. A probability distribution on the space of strategy profiles is a correlated equilibrium (CE, for short) if no player has an incentive to unilaterally deviate from the mediator’s recommendation. In case the mediator uses a product probability distribution, we are back to a Nash equilibrium in mixed strategies. A classical concrete example of a mediator is that of a traffic light; see, for instance, Section 13.1.4 in [Roughgarden, 2016] for more details. Moreover, we notice that the notion of CE admits other equivalent interpretations than that of a mediator. For instance, in [Bárány, 1992] it is shown that a CE of a non-cooperative  $N$ -person game ( $N \geq 4$ ) coincides with a Nash equilibrium of an extended game where the players are allowed to communicate before the original game starts.

Originally introduced in the context of static games with complete information by Aumann, the new notion of CE gave rise to a huge literature in game theory as well as economics along many directions. We refer to the survey by Forges [2012] on several aspects of the more general notion of communication equilibrium and extensions of CE to dynamic games, possibly stochastic and with incomplete information. More specifically on CE in stochastic games whose framework is close to ours, an extensive study has been performed in [Solan, 2000, 2001, Solan and Vieille, 2002].

The important role the concept of CE plays in game theory and economics can be explained by its many appealing properties, as compared to Nash equilibria. For instance, higher equilibrium payoffs can be reached, possibly outside the convex hull of Nash equilibrium payoffs. The computational complexity of CE is generally lower than for Nash equilibria [see Gilboa and Zemel, 1989]. In the evolutionary game theory literature, it has been proved that if all players follow natural learning procedures then the empirical distribution of their actions converges to CE distributions [for instance, Hart, 2005]. Moreover, given their interpretation in terms of mediator's recommendations, correlated equilibria can be seen as intermediate configurations between the two extreme cases of decentralized solutions such as Nash equilibria on the one hand, and centrally planned optimal solutions that are forced on the players on the other hand.

Here, we consider correlated equilibria for a simple class of symmetric finite horizon  $N$ -player games and their natural MFG counterpart as  $N \rightarrow \infty$ . In the  $N$ -player setting, the state variables evolve in discrete time, both state space and the set of control actions are finite and, more importantly, the players are allowed to use only restricted strategies, that is, feedback strategies that depend only on time and the corresponding individual state variable. We believe that further extensions of our results to games in continuous time and with common noise are possible. However, they are postponed to future research as they will most probably require different techniques.

Within this framework, we propose a notion of *correlated MFG solution*, defined as a probability distribution over the space of all pairs of strategies and flows of measures such that (i) the "representative player" has no incentive to deviate from the mediator's recommendation; (ii) the flow of measures at any time  $t$  equals the marginal law of the state variable at time  $t$  conditioned on the  $\sigma$ -algebra generated by the whole flow of measures up to the terminal time. The main contribution of the paper consists in justifying the definition of correlated MFG solutions in the following two ways:

- We prove that any sequence of symmetric approximate correlated equilibria in restricted strategies for the  $N$ -player games subsequentially converges towards some correlated MFG solution according to the definition above (forward approximation).
- We also prove the converse backward approximation, that is, any correlated MFG solution arises as limit of symmetric approximate CE in the  $N$ -player games as

$N \rightarrow \infty$  or, in other terms, any correlated MFG solution can be implemented in a natural way by some mediator willing to recommend strategies to the players.

Both approximation results will be proved using a purely probabilistic approach heavily relying on the theory of weak convergence of probability measures as well as coupling arguments.

The rest of the paper is structured as follows. In Section 2, we introduce the notation and basic elements for the objects of our study. In Section 3, we describe the underlying  $N$ -player games and give the definition of (approximate) correlated equilibrium in restricted strategies. Moreover, we also prove that  $N$ -player correlated equilibria exist in the class of symmetric profile distributions. Section 4 is dedicated to the mean field limit model. There, we give our definition of correlated MFG solution. **An example of a correlated MFG with explicit solutions is provided in Section 5.** In Section 6, we show that symmetric  $N$ -player correlated equilibria concentrate, in the limit as  $N \rightarrow \infty$ , on correlated MFG solutions, while Section 7 contains the converse result, that is, any correlated MFG solution arises as a limit of symmetric approximate  $N$ -player correlated equilibria as  $N \rightarrow \infty$ . In Appendix A, we collect some auxiliary results.

## 2 Preliminaries

For a Polish space  $\mathcal{S}$ , we denote by  $\mathcal{P}(\mathcal{S})$  the space of probability measures on  $\mathcal{B}(\mathcal{S})$ , the Borel sets of  $\mathcal{S}$ , and endow  $\mathcal{P}(\mathcal{S})$  with the topology of weak convergence of measures. Many of the spaces of interest here are simply finite sets. We endow a finite set with the discrete topology, which makes it a Polish space (a compatible metric being the discrete metric).

If  $\mathcal{S}$  is finite, then a metric on  $\mathcal{P}(\mathcal{S})$  compatible with the weak convergence topology is given by the following  $L^1$ -distance, which we indicate by  $\mathbf{dist}$  when the underlying space is clear from the context:

$$\mathbf{dist}(m, \tilde{m}) \doteq \frac{1}{2} \sum_{x \in \mathcal{S}} |m(x) - \tilde{m}(x)|, \quad m, \tilde{m} \in \mathcal{P}(\mathcal{S}).$$

Notice that weak convergence and convergence in total variation coincide for probability measures over a finite set. If  $m, \tilde{m}$  are empirical measures of the same size, that is, if  $m = \frac{1}{N} \sum_{i=1}^N \delta_{s_i}$ ,  $\tilde{m} = \frac{1}{N} \sum_{i=1}^N \delta_{\tilde{s}_i}$  for some  $s_i, \tilde{s}_i \in \mathcal{S}$ ,  $i \in \{1, \dots, N\}$ , then

$$(2.1) \quad \mathbf{dist}(m, \tilde{m}) \leq \min_{\sigma \text{ permutation of } \{1, \dots, N\}} \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{s_i \neq \tilde{s}_{\sigma(i)}} \leq \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{s_i \neq \tilde{s}_i}.$$

We consider symmetric dynamic games in discrete time over a finite time horizon with individual state and action spaces given by finite sets. Admissible strategies will

have Markov feedback form but with information restricted to player's individual states (sometimes called "Markov open-loop"). To fix the notation, we choose

- $T \in \mathbb{N}$ , representing the finite time horizon (with initial time zero);
- non-empty finite sets  $\mathcal{X}$  and  $\Gamma$ , the set of individual states and control actions, respectively;
- a measurable function  $\Psi: \{0, \dots, T-1\} \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma \times \mathcal{Z} \rightarrow \mathcal{X}$ , the system function, determining the one-step individual state dynamics, where  $\mathcal{Z} \doteq [0, 1]$  is the space of noise states;
- a bounded measurable function  $f: \{0, \dots, T-1\} \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma \rightarrow \mathbb{R}$ , representing the running costs;
- a bounded measurable function  $F: \mathcal{X} \times \mathcal{P}(\mathcal{X}) \rightarrow \mathbb{R}$ , representing the terminal costs.

Denote by  $\nu$  the uniform distribution on the Borel sets of  $\mathcal{Z} = [0, 1]$ ;  $\nu$  will be the common distribution of the random variables representing the idiosyncratic noise.

Let  $\mathcal{R}$  denote the set of Markov feedback strategies over players' own states (*restricted strategies*):

$$\mathcal{R} \doteq \{\varphi: \{0, \dots, T-1\} \times \mathcal{X} \rightarrow \Gamma\}.$$

Notice that  $\mathcal{R}$  is a finite set; it will hence be endowed with the discrete topology. Let  $\mathcal{U}$  denote the set of mappings from  $\mathcal{R}$  to  $\mathcal{R}$ :

$$\mathcal{U} \doteq \{u: \mathcal{R} \rightarrow \mathcal{R}\}.$$

Since  $\mathcal{R}$  is a finite set,  $\mathcal{U}$  is finite, too; it will therefore be endowed with the discrete topology. Any element of  $\mathcal{U}$ , that is, any function  $u: \mathcal{R} \rightarrow \mathcal{R}$  (which is automatically measurable) will be referred to as a *strategy modification*.

For the  $N$ -player game, we have to consider probability measures on strategy vectors (or strategy profiles). Any such probability measure, that is, any element of  $\mathcal{P}(\mathcal{R}^N)$ , will be called a *correlated profile*. For the mean field game, we will consider probability measures on individual strategies times flows of state distributions. Any such probability measure, that is, any element of  $\mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$ , will be called a *correlated flow*.

### 3 The $N$ -player games

Fix  $N \in \mathbb{N}$ . Choose  $\mathbf{m}^N \in \mathcal{P}(\mathcal{X}^N)$ , the joint distribution of the players' states at time zero; for instance,  $\mathbf{m}^N = \otimes^N \mathbf{m}_0$  for some  $\mathbf{m}_0 \in \mathcal{P}(\mathcal{X})$ . Let  $\gamma^N \in \mathcal{P}(\mathcal{R}^N)$  be an  $N$ -player correlated profile, and let  $u \in \mathcal{U}$  be a strategy modification.

A tuple  $((\Omega, \mathcal{F}, \mathbf{P}), \Phi_1^N, \dots, \Phi_N^N, X_1^N(\cdot), \dots, X_N^N(\cdot), \xi_1^N(\cdot), \dots, \xi_N^N(\cdot))$  is called a *realization* of the triple  $(\mathbf{m}^N, \gamma^N, u)$  for player  $i$  if  $\Phi_1^N, \dots, \Phi_N^N$  are  $\mathcal{R}$ -valued random variables,  $X_j^N(t)$ ,  $j \in \{1, \dots, N\}$ ,  $t \in \{0, \dots, T\}$ ,  $\mathcal{X}$ -valued random variables, and  $\xi_j^N(t)$ ,  $j \in \{1, \dots, N\}$ ,  $t \in \{1, \dots, T\}$ , are  $\mathcal{Z}$ -valued random variables all defined on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  such that

- (i)  $\mathbf{P} \circ (X_1^N(0), \dots, X_N^N(0))^{-1} = \mathbf{m}^N$ ;
- (ii)  $\mathbf{P} \circ (\Phi_1^N, \dots, \Phi_N^N)^{-1} = \gamma^N$ ;
- (iii)  $\xi_j^N(t)$ ,  $j \in \{1, \dots, N\}$ ,  $t \in \{1, \dots, T\}$ , are independent and identically distributed (i.i.d.) with common distribution  $\mathbf{P} \circ (\xi_j^N(t))^{-1} = \nu$ ;
- (iv)  $(\xi_j^N(t))_{j \in \{1, \dots, N\}, t \in \{1, \dots, T\}}$ ,  $(X_j^N(0))_{j \in \{1, \dots, N\}}$ , and  $(\Phi_j^N)_{j \in \{1, \dots, N\}}$  are independent as random variables with values in  $\mathcal{Z}^{N \cdot T}$ ,  $\mathcal{X}^N$ , and  $\mathcal{R}^N$ , respectively;
- (v)  $\mathbf{P}$ -almost surely, for every  $t \in \{0, \dots, T-1\}$ ,

$$\begin{aligned} X_i^N(t+1) &= \Psi(t, X_i^N(t), \mu_i^N(t), u \circ \Phi_i^N(t, X_i^N(t)), \xi_i^N(t+1)), \\ X_j^N(t+1) &= \Psi(t, X_j^N(t), \mu_j^N(t), \Phi_j^N(t, X_j^N(t)), \xi_j^N(t+1)), \quad j \neq i, \end{aligned}$$

where  $\mu_i^N(t)$  is the empirical measure of the states at time  $t$  of all players except player  $i$ :

$$\mu_l^N(t) \doteq \frac{1}{N-1} \sum_{j \neq l} \delta_{X_j^N(t)}, \quad l \in \{1, \dots, N\}.$$

The difference in the dynamics of  $X_i^N$  and  $X_j^N$ ,  $j \neq i$ , above is that the former includes the strategy modification  $u$ . Any realization

$$((\Omega, \mathcal{F}, \mathbf{P}), \Phi_1^N, \dots, \Phi_N^N, X_1^N(\cdot), \dots, X_N^N(\cdot), \xi_1^N(\cdot), \dots, \xi_N^N(\cdot))$$

of the triple  $(\mathbf{m}^N, \gamma^N, u)$  for player  $i$  can be interpreted in the following way. The random variables  $X_1^N(0), \dots, X_N^N(0)$  represent the initial states of players 1 through  $N$ , and their joint distribution is given by  $\mathbf{m}^N$ . The random variable  $\Phi_i^N$  represents the recommendation (or signal) the mediator sends to player  $i$  before the game starts. While the joint distribution of  $\Phi_1^N, \dots, \Phi_N^N$ , which is equal to  $\gamma^N$ , is common knowledge, no player can directly see the recommendations received by the others. This feature is made precise in the way the state dynamics are formulated in (v): Player  $i$ , the player who might deviate, chooses a strategy by modifying the recommendation  $\Phi_i^N$  through the application of a mapping  $u: \mathcal{R} \rightarrow \mathcal{R}$ , while the other players follow the recommendation they receive from the mediator. Player  $i$  thus uses the random strategy  $u \circ \Phi_i^N$  instead of simply  $\Phi_i^N$ . Clearly,  $u \circ \Phi_i^N$  is  $\sigma(\Phi_i^N)$ -measurable.

*Remark 3.1.* The independence assumption in (iv) is crucial. Clearly, the vector  $(\xi_j^N(t))$  of noise variables and the vector  $(X_j^N(0))$  of initial states have to be independent. But we also require them to be independent of the vector  $(\Phi_j^N)$  of recommendation variables. This makes precise the idea that the mediator gives recommendations to the players before the game starts. Recall that what the mediator suggests are feedback strategies. If player  $j$  accepts the mediator's recommendation, then, given a scenario  $\omega \in \Omega$ , he will use the feedback strategy  $\Phi_j^N(\omega)$ . In view of the dynamics according to (v), he will therefore select, at any time  $t$ , the control action  $\Phi_j^N(\omega)(t, X_j^N(t, \omega))$ . The control action at time  $t$  is thus in general not independent of the noise variables up to time  $t$ , nor is it independent of the initial states. An analogous observation holds for player  $i$ , who might modify the mediator's recommendation.

The costs for player  $i$  associated with initial distribution  $\mathbf{m}^N$ , correlated profile  $\gamma^N$ , and a strategy modification  $u$  are given by

$$J_i^N(\mathbf{m}^N, \gamma^N, u) \doteq \mathbf{E} \left[ \sum_{t=0}^{T-1} f(t, X_i^N(t), \mu_i^N(t), u \circ \Phi_i^N(t, X_i^N(t))) + F(X_i^N(T), \mu_i^N(T)) \right],$$

where the expected value on the right-hand side above is computed with respect to any realization of the triple  $(\mathbf{m}^N, \gamma^N, u)$  for player  $i$ . Thanks to the independence assumption (iv), the costs are well defined in that they do not depend on the choice of the realization.

**Definition 3.1.** Let  $\varepsilon \geq 0$ . A correlated profile  $\gamma^N \in \mathcal{P}(\mathcal{R}^N)$  is called an  $\varepsilon$ -*correlated equilibrium* in restricted strategies with initial distribution  $\mathbf{m}^N$  if for every  $i \in \{1, \dots, N\}$ , every strategy modification  $u \in \mathcal{U}$ ,

$$J_i^N(\mathbf{m}^N, \gamma^N, \text{Id}) \leq J_i^N(\mathbf{m}^N, \gamma^N, u) + \varepsilon.$$

When  $\varepsilon = 0$ , we say that  $\gamma^N$  is a correlated equilibrium in restricted strategies.

An  $\varepsilon$ -correlated equilibrium is called *symmetric* if it is symmetric as a probability measure on  $\mathcal{R}^N$  (i.e. invariant under permutations of the components).

*Remark 3.2.* Nash equilibria are particular cases of correlated equilibria. According to Definition 3.1, a Nash equilibrium in mixed strategies corresponds to a correlated profile  $\gamma^N$  that has product form, while a Nash equilibrium in pure strategies corresponds to a correlated profile which is the product of Dirac measures concentrated in the strategies of the Nash profile.

Next, we prove that there always exists a symmetric correlated equilibrium for the  $N$ -player game. To this end, instead of relying on the existence of symmetric Nash equilibria, which would hold in this setting, we rather follow a more direct approach by applying the existence result in Hart and Schmeidler [1989] through a simple symmetrization argument.

**Proposition 3.1.** *Let  $\mathbf{m}^N \in \mathcal{P}(\mathcal{X}^N)$  be symmetric. Then there exists a symmetric correlated equilibrium with initial distribution  $\mathbf{m}^N$ .*

*Proof.* Applying Hart and Schmeidler [1989, Theorem 1] to our setting, we obtain the existence of a correlated equilibrium  $\gamma \in \mathcal{P}(\mathcal{R}^N)$  for the  $N$ -player game described above, i.e. for all  $i \in \{1, \dots, N\}$  and all  $u \in \mathcal{U}$  we have

$$(3.1) \quad \sum_{\varphi \in \mathcal{R}^N} \gamma(\varphi) (J_i^N(\mathbf{m}^N, \delta_\varphi, u) - J_i^N(\mathbf{m}^N, \delta_\varphi, \text{Id})) \geq 0.$$

Since  $\gamma$  is not necessarily symmetric, we symmetrize it by defining

$$\tilde{\gamma}(\varphi^1, \dots, \varphi^N) := \frac{1}{N!} \sum_{\sigma} \gamma(\varphi^{\sigma(1)}, \dots, \varphi^{\sigma(N)}),$$

where  $\sigma$  varies in the set of all permutations of  $\{1, \dots, N\}$ . We check that also  $\tilde{\gamma}$  is a CE for the  $N$ -player game. Letting  $\varphi^\sigma = (\varphi^{\sigma(1)}, \dots, \varphi^{\sigma(N)})$ , we can write

$$\begin{aligned} & \sum_{\varphi \in \mathcal{R}^N} \tilde{\gamma}(\varphi) (J_i^N(\mathbf{m}^N, \delta_\varphi, u) - J_i^N(\mathbf{m}^N, \delta_\varphi, \text{Id})) \\ &= \frac{1}{N!} \sum_{\sigma} \sum_{\varphi \in \mathcal{R}^N} \gamma(\varphi^{\sigma(1)}, \dots, \varphi^{\sigma(N)}) (J_i^N(\mathbf{m}^N, \delta_\varphi, u) - J_i^N(\mathbf{m}^N, \delta_\varphi, \text{Id})) \\ &= \frac{1}{N!} \sum_{\sigma} \sum_{\varphi \in \mathcal{R}^N} \gamma(\varphi^\sigma) (J_{\sigma(i)}^N(\mathbf{m}^N, \delta_{\varphi^\sigma}, u) - J_{\sigma(i)}^N(\mathbf{m}^N, \delta_{\varphi^\sigma}, \text{Id})) \geq 0, \end{aligned}$$

where the second equality is due to symmetry and the final inequality follows from (3.1).  $\square$

## 4 The correlated mean field game

Choose  $\mathbf{m}_0 \in \mathcal{P}(\mathcal{X})$ , the distribution of the representative player's state at time zero. Let  $\rho \in \mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$  be a correlated flow, and let  $u \in \mathcal{U}$  be a strategy modification.

A tuple  $((\Omega, \mathcal{F}, \mathbf{P}), \Phi, X(\cdot), \mu(\cdot), \xi(\cdot))$  is called a *realization* of the triple  $(\mathbf{m}_0, \rho, u)$  if  $\Phi$  is an  $\mathcal{R}$ -valued random variable,  $X(0), \dots, X(T)$  are  $\mathcal{X}$ -valued random variables,  $\mu(0), \dots, \mu(T)$  are  $\mathcal{P}(\mathcal{X})$ -valued random variables, and  $\xi(1), \dots, \xi(T)$  are  $\mathcal{Z}$ -valued random variables all defined on a common probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  such that

- (i)  $\mathbf{P} \circ (X(0))^{-1} = \mathbf{m}_0$ ;
- (ii)  $\mathbf{P} \circ (\Phi, \mu(0), \dots, \mu(T))^{-1} = \rho$ ;
- (iii)  $\xi(t)$ ,  $t \in \{1, \dots, T\}$ , are i.i.d. with common distribution  $\mathbf{P} \circ (\xi(t))^{-1} = \nu$ ;



(iv)  $\xi(\cdot)$ ,  $X(0)$ , and  $(\Phi, \mu(\cdot))$  are independent as random variables with values in  $\mathcal{Z}^T$ ,  $\mathcal{X}$ , and  $\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1}$ , respectively;

(v)  $\mathbf{P}$ -almost surely, for every  $t \in \{0, \dots, T-1\}$ ,

$$(4.1) \quad X(t+1) = \Psi(t, X(t), \mu(t), u \circ \Phi(t, X(t)), \xi(t+1)).$$

Recalling the heuristic connection between  $N$ -player games and mean field game, we can interpret a realization  $((\Omega, \mathcal{F}, \mathbf{P}), \Phi, X(\cdot), \mu(\cdot), \xi(\cdot))$  of the triple  $(\mathbf{m}_0, \rho, u)$  as follows. The random variable  $\Phi$  represents the recommendation that one representative player receives from the mediator, whereas  $X(\cdot)$  gives the representative player's state sequence, which is recursively determined through Eq. (4.1). There,  $\xi(t)$ ,  $t \in \{1, \dots, T\}$ , are the noise variables, while  $\mu(\cdot)$  represents a stochastic flow of measures.

*Remark 4.1.* The flow of measures  $\mu(\cdot)$  should be thought of as a limit point of the  $N$ -player flows of empirical measures. As such, it will in general be stochastic and not independent of the recommendation variable  $\Phi$ , which in turn should be thought of as a limit point of the recommendation variables for one fixed player, say the first, in the  $N$ -player games. It is therefore necessary to prescribe the joint distribution of  $\Phi$  and  $\mu$ , as done in (ii) through the correlated flow  $\rho$ . Similarly, in (iv), which should be compared to the independence assumption (iv) of the  $N$ -player games, we require independence of  $\xi(\cdot)$ ,  $X(0)$ , and  $(\Phi, \mu(\cdot))$ , not just of  $\xi(\cdot)$ ,  $X(0)$ , and  $\Phi$ . We stress that in general  $\Phi$  and  $\mu$  will not be independent.

The costs for a representative player associated with initial distribution  $\mathbf{m}_0$ , correlated flow  $\rho$ , and a strategy modification  $u: \mathcal{R} \rightarrow \mathcal{R}$  are given by

$$J(\mathbf{m}_0, \rho, u) \doteq \mathbf{E} \left[ \sum_{t=0}^{T-1} f(t, X(t), \mu(t), u \circ \Phi(t, X(t))) + F(X(T), \mu(T)) \right],$$

where the expected value on the right-hand side above is computed with respect to any realization of  $(\mathbf{m}_0, \rho, u)$ . Thanks to (iv), any two realizations of  $(\mathbf{m}_0, \rho, u)$  generate the same expected value. The cost functional  $J$  is thus well defined.

**Definition 4.1.** A correlated flow  $\rho \in \mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$  is called a *correlated solution of the mean field game* in restricted strategies with initial distribution  $\mathbf{m}_0$  if the following two conditions hold:

(i) Optimality: For every strategy modification  $u \in \mathcal{U}$ ,

$$J(\mathbf{m}_0; \rho, \text{Id}) \leq J(\mathbf{m}_0; \rho, u).$$

(ii) Consistency: If  $((\Omega, \mathcal{F}, \mathbf{P}), \Phi, X(\cdot), \mu(\cdot), \xi(\cdot))$  is a realization of the triple  $(\mathbf{m}_0, \rho, \text{Id})$ , then for every  $t \in \{0, \dots, T\}$ ,

$$\mu(t)(\cdot) = \mathbf{P}(X(t) \in \cdot \mid \mathcal{F}^\mu),$$

where  $\mathcal{F}^\mu \doteq \sigma(\mu) = \sigma(\mu(s) : s \in \{0, \dots, T\})$ .

The consistency condition in Definition 4.1 is to be understood in the sense that  $\mu(t)$  is a regular conditional distribution of  $X(t)$  given  $\mathcal{F}^\mu$ . Notice that  $\mathcal{F}^\mu$  is the  $\sigma$ -algebra generated by the entire flow of measures  $\mu$ , up to terminal time  $T$ . Intuitively, the consistency condition can be interpreted as follows: The moderator has an idea of the flow  $\mu$  on the whole time interval, on the basis of which he recommends strategies to the players. If each player follows his recommendations, then that flow  $\mu$  will arise from aggregation of the individual behaviors.

*Remark 4.2.* Definition 4.1 should be compared to the definition of weak MFG solution, more precisely *weak semi-Markov mean field equilibrium*, given in Lacker [2020, Definition 2.5]. An obvious difference lies in the dynamics: While Lacker works with controlled Itô diffusions driven by non-degenerate additive Wiener noise (as here, idiosyncratic, no common noise), here we consider simple discrete time dynamics with finite state and control space. Conceptually more important is the fact that the admissible strategies here are restricted to functions that depend only on time and the player's current state, while Lacker allows for an additional dependence on the flow of measures up to current time. Notice that the flow of measures may be stochastic in both cases. Clearly, there is no mediator or correlation device in Lacker [2020]. If in our situation we take the recommendation variable to be (almost surely) constant, hence with Dirac distribution **concentrated at some feedback strategy**  $\varphi \in \mathcal{R}$ , then the optimality condition in Definition 4.1 above can be seen to be analogous to the optimality condition in Lacker's definition (point (5) there). His consistency condition (point (6) there) is apparently different in that the conditional distribution is taken with respect to the  $\sigma$ -algebra generated by the flow of measures up to current time, not up to terminal time as in our definition. However, if the recommendation variable is (almost surely) constant, or absent as in Lacker's work, then the two ways of conditioning lead to equivalent consistency conditions, thanks to the (semi-)Markov property of the state process. Indeed, when  $\Phi$  is constant,  $X(t)$  and  $\mu$  are conditionally independent given  $\mathcal{F}^\mu(t) = \sigma(\mu(s) : s \in \{0, \dots, t\})$ , for all  $t \in \{0, \dots, T\}$ . Therefore, in this case, the property  $\mu(t)(\cdot) = \mathbf{P}(X(t) \in \cdot \mid \mathcal{F}^\mu(t))$ ,  $t \in \{0, \dots, T\}$ , implies the consistency condition (ii) above. In this way, and under the simplifying assumptions made here, one can interpret weak MFG solutions as a special case of correlated solutions.

## 5 Example of a correlated mean field game

In this section, we give an example of a two-state mean field game possessing correlated solutions with non-deterministic flow of measures.

Set  $\mathcal{X} \doteq \{-1, 1\}$ ,  $\Gamma \doteq \{0, 1\}$ , and  $T \doteq 2$ . Define the system function  $\Psi$  by

$$\Psi(t, x, m, \gamma, z) \doteq \Psi(x, \gamma, z) \doteq \begin{cases} x & \text{if } \gamma = 0 \text{ and } z \in [0, \frac{1}{2}], \\ -x & \text{if } \gamma = 0 \text{ and } z \in (\frac{1}{2}, 1], \\ x & \text{if } \gamma = 1 \text{ and } z \in [0, \frac{3}{4}], \\ -x & \text{if } \gamma = 1 \text{ and } z \in (\frac{3}{4}, 1], \end{cases}$$

Notice that  $\Psi$  is time-homogeneous and independent of the measure variable, which justifies our slight abuse of notation. According to  $\Psi$ , when moving one step in time, the player's state switches with probability  $1/2$  if action  $\gamma = 0$  is chosen, while it will change only with probability  $1/4$  if action  $\gamma = 1$  is played. Choose running costs  $f$  and terminal costs  $F$  according to

$$f(t, x, m, \gamma) \doteq \begin{cases} c_0 \cdot \gamma & \text{if } t = 0, \\ c_1 \cdot \gamma - x \cdot \mathbf{M}(m) & \text{if } t = 1, \end{cases} \quad F(x, m) \doteq -x \cdot \mathbf{M}(m),$$

where  $c_0, c_1 > 0$  will be chosen below and  $\mathbf{M}(m) \doteq m(\{1\}) - m(\{-1\})$  is the mean of a probability measure  $m \in \mathcal{P}(\{-1, 1\})$ .

Define strategies  $\varphi_+, \hat{\varphi}_+, \varphi_-, \hat{\varphi}_-, \varphi_o \in \mathcal{R}$  according to

$$\begin{aligned} \varphi_+(t, x) &\doteq \begin{cases} 0 & \text{if } x = -1 \\ 1 & \text{if } x = 1, \end{cases} & \hat{\varphi}_+(t, x) &\doteq \begin{cases} 0 & \text{if } x = -1 \text{ or } t = 1, \\ 1 & \text{if } x = 1 \text{ and } t = 0, \end{cases} \\ \varphi_-(t, x) &\doteq \begin{cases} 0 & \text{if } x = 1, \\ 1 & \text{if } x = -1, \end{cases} & \hat{\varphi}_-(t, x) &\doteq \begin{cases} 0 & \text{if } x = 1 \text{ or } t = 1, \\ 1 & \text{if } x = -1 \text{ and } t = 0, \end{cases} \\ \varphi_o(t, x) &\doteq 0, & t &\in \{0, 1\}, x \in \{-1, 1\}. \end{aligned}$$

Strategy  $\varphi_+$  has the effect of maximizing the probability of being in state 1 at times 1 and 2, while  $\hat{\varphi}_+$  only maximizes that probability at time 1. The effect of  $\varphi_-$ ,  $\hat{\varphi}_-$  is analogous, with the roles of states 1 and  $-1$  inverted. Under strategy  $\varphi_o$  (“do nothing”), on the other hand, the two states will have equal probability at both time 1 and time 2, independently of the initial state.

Choose  $\beta_1, \dots, \beta_4 > 0$  such that

$$(5.1) \quad \beta_1 + \beta_2 + \beta_3 + \beta_4 = \frac{1}{2}, \quad \frac{5\beta_1 + 4\beta_2}{8(\beta_1 + \beta_2)} = \frac{5\beta_3 + 4\beta_4}{8(\beta_3 + \beta_4)}.$$

An obvious choice satisfying (5.1) is to set  $\beta_i \doteq 1/8$  for all  $i \in \{1, \dots, 4\}$ . Define

distributions in  $\mathcal{P}(\{-1, 1\})$  according to

$$\begin{aligned} m_0 &\doteq \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}, \\ m_1^+ &\doteq \frac{5\beta_1 + 4\beta_2}{8(\beta_1 + \beta_2)}\delta_1 + \frac{3\beta_1 + 4\beta_2}{8(\beta_1 + \beta_2)}\delta_{-1}, & m_2^+ &\doteq \frac{21\beta_1 + 16\beta_2}{32(\beta_1 + \beta_2)}\delta_1 + \frac{11\beta_1 + 16\beta_2}{32(\beta_1 + \beta_2)}\delta_{-1}, \\ m_1^- &\doteq \frac{3\beta_1 + 4\beta_2}{8(\beta_1 + \beta_2)}\delta_1 + \frac{5\beta_1 + 4\beta_2}{8(\beta_1 + \beta_2)}\delta_{-1}, & m_2^- &\doteq \frac{11\beta_1 + 16\beta_2}{32(\beta_1 + \beta_2)}\delta_1 + \frac{21\beta_1 + 16\beta_2}{32(\beta_1 + \beta_2)}\delta_{-1}. \end{aligned}$$

In terms of the above distributions and strategies, define a correlated flow  $\rho$  by

$$\begin{aligned} \rho &\doteq \beta_1\delta_{(\varphi_+, (m_0, m_1^+, m_2^+))} + \beta_2\delta_{(\varphi_o, (m_0, m_1^+, m_2^+))} + \beta_3\delta_{(\hat{\varphi}_+, (m_0, m_1^+, m_0))} + \beta_4\delta_{(\varphi_o, (m_0, m_1^+, m_0))} \\ &+ \beta_1\delta_{(\varphi_-, (m_0, m_1^-, m_2^-))} + \beta_2\delta_{(\varphi_o, (m_0, m_1^-, m_2^-))} + \beta_3\delta_{(\hat{\varphi}_-, (m_0, m_1^-, m_0))} + \beta_4\delta_{(\varphi_o, (m_0, m_1^-, m_0))}. \end{aligned}$$

In Figure 1, we illustrate the four measure trajectories that have strictly positive probability according to  $\rho$ , namely  $(m_0, m_1^+, m_2^+)$ ,  $(m_0, m_1^+, m_0)$ ,  $(m_0, m_1^-, m_2^-)$ , and  $(m_0, m_1^-, m_0)$ . In analogy with the classical definition of correlated equilibria, we suppose that  $\rho$ , which here gives the joint distribution of the mediator's recommendations and the flow of measures, is common knowledge. Thus, if the representative player receives the recommendation to play  $\varphi_+$ , she can infer that with probability one the flow of measures will be concentrated at the measure trajectory  $(m_0, m_1^+, m_2^+)$ . Similarly, upon receiving recommendation  $\hat{\varphi}_+$ , she can deduce that the flow of measures will be concentrated at  $(m_0, m_1^+, m_0)$ , and analogously for  $\varphi_-$ ,  $\hat{\varphi}_-$ . If, on the other hand, the mediator recommends to play  $\varphi_o$ , then all four measure trajectories are possible. In fact,  $(m_0, m_1^+, m_2^+)$  and  $(m_0, m_1^-, m_2^-)$  will both have (conditional) probability equal to  $\frac{\beta_2}{2(\beta_2 + \beta_4)}$ , while  $(m_0, m_1^+, m_0)$  and  $(m_0, m_1^-, m_0)$  will both have (conditional) probability equal to  $\frac{\beta_4}{2(\beta_2 + \beta_4)}$ ; also see Remark 5.1 below.

We are going to show that  $\rho$  is a correlated solution of the mean field game in the sense of Definition 4.1 provided  $c_0, c_1 > 0$  are taken sufficiently small. To this end, let  $((\Omega, \mathcal{F}, \mathbf{P}), \Phi, X(\cdot), \mu(\cdot), \xi(\cdot))$  be a realization of the triple  $(\mathbf{m}_0, \rho, \text{Id})$ . Then, for every  $t \in \{0, 1, 2\}$ ,  $\mathbf{P}$ -almost surely,

$$\mathbf{P}(X(t) \in \cdot \mid \mathcal{F}^\mu) = \begin{cases} \frac{\beta_1}{\beta_1 + \beta_2} \mathbf{P} \circ (X_+(t))^{-1} + \frac{\beta_2}{\beta_1 + \beta_2} \mathbf{P} \circ (X_o(t))^{-1} & \text{if } \mu = (m_0, m_1^+, m_2^+), \\ \frac{\beta_3}{\beta_3 + \beta_4} \mathbf{P} \circ (\hat{X}_+(t))^{-1} + \frac{\beta_4}{\beta_3 + \beta_4} \mathbf{P} \circ (X_o(t))^{-1} & \text{if } \mu = (m_0, m_1^+, m_0), \\ \frac{\beta_1}{\beta_1 + \beta_2} \mathbf{P} \circ (X_-(t))^{-1} + \frac{\beta_2}{\beta_1 + \beta_2} \mathbf{P} \circ (X_o(t))^{-1} & \text{if } \mu = (m_0, m_1^-, m_2^-), \\ \frac{\beta_3}{\beta_3 + \beta_4} \mathbf{P} \circ (\hat{X}_-(t))^{-1} + \frac{\beta_4}{\beta_3 + \beta_4} \mathbf{P} \circ (X_o(t))^{-1} & \text{if } \mu = (m_0, m_1^-, m_0), \end{cases}$$

where  $X_o, X_+, \hat{X}_+, X_-, \hat{X}_-$  are the state processes that result from applying feedback strategies  $\varphi_o, \varphi_+, \hat{\varphi}_+, \varphi_-$ , and  $\hat{\varphi}_-$ , respectively, with initial distribution  $m_0$ . Notice that these processes can be recursively defined through  $\Psi$  on the given probability space

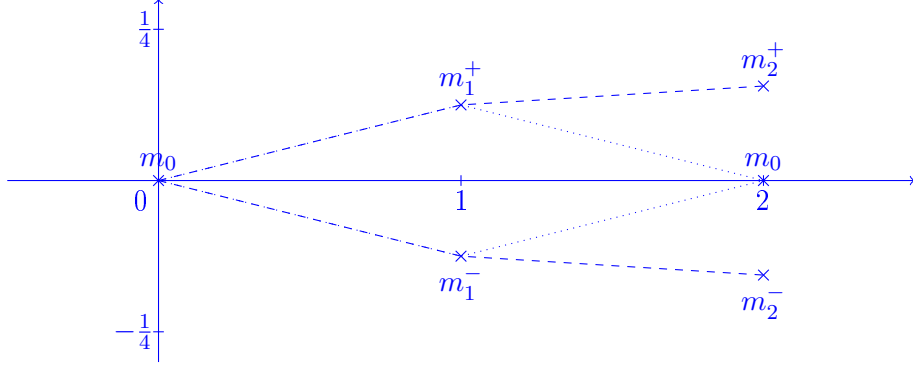


Figure 1: Measure trajectories with strictly positive probability according to the correlated flow  $\rho$ . Elements of  $\mathcal{P}(\{-1, 1\})$  are identified with their mean, measured along the vertical axis. Values have been computed for the case  $\beta_1 = \beta_2 = \beta_3 = \beta_4 = 1/8$ . Time is measured along the horizontal axis. Only times 0, 1, 2 are relevant for the model, though for the sake of illustration trajectories are represented as if time was continuous.

$(\Omega, \mathcal{F}, \mathbf{P})$  in terms of the noise variables  $\xi(1), \xi(2)$ , and the initial state  $X(0)$ . For instance,  $X_+(\cdot)$  is recursively determined by setting

$$X_+(0) \doteq X(0), \quad X_+(t+1) = \Psi(X_+(t), \varphi_+(t, X_+(t)), \xi(t+1)), \quad t \in \{0, 1\},$$

where  $\Psi$  is seen as a function of state, control, and noise only, according to its definition above. Recalling that  $m_0 = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$ , we find:

$$\begin{aligned} \mathbf{P} \circ (X_o(t))^{-1} &= m_0, \quad t \in \{0, 1, 2\}, \\ \mathbf{P} \circ (X_+(0))^{-1} &= m_0, \quad \mathbf{P} \circ (X_+(1))^{-1} = \frac{5}{8}\delta_1 + \frac{3}{8}\delta_{-1}, \quad \mathbf{P} \circ (X_+(2))^{-1} = \frac{21}{32}\delta_1 + \frac{11}{32}\delta_{-1}, \\ \mathbf{P} \circ (\hat{X}_+(0))^{-1} &= m_0, \quad \mathbf{P} \circ (\hat{X}_+(1))^{-1} = \frac{5}{8}\delta_1 + \frac{3}{8}\delta_{-1}, \quad \mathbf{P} \circ (\hat{X}_+(2))^{-1} = m_0, \\ \mathbf{P} \circ (X_-(0))^{-1} &= m_0, \quad \mathbf{P} \circ (X_-(1))^{-1} = \frac{3}{8}\delta_1 + \frac{5}{8}\delta_{-1}, \quad \mathbf{P} \circ (X_-(2))^{-1} = \frac{11}{32}\delta_1 + \frac{21}{32}\delta_{-1}, \\ \mathbf{P} \circ (\hat{X}_-(0))^{-1} &= m_0, \quad \mathbf{P} \circ (\hat{X}_-(1))^{-1} = \frac{3}{8}\delta_1 + \frac{5}{8}\delta_{-1}, \quad \mathbf{P} \circ (\hat{X}_-(2))^{-1} = m_0. \end{aligned}$$

It follows that for  $\mathbf{P}$ -almost all  $\omega \in \Omega$ ,

$$\begin{aligned} \mathbf{P}(X(0) \in \cdot \mid \mathcal{F}^\mu)(\omega) &= m_0, \\ \mathbf{P}(X(1) \in \cdot \mid \mathcal{F}^\mu)(\omega) &= \begin{cases} \frac{5\beta_1+4\beta_2}{8(\beta_1+\beta_2)}\delta_1 + \frac{3\beta_1+4\beta_2}{8(\beta_1+\beta_2)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^+, m_2^+), \\ \frac{5\beta_3+4\beta_4}{8(\beta_3+\beta_4)}\delta_1 + \frac{3\beta_3+4\beta_4}{8(\beta_3+\beta_4)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^+, m_0), \\ \frac{3\beta_1+4\beta_2}{8(\beta_1+\beta_2)}\delta_1 + \frac{5\beta_1+4\beta_2}{8(\beta_1+\beta_2)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^-, m_2^-), \\ \frac{3\beta_3+4\beta_4}{8(\beta_3+\beta_4)}\delta_1 + \frac{5\beta_3+4\beta_4}{8(\beta_3+\beta_4)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^-, m_0), \end{cases} \end{aligned}$$

$$\mathbf{P}(X(2) \in \cdot \mid \mathcal{F}^\mu)(\omega) = \begin{cases} \frac{21\beta_1+16\beta_2}{32(\beta_1+\beta_2)}\delta_1 + \frac{11\beta_1+16\beta_2}{32(\beta_1+\beta_2)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^+, m_2^+), \\ \frac{\beta_3+\beta_4}{2(\beta_3+\beta_4)}\delta_1 + \frac{\beta_3+\beta_4}{2(\beta_3+\beta_4)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^+, m_0), \\ \frac{11\beta_1+16\beta_2}{32(\beta_1+\beta_2)}\delta_1 + \frac{21\beta_1+16\beta_2}{32(\beta_1+\beta_2)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^-, m_2^-), \\ \frac{\beta_3+\beta_4}{2(\beta_3+\beta_4)}\delta_1 + \frac{\beta_3+\beta_4}{2(\beta_3+\beta_4)}\delta_{-1} & \text{if } \mu_\omega = (m_0, m_1^-, m_0). \end{cases}$$

Thanks to the choice of  $\beta_i$ ,  $i \in \{1, \dots, 4\}$ , according to (5.1), we find that the consistency condition of Definition 4.1 is satisfied.

As to optimality, let  $u \in \mathcal{U}$  be any strategy modification. Since  $\Phi$  takes values in  $\{\varphi_o, \varphi_+, \hat{\varphi}_+, \varphi_-, \hat{\varphi}_-\}$  with probability one, we set

$$\psi_o \doteq u(\varphi_o), \quad \psi_+ \doteq u(\varphi_+), \quad \hat{\psi}_+ \doteq u(\hat{\varphi}_+), \quad \psi_- \doteq u(\varphi_-), \quad \hat{\psi}_- \doteq u(\hat{\varphi}_-).$$

Let  $Y_o, Y_+, \hat{Y}_+, Y_-, \hat{Y}_-$  be the corresponding state processes, all starting from  $X(0)$ , hence with initial distribution  $m_0$ . Using that  $\mathbf{M}(m_0) = 0$ , while  $\mathbf{M}(m_t^+) = -\mathbf{M}(m_t^-)$ ,  $t \in \{1, 2\}$ , we obtain

$$\begin{aligned} & J(\mathbf{m}_0; \rho, u) \\ &= \left( c_0 \cdot \mathbf{P}(\psi_+(0, Y_+(0)) = 1) + c_1 \cdot \mathbf{P}(\psi_+(1, Y_+(1)) = 1) + (1 - 2\mathbf{P}(Y_+(1) = 1)) \cdot \mathbf{M}(m_1^+) \right. \\ &\quad \left. + (1 - 2\mathbf{P}(Y_+(2) = 1)) \cdot \mathbf{M}(m_2^+) \right) \cdot \beta_1 \\ &+ \left( c_0 \cdot \mathbf{P}(\hat{\psi}_+(0, \hat{Y}_+(0)) = 1) + c_1 \cdot \mathbf{P}(\hat{\psi}_+(1, \hat{Y}_+(1)) = 1) \right. \\ &\quad \left. + (1 - 2\mathbf{P}(\hat{Y}_+(1) = 1)) \cdot \mathbf{M}(m_1^+) \right) \cdot \beta_3 \\ &+ \left( c_0 \cdot \mathbf{P}(\psi_-(0, Y_-(0)) = 1) + c_1 \cdot \mathbf{P}(\psi_-(1, Y_-(1)) = 1) + (1 - 2\mathbf{P}(Y_-(1) = 1)) \cdot \mathbf{M}(m_1^-) \right. \\ &\quad \left. + (1 - 2\mathbf{P}(Y_-(2) = 1)) \cdot \mathbf{M}(m_2^-) \right) \cdot \beta_1 \\ &+ \left( c_0 \cdot \mathbf{P}(\hat{\psi}_-(0, \hat{Y}_-(0)) = 1) + c_1 \cdot \mathbf{P}(\hat{\psi}_-(1, \hat{Y}_-(1)) = 1) \right. \\ &\quad \left. + (1 - 2\mathbf{P}(\hat{Y}_-(1) = 1)) \cdot \mathbf{M}(m_1^-) \right) \cdot \beta_3 \\ &+ (c_0 \cdot \mathbf{P}(\psi_o(0, Y_o(0)) = 1) + c_1 \cdot \mathbf{P}(\psi_o(1, Y_o(1)) = 1)) \cdot (2\beta_2 + 2\beta_4). \end{aligned}$$

The last line above shows that if  $c_0, c_1 > 0$ , then taking  $\psi_o = \varphi_o$  is optimal on the event  $\{\Phi = \varphi_o\}$  since only in this case  $\mathbf{P}(\psi_o(t, Y_o(t)) = 1) = 0$  for  $t \in \{0, 1\}$ . By symmetry of construction, it remains to show that  $\varphi_+$  is optimal on the event  $\{\Phi = \varphi_+\}$  and  $\hat{\varphi}_+$  on  $\{\Phi = \hat{\varphi}_+\}$  provided  $c_0, c_1 > 0$  are sufficiently small.

In verifying optimality, we will make use of the principle of dynamic programming.

First, to show that  $\varphi_+$  is optimal on the event  $\{\Phi = \varphi_+\}$ , set, for  $x \in \{-1, 1\}$ ,

$$\begin{aligned} V_+(2, x) &\doteq F(x, m_2^+) = -x \cdot \mathbf{M}(m_2^+), \\ V_+(1, x) &\doteq \min_{\gamma \in \{0,1\}} \left\{ c_1 \cdot \gamma - x \cdot \mathbf{M}(m_1^+) + \mathbf{E} [V_+(2, \Psi(x, \gamma, \xi(2)))] \right\}, \\ V_+(0, x) &\doteq \min_{\gamma \in \{0,1\}} \left\{ c_0 \cdot \gamma + \mathbf{E} [V_+(1, \Psi(x, \gamma, \xi(1)))] \right\}. \end{aligned}$$

As we are interested in finding optimal control actions, in defining  $V_+$  we have omitted the weight factor  $\beta_1$ . This corresponds to computing costs with respect to the conditional probability  $\mathbf{P}(\cdot | \Phi = \varphi_+)$ . Notice that  $V_+$  is the value function of the optimal control problem the representative player faces when being told to play  $\varphi_+$  by the mediator. We have

$$V_+(2, 1) = -\frac{5\beta_1}{16(\beta_1 + \beta_2)}, \quad V_+(2, -1) = \frac{5\beta_1}{16(\beta_1 + \beta_2)}$$

by choice of  $m_2^+$ , hence, also recalling  $m_1^+$  and the definition of  $\Psi$ ,

$$\begin{aligned} V_+(1, 1) &= \min \left\{ -\frac{\beta_1}{4(\beta_1 + \beta_2)} + \frac{1}{2}V_+(2, 1) + \frac{1}{2}V_+(2, -1), \right. \\ &\quad \left. c_1 - \frac{\beta_1}{4(\beta_1 + \beta_2)} + \frac{3}{4}V_+(2, 1) + \frac{1}{4}V_+(2, -1) \right\} \\ &= \min \left\{ -\frac{\beta_1}{4(\beta_1 + \beta_2)}, c_1 - \frac{\beta_1}{4(\beta_1 + \beta_2)} - \frac{5\beta_1}{32(\beta_1 + \beta_2)} \right\}, \\ V_+(1, -1) &= \min \left\{ \frac{\beta_1}{4(\beta_1 + \beta_2)} + \frac{1}{2}V_+(2, 1) + \frac{1}{2}V_+(2, -1), \right. \\ &\quad \left. c_1 + \frac{\beta_1}{4(\beta_1 + \beta_2)} + \frac{1}{4}V_+(2, 1) + \frac{3}{4}V_+(2, -1) \right\} \\ &= \min \left\{ \frac{\beta_1}{4(\beta_1 + \beta_2)}, c_1 + \frac{\beta_1}{4(\beta_1 + \beta_2)} + \frac{5\beta_1}{32(\beta_1 + \beta_2)} \right\}. \end{aligned}$$

Above, the first expression inside the min corresponds to control action  $\gamma = 0$ , the second to  $\gamma = 1$ . We see that when being in state  $x = -1$  at time  $t = 1$ , it is optimal to choose  $\gamma = 0$ , while when being in state  $x = 1$ , it is optimal to choose  $\gamma = 1$  provided that

$$(5.2) \quad 0 < c_1 < \frac{5\beta_1}{32(\beta_1 + \beta_2)}.$$

Assume from now on that (5.2) holds. Then  $\varphi_+$  gives the optimal control actions at time

$t = 1$ . As to time  $t = 0$ , we have

$$\begin{aligned} V_+(0, 1) &= \min \left\{ \frac{1}{2}V_+(1, 1) + \frac{1}{2}V_+(1, -1), c_0 + \frac{3}{4}V_+(1, 1) + \frac{1}{4}V_+(1, -1) \right\} \\ &= \min \left\{ -\frac{1}{2} \left( \frac{5\beta_1}{32(\beta_1 + \beta_2)} - c_1 \right), c_0 - \frac{3}{4} \left( \frac{5\beta_1}{32(\beta_1 + \beta_2)} - c_1 \right) - \frac{\beta_1}{8(\beta_1 + \beta_2)} \right\}, \\ V_+(0, -1) &= \min \left\{ \frac{1}{2}V_+(1, 1) + \frac{1}{2}V_+(1, -1), c_0 + \frac{1}{4}V_+(1, 1) + \frac{3}{4}V_+(1, -1) \right\} \\ &= \min \left\{ -\frac{1}{2} \left( \frac{5\beta_1}{32(\beta_1 + \beta_2)} - c_1 \right), c_0 - \frac{1}{4} \left( \frac{5\beta_1}{32(\beta_1 + \beta_2)} - c_1 \right) + \frac{\beta_1}{8(\beta_1 + \beta_2)} \right\}. \end{aligned}$$

We see that when being in state  $x = -1$  at time  $t = 0$ , it is optimal to choose  $\gamma = 0$ , while when being in state  $x = 1$ , it is certainly optimal to choose  $\gamma = 1$  if

$$(5.3) \quad 0 < c_0 < \frac{\beta_1}{8(\beta_1 + \beta_2)}.$$

Assume from now on that (5.3) holds. Then  $\varphi_+$  gives the optimal control actions also at time  $t = 0$ . Thus, under (5.2) and (5.3),  $\varphi_+$  is optimal on  $\{\Phi = \varphi_+\}$ .

We proceed similarly to verify that  $\hat{\varphi}_+$  is optimal on the event  $\{\Phi = \hat{\varphi}_+\}$ . For  $x \in \{-1, 1\}$ , set

$$\begin{aligned} \hat{V}_+(2, x) &\doteq F(x, m_0) = 0, \\ \hat{V}_+(1, x) &\doteq \min_{\gamma \in \{0,1\}} \left\{ c_1 \cdot \gamma - x \cdot \mathbf{M}(m_1^+) + \mathbf{E} \left[ \hat{V}_+(2, \Psi(x, \gamma, \xi(2))) \right] \right\}, \\ \hat{V}_+(0, x) &\doteq \min_{\gamma \in \{0,1\}} \left\{ c_0 \cdot \gamma + \mathbf{E} \left[ \hat{V}_+(1, \Psi(x, \gamma, \xi(1))) \right] \right\}. \end{aligned}$$

We have again omitted the weight factor, here  $\beta_3$ . Costs are thus computed with respect to the conditional probability  $\mathbf{P}(\cdot | \Phi = \hat{\varphi}_+)$ . Notice that  $\hat{V}_+$  is the value function of the optimal control problem the representative player faces when being told to play  $\hat{\varphi}_+$  by the mediator. We have  $\hat{V}_+(2, \cdot) \equiv 0$ ,

$$\begin{aligned} \hat{V}_+(1, 1) &= \min \left\{ -\frac{\beta_1}{4(\beta_1 + \beta_2)}, c_1 - \frac{\beta_1}{4(\beta_1 + \beta_2)} \right\}, \\ \hat{V}_+(1, -1) &= \min \left\{ \frac{\beta_1}{4(\beta_1 + \beta_2)}, c_1 + \frac{\beta_1}{4(\beta_1 + \beta_2)} \right\}. \end{aligned}$$

The first expression inside the min above corresponds again to control action  $\gamma = 0$ , the second to  $\gamma = 1$ . We see that at time  $t = 1$  it is always optimal to choose  $\gamma = 0$ . This is



exactly what  $\hat{\varphi}_+$  prescribes at  $t = 1$ . As to time  $t = 0$ , we have

$$\begin{aligned}\hat{V}_+(0, 1) &= \min \left\{ \frac{1}{2}\hat{V}_+(1, 1) + \frac{1}{2}\hat{V}_+(1, -1), c_0 + \frac{3}{4}\hat{V}_+(1, 1) + \frac{1}{4}\hat{V}_+(1, -1) \right\} \\ &= \min \left\{ 0, c_0 - \frac{\beta_1}{8(\beta_1 + \beta_2)} \right\}, \\ \hat{V}_+(0, -1) &= \min \left\{ \frac{1}{2}\hat{V}_+(1, 1) + \frac{1}{2}\hat{V}_+(1, -1), c_0 + \frac{1}{4}\hat{V}_+(1, 1) + \frac{3}{4}\hat{V}_+(1, -1) \right\} \\ &= \min \left\{ 0, c_0 + \frac{\beta_1}{8(\beta_1 + \beta_2)} \right\}.\end{aligned}$$

We see that when being in state  $x = -1$  at time  $t = 0$ , it is optimal to choose  $\gamma = 0$ , while when being in state  $x = 1$ , it is optimal to choose  $\gamma = 1$  if (5.3) holds. This is what  $\hat{\varphi}_+$  prescribes at time  $t = 0$ . Therefore, under (5.3),  $\hat{\varphi}_+$  is optimal on  $\{\Phi = \hat{\varphi}_+\}$ .

We have thus established that the correlated flow  $\rho$  defined above is a correlated solution of the mean field game with initial distribution  $m_0$  provided the coefficients satisfy (5.1), (5.2), and (5.3). In particular, given any  $\alpha \in (0, 1)$ , we have that  $\rho$  is a solution if

$$\beta_1 = \beta_2 = \frac{\alpha}{4}, \quad \beta_3 = \beta_4 = \frac{1 - \alpha}{4}, \quad 0 < c_0 < \frac{1}{16}, \quad 0 < c_1 < \frac{5}{64}.$$

*Remark 5.1.* The correlated flow constructed in the example above does not simply arise by randomizing among solutions with deterministic flow of measures. While the representative player can infer with probability one what the flow of measures will be when receiving the recommendation to play strategy  $\varphi_+$ ,  $\hat{\varphi}_+$ ,  $\varphi_-$ , or  $\hat{\varphi}_-$ , she cannot do so when being told to play strategy  $\varphi_o$ . In this case, the final evolution of the flow of measures will be uncertain for the player not only at time zero, but also at time  $t = 1$ . Also notice that in our example there are multiple, actually infinitely many solutions.

## 6 Convergence of correlated equilibria

For  $N \in \mathbb{N}$ , let  $\mathbf{m}^N \in \mathcal{P}(\mathcal{X}^N)$ , let  $\gamma^N \in \mathcal{P}(\mathcal{R}^N)$  be a strategy profile, and let  $(\varepsilon_N)_{N \in \mathbb{N}} \subset [0, \infty)$ . Moreover, let  $\mathbf{m}_0 \in \mathcal{P}(\mathcal{X})$ . We make the following assumptions:

- (A1) Continuity property of the system function  $\Psi$ : There exists a measurable function  $\mathbf{w} : [0, \infty) \rightarrow [0, 1]$  with  $\mathbf{w}(s) \rightarrow 0$  as  $s \rightarrow 0+$  such that for every  $(t, x, a) \in \{0, \dots, T - 1\} \times \mathcal{X} \times \Gamma$ , all  $m, \tilde{m} \in \mathcal{P}(\mathcal{X})$ ,

$$\int_{\mathcal{Z}} \mathbf{1}_{\Psi(t, x, m, a, z) \neq \Psi(t, x, \tilde{m}, a, z)} \nu(dz) \leq \mathbf{w}(\text{dist}(m, \tilde{m})).$$

Moreover, for every  $t \in \{0, \dots, T - 1\}$ , every  $\tau \in \mathcal{P}(\mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma)$ ,  $\Psi(t, \cdot)$  is  $\tau \otimes \nu$ -almost everywhere continuous.

(A2) The cost coefficients  $f, F$  are continuous.

(A3) For every  $N \in \mathbb{N}$ ,  $\gamma^N$  is a symmetric  $\varepsilon_N$ -correlated equilibrium in restricted strategies with initial distribution  $\mathbf{m}^N$ .

(A4) The sequence  $(\varepsilon_N)_{N \in \mathbb{N}}$  converges to zero as  $N \rightarrow \infty$ .

(A5) Initial distributions:  $\mathbf{m}^N = \otimes^N \mathbf{m}_{0,N}$  where  $\mathbf{m}_{0,N} \rightarrow \mathbf{m}_0$  as  $N \rightarrow \infty$ .

*Remark 6.1.* The continuity property (A1) is satisfied, for instance, if  $\Psi$  is defined as follows. Choose  $L > 0$ , let  $d \doteq |\mathcal{X}|$  be the number of states, and let  $\sigma: \{1, \dots, d\} \rightarrow \mathcal{X}$  be a bijection. For  $(t, x, a) \in \{0, \dots, T-1\} \times \mathcal{X} \times \Gamma$ , choose functions  $a_{1,t,x,a}, \dots, a_{d,t,x,a}: \mathcal{P}(\mathcal{X}) \rightarrow [0, 1]$  that are  $L$ -Lipschitz continuous and such that  $\sum_{i=1}^d a_{i,t,x,a} = 1$ . Now set

$$\Psi(t, x, m, a, z) \doteq \sigma \left( \operatorname{argmin} \left\{ j \in \{1, \dots, d\} : \sum_{i=1}^j a_{i,t,x,a}(m) \geq z \right\} \right),$$

$$(t, x, m, a, z) \in \{0, \dots, T-1\} \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma \times \mathcal{Z}.$$

Recall that  $\nu$  indicates the uniform distribution on  $\mathcal{Z} = [0, 1]$ . With the above definition of  $\Psi$ , we have for all  $m, \tilde{m} \in \mathcal{P}(\mathcal{X})$ ,

$$\int_{\mathcal{Z}} \mathbf{1}_{\Psi(t,x,m,a,z) \neq \Psi(t,x,\tilde{m},a,z)} \nu(dz) \leq \sum_{j=1}^{d-1} \left| \sum_{i=1}^j a_{i,t,x,a}(m) - \sum_{i=1}^j a_{i,t,x,a}(\tilde{m}) \right|$$

$$\leq L \frac{d(d-1)}{2} \operatorname{dist}(m, \tilde{m}).$$

The first part of (A1) is thus satisfied with  $\mathbf{w}(s) = L \frac{d(d-1)}{2} s$ . This modulus of continuity changes if the functions  $a_{1,t,x,a}, \dots, a_{d,t,x,a}$  are (uniformly) continuous, but not Lipschitz. In order to check the second part of (A1), fix  $t \in \{0, \dots, T-1\}$  and let  $D_t$  denote the set of points of discontinuity of  $\Psi(t, \cdot)$ . By construction, the functions  $a_{1,t,x,a}, \dots, a_{d,t,x,a}$  are continuous on  $\mathcal{P}(\mathcal{X})$ , and they depend continuously also on  $x, a$  since  $\mathcal{X}, \Gamma$  are finite sets. In view of the definition of  $\Psi$ , it follows that

$$D_t \subseteq \bigcup_{j=1}^d \left\{ (x, m, a, z) \in \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma \times \mathcal{Z} : \sum_{i=1}^j a_{i,t,x,a}(m) = z \right\}.$$

The assertion of the second part of (A1) is now a consequence of Fubini's theorem as  $\nu$  assigns measure zero to any finite subset of  $\mathcal{Z}$ .

*Remark 6.2.* The example from Section 4 satisfies Assumptions (A1) and (A2) above. The system function  $\Psi$  there can be written as in Remark 6.1. In particular,  $\Psi(t, \cdot)$  is almost surely continuous with respect to  $\tau \otimes \nu$  given any  $\tau \in \mathcal{P}(\mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma)$ , but discontinuous as a function  $\mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \Gamma \times \mathcal{Z} \rightarrow \mathcal{X}$ .

For  $N \in \mathbb{N} \setminus \{1\}$ , let

$$((\Omega_N, \mathcal{F}_N, \mathbf{P}_N), \Phi_1^N, \dots, \Phi_N^N, X_1^N(\cdot), \dots, X_N^N(\cdot), \xi_1^N(\cdot), \dots, \xi_N^N(\cdot))$$

be a realization of the triple  $(\mathbf{m}^N, \gamma^N, \text{Id})$ , and set

$$\rho^N \doteq \mathbf{P}_N \circ (\Phi_1^N, \mu_1^N(0), \dots, \mu_1^N(T))^{-1},$$

where  $\mu_1^N(t) = \frac{1}{N-1} \sum_{j=2}^N \delta_{X_j^N(t)}$ ,  $t \in \{0, \dots, T\}$ , as above. We then have the following convergence result:

**Theorem 6.1.** *Grant (A1) – (A5). Then  $(\rho^N)_{N \in \mathbb{N}}$  is relatively compact as a subset of  $\mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$ , and any limit point is a correlated solution of the mean field game in restricted strategies with initial distribution  $\mathbf{m}_0$ .*

*Proof.* If  $\mathcal{S}$  is a compact Polish space, then  $\mathcal{P}(\mathcal{S})$  is compact with respect to the topology of weak convergence of measures. Since  $\mathcal{X}, \mathcal{R}$  are finite sets, hence compact Polish spaces under the discrete topology, we have that  $\mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$  is compact. This in turn implies the relative compactness of  $(\rho^N)_{N \in \mathbb{N}}$  in  $\mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$ .

In order to identify the limit points of  $(\rho^N)_{N \in \mathbb{N}}$ , set

$$\eta^N \doteq \mathbf{P}_N \circ (\Phi_1^N, X_1^N(0), \dots, X_1^N(T), \xi_1^N(1), \dots, \xi_1^N(T), \mu_1^N(0), \dots, \mu_1^N(T))^{-1}.$$

Clearly,  $\rho^N$  coincides with the image (push forward) measure of  $\eta^N$  under the natural projection  $\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1} \rightarrow \mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1}$ . Moreover,  $(\eta^N)_{N \in \mathbb{N}}$  is relatively compact in  $\mathcal{P}(\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1})$  since the space  $\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1}$  is compact, too. To identify the limit points of  $(\rho^N)_{N \in \mathbb{N}}$  it is therefore enough to characterize the limits of convergent subsequences of  $(\eta^N)_{N \in \mathbb{N}}$ . We will proceed in several steps.

**Step One.** Let  $(\eta^{N_k})_{k \in \mathbb{N}}$  be any convergent subsequence of  $(\eta^N)_{N \in \mathbb{N}}$ , and denote its limit by  $\eta$ . Let  $(\Phi, X(\cdot), \xi(\cdot), \mu(\cdot))$  be a  $\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1}$ -valued random element on some probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  such that

$$\eta = \mathbf{P} \circ (\Phi, X(0), \dots, X(T), \xi(1), \dots, \xi(T), \mu(0), \dots, \mu(T))^{-1}.$$

Set

$$\rho \doteq \mathbf{P} \circ (\Phi, \mu(0), \dots, \mu(T))^{-1}.$$

Then the following properties hold:

- (a)  $\mathbf{P} \circ (X(0))^{-1} = \mathbf{m}_0$ ;
- (b)  $\xi(t)$ ,  $t \in \{1, \dots, T\}$ , are i.i.d. with common distribution  $\nu$ ;

(c)  $\xi(\cdot)$ ,  $X(0)$ , and  $(\Phi, \mu(\cdot))$  are independent;

(d)  $\mathbf{P}$ -almost surely, for every  $t \in \{0, \dots, T-1\}$ ,

$$X(t+1) = \Psi(t, X(t), \mu(t), \Phi(t, X(t)), \xi(t+1));$$

(e)  $\lim_{k \rightarrow \infty} J_1^{N_k}(\mathbf{m}^{N_k}; \gamma^{N_k}, \text{Id}) = J(\mathbf{m}_0; \rho, \text{Id})$ .

Point (a) is a consequence of assumption (A5). Point (b) follows from the corresponding independence property of  $\xi_1^N(t)$ ,  $t \in \{1, \dots, T\}$ , and their joint convergence in distribution.

Points (c) and (d) will follow from analogous properties established in the next two steps by taking  $u = \text{Id}$ . The convergence of costs according to (e) is again a consequence of convergence in distribution in conjunction with assumption (A2).

**Step Two.** Let  $u: \mathcal{R} \rightarrow \mathcal{R}$  be any strategy modification. We define a realization of  $(\mathbf{m}_0, \rho, u)$  with the same noises and in the same probability space as the realization of  $(\mathbf{m}_0, \rho, \text{Id})$  given in Step One. For  $N \in \mathbb{N} \setminus \{1\}$ , define  $\mathcal{X}$ -valued random variables  $\tilde{X}_j^N(t)$ ,  $j \in \{1, \dots, N\}$ ,  $t \in \{0, \dots, T\}$ , on  $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$  recursively through

$$\begin{aligned} \tilde{X}_i^N(0) &\doteq X_i^N(0) \text{ for every } i \in \{1, \dots, N\}, \\ \tilde{X}_1^N(t+1) &\doteq \Psi\left(t, \tilde{X}_1^N(t), \tilde{\mu}_1^N(t), u \circ \Phi_1^N\left(t, \tilde{X}_1^N(t)\right), \xi_1^N(t+1)\right), \\ \tilde{X}_j^N(t+1) &= \Psi\left(t, \tilde{X}_j^N(t), \tilde{\mu}_j^N(t), \Phi_j^N\left(t, \tilde{X}_j^N(t)\right), \xi_j^N(t+1)\right), \quad j \neq 1, \\ &t \in \{0, \dots, T-1\}, \end{aligned}$$

where

$$\tilde{\mu}_i^N(t) \doteq \frac{1}{N-1} \sum_{l \neq i} \delta_{\tilde{X}_l^N(t)}, \quad i \in \{1, \dots, N\}.$$

Set

$$\tilde{\eta}^N \doteq \mathbf{P}_N \circ \left( \Phi_1^N, \tilde{X}_1^N(0), \dots, \tilde{X}_1^N(T), \xi_1^N(1), \dots, \xi_1^N(T), \tilde{\mu}_1^N(0), \dots, \tilde{\mu}_1^N(T) \right)^{-1}.$$

Reasoning as in Step One, we have that  $(\tilde{\eta}^N)_{N \in \mathbb{N}}$  is relatively compact in  $\mathcal{P}(\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1})$ , and so is  $(\tilde{\eta}^{N_k})_{k \in \mathbb{N}}$ . Choose any convergent subsequence of  $(\tilde{\eta}^{N_k})_{k \in \mathbb{N}}$ , which we continue to indicate by  $(\tilde{\eta}^{N_k})_{k \in \mathbb{N}}$ , thus omitting the sub-subscript. Denote its limit by  $\tilde{\eta}$ , and let  $(\tilde{\Phi}, \tilde{X}(\cdot), \tilde{\xi}(\cdot), \tilde{\mu}(\cdot))$  be a  $\mathcal{R} \times \mathcal{X}^{T+1} \times \mathcal{Z}^T \times \mathcal{P}(\mathcal{X})^{T+1}$ -valued random element on some probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbf{P}})$  such that

$$\tilde{\eta} = \tilde{\mathbf{P}} \circ \left( \tilde{\Phi}, \tilde{X}(0), \dots, \tilde{X}(T), \tilde{\xi}(1), \dots, \tilde{\xi}(T), \tilde{\mu}(0), \dots, \tilde{\mu}(T) \right)^{-1}.$$

Set

$$\tilde{\rho} \doteq \tilde{\mathbf{P}} \circ \left( \tilde{\Phi}, \tilde{\mu}(0), \dots, \tilde{\mu}(T) \right)^{-1}.$$

Then the following properties hold:

- (a)  $\mathbf{P} \circ (\tilde{X}(0))^{-1} = \mathbf{m}_0$ ;
- (b)  $\tilde{\xi}(t)$ ,  $t \in \{1, \dots, T\}$ , are i.i.d. with common distribution  $\nu$ ;
- (c)  $\tilde{\xi}(\cdot)$ ,  $\tilde{X}(0)$ , and  $(\tilde{\Phi}, \tilde{\mu}(\cdot))$  are independent;
- (d)  $\tilde{\mathbf{P}}$ -almost surely, for every  $t \in \{0, \dots, T-1\}$ ,

$$\tilde{X}(t+1) = \Psi \left( t, \tilde{X}(t), \tilde{\mu}(t), u \circ \tilde{\Phi}(t, \tilde{X}(t)), \tilde{\xi}(t+1) \right);$$

- (e)  $\lim_{k \rightarrow \infty} J_1^{N_k}(\mathbf{m}^{N_k}; \gamma^{N_k}, u) = J(\mathbf{m}_0; \tilde{\rho}, u)$ .

Points (a), (b), and (e) follow as in Step One. The independence property (c) will be established in Step Three. To verify Property (d), define functions  $G_t: \mathcal{R} \times \mathcal{X} \times \mathcal{P}(\mathcal{X}) \times \mathcal{Z} \rightarrow \mathcal{X}$ ,  $t \in \{1, \dots, T\}$ , by setting

$$G_t(\varphi, x, m, z) \doteq \Psi(t, x, m, u(\varphi)(t, x), z).$$

The function  $G_t$  is  $\sigma \otimes \nu$ -almost everywhere continuous given any measure  $\sigma \in \mathcal{P}(\mathcal{R} \times \mathcal{X} \times \mathcal{P}(\mathcal{X}))$ . This follows from the second part of assumption (A1) and the fact that the spaces  $\mathcal{R}$  and  $\mathcal{X}$  are finite. In particular, the mapping  $\mathcal{R} \times \mathcal{X} \rightarrow \Gamma$  given by  $(\varphi, x) \mapsto \varphi(t, x)$  is continuous for every  $t \in \{0, \dots, T-1\}$ . Property (d) is now a consequence of Lemma A.1, the almost everywhere continuity of  $G_t$  in conjunction with the independence properties (b) and (c), and the convergence in distribution of

$$\left( \tilde{X}_1^{N_k}(t+1), \left( \tilde{\Phi}_1^{N_k}, \tilde{X}_1^{N_k}(t), \tilde{\mu}_1^{N_k}(t), \xi_1^{N_k}(t+1) \right) \right)$$

to

$$\left( \tilde{X}(t+1), \left( \tilde{\Phi}, \tilde{X}(t), \tilde{\mu}(t), \tilde{\xi}(t+1) \right) \right)$$

as  $k \rightarrow \infty$ , by the mapping theorem.

Point (e) above, together with the corresponding property established in Step One, entails thanks to assumptions (A3) and (A4) that

$$J(\mathbf{m}_0; \rho, \text{Id}) \leq J(\mathbf{m}_0; \tilde{\rho}, u).$$

It remains to show that  $\rho = \tilde{\rho}$  and that property (c) holds.

**Step Three.** For  $N \in \mathbb{N} \setminus \{1\}$ , recursively define  $\mathcal{X}$ -valued random variables  $\hat{X}_j^N(t)$ ,  $j \in \{2, \dots, N\}$ ,  $t \in \{0, \dots, T\}$ , on  $(\Omega_N, \mathcal{F}_N, \mathbf{P}_N)$  through

$$\begin{aligned} \hat{X}_j^N(0) &\doteq X_j^N(0) \text{ for every } j \in \{2, \dots, N\}, \\ \hat{X}_j^N(t+1) &= \Psi \left( t, \hat{X}_j^N(t), \hat{\mu}_1^N(t), \Phi_j^N \left( t, \hat{X}_j^N(t) \right), \xi_j^N(t+1) \right), \quad j \neq 1, \\ &t \in \{0, \dots, T-1\}, \end{aligned}$$

where

$$(6.1) \quad \hat{\mu}_1^N(t) \doteq \frac{1}{N-1} \sum_{l=2}^N \delta_{\hat{X}_l^N(t)}.$$

The main difference between  $\tilde{X}$  and  $\hat{X}$  is that for the latter we consider only players in  $\{2, \dots, N\}$ , thus excluding the contribution of Player 1; see the empirical measure (6.1) above.

Let  $\mathbf{dist}$  be the metric on  $\mathcal{P}(\mathcal{X})$  introduced in Section 2. We claim that for every  $t \in \{0, \dots, T\}$ ,

$$(6.2) \quad \frac{1}{N-1} \sum_{j=2}^N \mathbf{P}_N \left( \hat{X}_j^N(t) \neq \tilde{X}_j^N(t) \right) \xrightarrow{N \rightarrow \infty} 0.$$

We verify (6.2) by induction over  $t \in \{0, \dots, T\}$ . First notice that, by inequality (2.1), for all  $j \in \{1, \dots, N\}$ , including  $j = 1$ , all  $t \in \{0, \dots, T\}$ ,

$$(6.3) \quad \mathbf{E}_N [\mathbf{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_j^N(t))] \leq \frac{1}{N-1} + \frac{1}{N-1} \sum_{l=2}^N \mathbf{P}_N \left( \hat{X}_l^N(t) \neq \tilde{X}_l^N(t) \right).$$

It is clear that relation (6.2) holds if  $t = 0$ , since  $\hat{X}_j^N(0) = \tilde{X}_j^N(0) = X_j^N(0)$  for all  $j \in \{2, \dots, N\}$ . Now, suppose that (6.2) holds for some  $t \in \{0, \dots, T-1\}$ . For each  $j \in \{2, \dots, N\}$ , we have

$$\begin{aligned} \mathbf{P}_N \left( \hat{X}_j^N(t+1) \neq \tilde{X}_j^N(t+1) \right) &\leq \mathbf{P}_N \left( \hat{X}_j^N(t) \neq \tilde{X}_j^N(t) \right) \\ &+ \mathbf{E}_N \left[ \mathbf{1}_{\Psi(t, \hat{X}_j^N(t), \hat{\mu}_1^N(t), \Phi_j^N(t, \hat{X}_j^N(t)), \xi_j^N(t+1)) \neq \Psi(t, \hat{X}_j^N(t), \tilde{\mu}_j^N(t), \Phi_j^N(t, \hat{X}_j^N(t)), \xi_j^N(t+1))} \right]. \end{aligned}$$

Using the Fubini-Tonelli theorem, the independence of  $\xi_j^N(t+1)$ , as well as assumption (A1), we find the expected value in the display above to be less than or equal to

$$\begin{aligned} \mathbf{E}_N \left[ \int_{\mathcal{Z}} \mathbf{1}_{\Psi(t, \hat{X}_j^N(t), \hat{\mu}_1^N(t), \Phi_j^N(t, \hat{X}_j^N(t)), z) \neq \Psi(t, \hat{X}_j^N(t), \tilde{\mu}_j^N(t), \Phi_j^N(t, \hat{X}_j^N(t)), z)} \nu(dz) \right] \\ \leq \mathbf{E}_N [\mathbf{w}(\mathbf{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_j^N(t)))] . \end{aligned}$$

The induction hypothesis and (6.3) imply that

$$\max_{j \in \{2, \dots, N\}} \mathbf{E}_N [\mathbf{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_j^N(t))] \xrightarrow{N \rightarrow \infty} 0.$$

This in turn entails, by Markov's inequality and the fact that  $\mathbf{w}$  is bounded non-negative with  $\mathbf{w}(s) \rightarrow 0$  as  $s \rightarrow 0+$ , that

$$\max_{j \in \{2, \dots, N\}} \mathbf{E}_N [\mathbf{w}(\mathbf{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_j^N(t)))] \xrightarrow{N \rightarrow \infty} 0.$$

Using again the induction hypothesis, we find that

$$\begin{aligned} & \frac{1}{N-1} \sum_{j=2}^N \mathbf{P}_N \left( \hat{X}_j^N(t+1) \neq \tilde{X}_j^N(t+1) \right) \\ & \leq \frac{1}{N-1} \sum_{j=2}^N \mathbf{P}_N \left( \hat{X}_j^N(t) \neq \tilde{X}_j^N(t) \right) + \max_{j \in \{2, \dots, N\}} \mathbf{E}_N [\mathbf{w}(\text{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_j^N(t)))] \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

This establishes (6.2) for all  $t \in \{0, \dots, T\}$ .

As a consequence of (6.2) and (6.3), we obtain

$$\mathbf{E}_N \left[ \sum_{t=0}^T \text{dist}(\hat{\mu}_1^N(t), \tilde{\mu}_1^N(t)) \right] \xrightarrow{N \rightarrow \infty} 0.$$

By Lemma A.2, this implies that the sequences

$$(\Phi_1^{N_k}, \tilde{X}_1^{N_k}(\cdot), \xi_1^{N_k}(\cdot), \tilde{\mu}_1^{N_k}(\cdot)), \quad \text{and} \quad (\Phi_1^{N_k}, \tilde{X}_1^{N_k}(\cdot), \xi_1^{N_k}(\cdot), \hat{\mu}_1^{N_k}(\cdot))$$

have the same limit in distribution, namely  $(\tilde{\Phi}, \tilde{X}(\cdot), \tilde{\xi}(\cdot), \tilde{\mu}(\cdot))$ . From equation (6.1) it is clear that the definition of  $\hat{\mu}_1^N(\cdot)$  through the random variables  $\hat{X}_j^N(t)$ , with  $j \geq 2$ , does not depend on the strategy modification  $u$ . Thus, also  $(\Phi_1^{N_k}, X_1^{N_k}(\cdot), \xi_1^{N_k}(\cdot), \mu_1^{N_k}(\cdot))$  and  $(\Phi_1^{N_k}, X_1^{N_k}(\cdot), \xi_1^{N_k}(\cdot), \hat{\mu}_1^{N_k}(\cdot))$  have the same limit in distribution, namely  $(\Phi, X(\cdot), \xi(\cdot), \mu(\cdot))$ . Since  $X_1^N(0) = \tilde{X}_1^N(0)$  for every  $N \in \mathbb{N}$ , we find that

$$\mathbf{P} \circ (\Phi, X(0), \xi(\cdot), \mu(\cdot))^{-1} = \tilde{\mathbf{P}} \circ (\tilde{\Phi}, \tilde{X}(0), \tilde{\xi}(\cdot), \tilde{\mu}(\cdot))^{-1}.$$

This implies, in particular, that  $\rho = \tilde{\rho}$ . In addition, by the fact that  $\hat{\mu}_1^N(\cdot)$  does not depend on Player 1's state (see (6.1)), the independence of  $X_1^N(0), \dots, X_N^N(0)$  according to (A5), and the independence of  $X_j^N(0), \xi_j^N(\cdot), \Phi_j^N, j \in \{1, \dots, N\}$ , we have that

$$X_1^N(0), \xi_1^N(\cdot), \text{ and } (\Phi_1^N, \hat{\mu}_1^N) \text{ are independent, for every } N \in \mathbb{N}.$$

Convergence in distribution now yields property (c). Thanks to Step Two, it follows that

$$J(\mathbf{m}_0; \rho, \text{Id}) \leq J(\mathbf{m}_0; \rho, u).$$

The optimality condition of Definition 4.1 is therefore satisfied.

**Step Four.** We verify the consistency condition in Definition 4.1. For  $N \in \mathbb{N}$ , let  $\mu^N(t)$  denote the empirical measure of the states at time  $t \in \{0, \dots, T\}$  of all players in the  $N$ -player game, and let  $\boldsymbol{\mu}^N$  denote the empirical measure of their state trajectories:

$$\mu^N(t) \doteq \frac{1}{N} \sum_{i=1}^N \delta_{X_i^N(t)}, \quad \boldsymbol{\mu}^N \doteq \frac{1}{N} \sum_{i=1}^N \delta_{(X_i^N(0), \dots, X_i^N(T))}.$$

Thus,  $\mu^N(t)$  is a  $\mathcal{P}(\mathcal{X})$ -valued random variable, for every  $t$ , while  $\boldsymbol{\mu}^N$  is a  $\mathcal{P}(\mathcal{X}^{T+1})$ -valued random variable.

As a consequence of the symmetry of the correlated profiles according to (A3), of the initial distributions according to (A5), and of the dynamics, we obtain that  $(X_1^N(t), \dots, X_N^N(t))$  is a finite exchangeable sequence of  $\mathcal{X}$ -valued random variables for every  $t$ , while  $(X_1^N(\cdot), \dots, X_N^N(\cdot))$  is a finite exchangeable sequence of  $\mathcal{X}^{T+1}$ -valued random variables. Lemma A.5 now yields the conditional distributions of the state and of the state trajectory of player 1 given the corresponding empirical measure:

$$\begin{aligned} \mathbf{P}_N (X_1^N(t) \in \cdot \mid \mu^N(t)) &= \mu^N(t)(\cdot), & t \in \{0, \dots, T\}, \\ \mathbf{P}_N ((X_1^N(0), \dots, X_1^N(T)) \in \cdot \mid \boldsymbol{\mu}^N) &= \boldsymbol{\mu}^N(\cdot). \end{aligned}$$

Applying the conditional distribution of the state trajectory of player 1 to sets of the form  $\mathcal{X}^t \times B \times \mathcal{X}^{T-t}$  shows that also

$$\mathbf{P}_N (X_1^N(t) \in \cdot \mid \boldsymbol{\mu}^N) = \mu^N(t)(\cdot) \text{ for every } t \in \{0, \dots, T\}.$$

The  $\sigma$ -algebra generated by  $\mu^N(\cdot) = (\mu^N(0), \dots, \mu^N(T))$ , the flow of empirical measures, is contained in  $\sigma(\boldsymbol{\mu}^N)$ , while it contains  $\sigma(\mu^N(t))$  for every  $t$ . In view of Lemma A.4, we thus find that

$$\mathbf{P}_N (X_1^N(t) \in \cdot \mid (\mu^N(0), \dots, \mu^N(T))) = \mu^N(t)(\cdot) \text{ for every } t \in \{0, \dots, T\}.$$

According to Step One (using the mapping theorem), we have that, for every  $t \in \{0, \dots, T\}$ , the random vector  $(X_1^{N_k}(t), (\mu_1^{N_k}(0), \dots, \mu_1^{N_k}(T)), \mu_1^{N_k}(t))$  converges in distribution to  $(X(t), (\mu(0), \dots, \mu(T)), \mu(t))$  as  $k \rightarrow \infty$ . Now, for every  $k \in \mathbb{N}$ , every  $\omega \in \Omega_{N_k}$ , every  $t \in \{0, \dots, T\}$ ,

$$\text{dist} \left( \mu_{1,\omega}^{N_k}(t), \mu_{\omega}^{N_k}(t) \right) \leq (N_k - 1) \left( \frac{1}{N_k - 1} - \frac{1}{N_k} \right) + \frac{1}{N_k} = \frac{2}{N_k},$$

where  $\text{dist}$  is the metric on  $\mathcal{P}(\mathcal{X})$  introduced in Section 2. This implies by Lemma A.2 that also the vector  $(X_1^{N_k}(t), (\mu^{N_k}(0), \dots, \mu^{N_k}(T)), \mu^{N_k}(t))$  converges in distribution to  $(X(t), (\mu(0), \dots, \mu(T)), \mu(t))$  as  $k \rightarrow \infty$ . By Lemma A.3, we now find the conditional distribution of  $X(t)$  given the flow of measures  $(\mu(0), \dots, \mu(T))$ :

$$\mathbf{P}(X(t) \in \cdot \mid (\mu(0), \dots, \mu(T))) = \mu(t)(\cdot) \text{ for every } t \in \{0, \dots, T\},$$

which yields the consistency condition.  $\square$

## 7 Approximate $N$ -player correlated equilibria

The next result shows how to construct a sequence of approximate  $N$ -player correlated equilibria with approximation error tending to zero as  $N \rightarrow \infty$  provided we have a



correlated solution to the mean field game. The construction can be roughly described as follows: first, the mediator draws some flow of measures from the second marginal of the correlated solution and second, conditioning on such a flow, he draws a sequence of i.i.d. recommendations that are privately communicated to the players in the  $N$ -player games.

In order to rigorously state the result, let  $\mathbf{m}_0 \in \mathcal{P}(\mathcal{X})$ , and let  $(\mathbf{m}^N)_{N \in \mathbb{N}}$  be such that assumption (A5) holds:  $\mathbf{m}^N = \otimes^N \mathbf{m}_{0,N}$  with  $\mathbf{m}_{0,N} \rightarrow \mathbf{m}_0$  as  $N \rightarrow \infty$ .

**Theorem 7.1.** *Grant (A1) and (A2). Suppose that  $\rho \in \mathcal{P}(\mathcal{R} \times \mathcal{P}(\mathcal{X})^{T+1})$  is a correlated solution of the mean field game with initial distribution  $\mathbf{m}_0$ . For  $N \in \mathbb{N}$ , define  $\gamma^N \in \mathcal{P}(\mathcal{R}^N)$  through*

$$\gamma^N(C_1 \times \dots \times C_N) \doteq \int_{\mathcal{P}(\mathcal{X})^{T+1}} \prod_{i=1}^N \rho_1(C_i | m) \rho_2(dm),$$

where  $\rho$  has been factorized according to

$$\rho(C \times B) = \int_B \rho_1(C | m) \rho_2(dm), \quad C \subseteq \mathcal{R}, B \in \otimes^{T+1} \mathcal{B}(\mathcal{P}(\mathcal{X})).$$

Then there exists a sequence  $(\varepsilon_N)_{N \in \mathbb{N}} \subset [0, \infty)$  such that  $\gamma^N$  is an  $\varepsilon_N$ -correlated equilibrium with initial distribution  $\mathbf{m}^N$ , for every  $N$ , and  $\varepsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ .

*Proof.* By symmetry, we may restrict attention to strategy modifications of player 1. For  $N \in \mathbb{N}$ , set

$$\varepsilon_N \doteq J_1^N(\mathbf{m}^N; \gamma^N, \text{Id}) - \inf_{u \in \mathcal{U}} J_1^N(\mathbf{m}^N; \gamma^N, u).$$

Then  $\gamma^N$  is an  $\varepsilon_N$ -correlated equilibrium with initial distribution  $\mathbf{m}^N$ . It remains to show that  $\varepsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ . To this end, choose a sequence of strategy modifications  $u^N$  such that

$$J_1^N(\mathbf{m}^N; \gamma^N, u^N) \leq \inf_{u \in \mathcal{U}} J_1^N(\mathbf{m}^N; \gamma^N, u) + \frac{1}{N}.$$

We have to show that  $\lim_{N \rightarrow \infty} J_1^N(\mathbf{m}^N; \gamma^N, \text{Id}) = J(\mathbf{m}_0; \rho, \text{Id})$  and that

$$\liminf_{N \rightarrow \infty} J_1^N(\mathbf{m}^N; \gamma^N, u^N) \geq J(\mathbf{m}_0; \rho, \text{Id}),$$

as this entails that  $\varepsilon_N \rightarrow 0$  as  $N \rightarrow \infty$ .

In our setting, the set of strategy modifications, i.e. of mappings  $u : \mathcal{R} \rightarrow \mathcal{R}$ , is finite since  $\mathcal{R}$  is finite. Therefore and by the optimality condition, the above limit inferior will be established as soon as  $\liminf_{N \rightarrow \infty} J_1^N(\mathbf{m}^N; \gamma^N, u) \geq J(\mathbf{m}_0; \rho, u)$  for every strategy modification  $u$ . It is therefore enough to show that

$$(7.1) \quad \lim_{N \rightarrow \infty} J_1^N(\mathbf{m}^N; \gamma^N, u) = J(\mathbf{m}_0; \rho, u) \quad \text{for every } u \in \mathcal{U}.$$

We proceed in three steps.

**Step One.** For  $N \in \mathbb{N}$ , set

$$\gamma_m^N \doteq \otimes^N \rho_1(\cdot | m), \quad m \in \mathcal{P}(\mathcal{X})^{T+1}.$$

Then, for every strategy modification  $u \in \mathcal{U}$ ,

$$J_1^N(\mathbf{m}^N; \gamma^N, u) = \int_{\mathcal{P}(\mathcal{X})^{T+1}} J_1^N(\mathbf{m}^N; \gamma_m^N, u) \rho_2(dm),$$

and also

$$J(\mathbf{m}_0; \rho, u) = \int_{\mathcal{P}(\mathcal{X})^{T+1}} J(\mathbf{m}_0; \rho_1(\cdot | m) \otimes \delta_m, u) \rho_2(dm).$$

To see this, recall that if we have a realization of  $(\mathbf{m}^N, \gamma^N, u)$  for player 1 in the  $N$ -player game or of  $(\mathbf{m}_0, \rho, u)$  in the mean field game, then the sequence of noise variables, the initial states (or state), and the random elements realizing the correlated profile  $\gamma^N$  (or the correlated flow  $\rho$ ) are independent.

**Step Two.** Fix a strategy modification  $u$ . Let  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ . As in the proof of Theorem 6.1, let  $(\Phi_m, X_m(\cdot), \xi(\cdot), \mu_m(\cdot))$  and  $(\tilde{\Phi}_m, \tilde{X}_m(\cdot), \tilde{\xi}(\cdot), \tilde{\mu}_m(\cdot))$  be the distributional limit along a convergent subsequence of realizations of  $(\mathbf{m}^N, \gamma_m^N, \text{Id})$  and  $(\mathbf{m}^N, \gamma_m^N, u)$ , respectively. Then properties (a)–(e) there hold, for both  $(\Phi_m, X_m(\cdot), \xi(\cdot), \mu_m(\cdot))$  and  $(\tilde{\Phi}_m, \tilde{X}_m(\cdot), \tilde{\xi}(\cdot), \tilde{\mu}_m(\cdot))$ . By construction and Step Three in the proof of Theorem 6.1, we also have

$$\mathbf{P}_m \circ (\Phi_m, X_m(0), \xi(\cdot), \mu_m(\cdot))^{-1} = \tilde{\mathbf{P}}_m \circ (\tilde{\Phi}_m, \tilde{X}_m(0), \tilde{\xi}(\cdot), \tilde{\mu}_m(\cdot))^{-1}$$

and  $\mathbf{P}_m \circ \Phi_m^{-1} = \tilde{\mathbf{P}}_m \circ \tilde{\Phi}_m^{-1} = \rho_1(\cdot | m)$ . Moreover, by (A1), (A5) and thanks to the fact that  $\gamma_m^N$  is the  $N$ -fold product of  $\rho_1(\cdot | m)$ , propagation of chaos holds for the convergent subsequence corresponding to  $(\mathbf{m}^N, \gamma_m^N, \text{Id})$  in the sense that

$$\mathbf{P}_m \circ (X_m(t), \mu_m(t))^{-1} = \hat{m}_m(t) \otimes \delta_{\hat{m}_m(t)}, \quad t \in \{0, \dots, T\},$$

for some deterministic flow of measures  $\hat{m}_m \in \mathcal{P}(\mathcal{X})^{T+1}$  with  $\hat{m}_m(0) = \mathbf{m}_0$ ; see, for instance, Theorem 4.2 in Gottlieb [1998]. In view of property (d), we therefore have  $\mathbf{P}_m$ -almost surely,

$$(7.2) \quad \begin{aligned} X_m(t+1) &= \Psi(t, X_m(t), \hat{m}_m(t), \Phi_m(t, X_m(t)), \xi(t+1)), \\ \mathbf{P}_m \circ X_m(t)^{-1} &= \hat{m}_m(t), \quad t \in \{0, \dots, T-1\}. \end{aligned}$$

Using the independence properties (b) and (c), we see by induction over the time variable  $t$  that Equation (7.2), together with the initial distribution  $\mathbf{P}_m \circ X_m(0)^{-1} = \mathbf{m}_0$  and the distribution  $\mathbf{P}_m \circ \Phi_m^{-1} = \rho_1(\cdot | m)$ , uniquely determines the (deterministic) flow of measures  $\hat{m}_m$ . This can be seen as a uniqueness property for a kind of McKean-Vlasov equation.

**Step Three.** We are going to show that  $\hat{m}_m = m$  for  $\rho_2$ -almost every  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ , where  $\hat{m}_m$  is the deterministic flow of measures identified in Step Two.

Let  $((\Omega, \mathcal{F}, \mathbf{P}^*), \Phi^*, X^*(\cdot), \mu^*(\cdot), \xi^*(\cdot))$  be a realization of the triple  $(\mathbf{m}_0, \rho, \text{Id})$ . The quintuple thus satisfies the dynamics given by Eq. (4.1), that is,  $\mathbf{P}^*$ -almost surely, for every  $t \in \{0, \dots, T-1\}$ ,

$$X^*(t+1) = \Psi(t, X^*(t), \mu^*(t), \Phi^*(t, X^*(t)), \xi^*(t+1)).$$

Moreover,  $\mathbf{P}^* \circ (\Phi^*, \mu^*(0), \dots, \mu^*(T))^{-1} = \rho$  and  $\mathbf{P}^* \circ (X^*(0))^{-1} = \mathbf{m}_0$ . By hypothesis,  $\rho$  is a correlated solution with initial distribution  $\mathbf{m}_0$ . In view of the consistency property, conditioning on  $\mu^*$  therefore yields for  $\rho_2$ -almost every  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ ,

$$\begin{aligned} \mathbf{P}^*(X^*(t) \in \cdot \mid \mu^* = m) &= m(t)(\cdot), & t \in \{0, \dots, T-1\}, \\ \mathbf{P}^*(\Phi^* \in \cdot \mid \mu^* = m) &= \rho_1(\cdot \mid m). \end{aligned}$$

Since the noise variables  $\xi^*(t)$ ,  $t \in \{1, \dots, T\}$ , are i.i.d. with common distribution  $\nu$  and  $\xi^*(\cdot)$ ,  $X^*(0)$ , and  $(\Phi^*, \mu^*(\cdot))$  are independent, it follows that for  $\rho_2$ -almost every flow  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ , the triple  $(X^*(\cdot), \Phi^*, \xi^*(\cdot))$  solves Equation (7.2)  $\mathbf{P}^*(\cdot \mid \mu^* = m)$ -almost surely with deterministic flow of measures  $\hat{m}_m = m$ . Uniqueness of solutions for Equation (7.2) now entails that

$$\hat{m}_m = m \quad \text{and} \quad \mathbf{P}_m \circ (\Phi_m, \mu_m)^{-1} = \tilde{\mathbf{P}}_m \circ (\tilde{\Phi}_m, \tilde{\mu}_m)^{-1} = \rho_1(\cdot \mid m) \otimes \delta_m$$

for  $\rho_2$ -almost every  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ . This also shows that, given a ( $\rho_2$ -typical) flow of measures  $m \in \mathcal{P}(\mathcal{X})^{T+1}$ , any convergent subsequence of realizations of  $(\mathbf{m}^N, \gamma_m^N, \text{Id})$  has the same limit in distribution, and analogously for realizations of  $(\mathbf{m}^N; \gamma_m^N, u)$ .

Convergence of costs according to property (e) and integration against  $\rho_2$  according to Step One, in conjunction with dominated convergence, finish the proof.  $\square$

## A Auxiliary results

Here, we collect some auxiliary results, mostly elementary, regarding weak convergence and exchangeable triangular arrays. We refer to Billingsley [1968] for the theory of weak convergence of probability measures.

Let  $\mathcal{Y}, \mathcal{Z}$  be Polish spaces. For  $n \in \mathbb{N}$ , let  $Y_n, Z_n$  be random variables on  $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$  with values in  $\mathcal{Y}$  and  $\mathcal{Z}$ , respectively.

**Lemma A.1.** *Let  $\Psi: \mathcal{Z} \rightarrow \mathcal{Y}$  be measurable. Suppose that  $(Y_n, Z_n)$  converges in distribution to  $(Y, Z)$  as  $n \rightarrow \infty$  for some  $\mathcal{Y} \times \mathcal{Z}$ -valued random variable  $(Y, Z)$  defined on  $(\Omega, \mathcal{F}, \mathbf{P})$ .*

*If  $Y_n = \Psi(Z_n)$   $\mathbf{P}_n$ -almost surely for every  $n \in \mathbb{N}$  and if  $\Psi$  is continuous  $\mathbf{P} \circ Z^{-1}$ -almost everywhere, then  $Y = \Psi(Z)$   $\mathbf{P}$ -almost surely.*

*Proof.* The hypothesis that  $\Psi$  is continuous  $\mathbf{P} \circ Z^{-1}$ -almost everywhere implies that the mapping  $\mathcal{Y} \times \mathcal{Z} \ni (y, z) \mapsto (y, \Psi(z))$  is continuous  $\mathbf{P} \circ (Y, Z)^{-1}$ -almost everywhere. By the convergence assumption and the mapping theorem Billingsley [1968, Theorem I.5.1], it follows that

$$(Y_n, \Psi(Z_n)) \xrightarrow{n \rightarrow \infty} (Y, \Psi(Z)) \text{ in distribution.}$$

Let  $D \doteq \{(y, \tilde{y}) \in \mathcal{Y} \times \mathcal{Y} : y = \tilde{y}\}$  be the diagonal in  $\mathcal{Y} \times \mathcal{Y}$ . Then  $D$  is closed in  $\mathcal{Y} \times \mathcal{Y}$ , hence

$$\limsup_{n \rightarrow \infty} \mathbf{P}_n((Y_n, \Psi(Z_n)) \in D) \leq \mathbf{P}((Y, \Psi(Z)) \in D)$$

by the Portmanteau theorem [Billingsley, 1968, Theorem I.2.1]. On the other hand, we have  $\mathbf{P}_n((Y_n, \Psi(Z_n)) \in D) = 1$  for every  $n \in \mathbb{N}$  since  $Y_n = \Psi(Z_n)$   $\mathbf{P}_n$ -almost surely by hypothesis. It follows that  $\mathbf{P}((Y, \Psi(Z)) \in D) = 1$ , that is,  $Y = \Psi(Z)$   $\mathbf{P}$ -almost surely.  $\square$

**Lemma A.2.** *Let  $d_{\mathcal{Z}}$  be a metric compatible with the topology of  $\mathcal{Z}$ . Let  $(Y_n, Z_n, \tilde{Z}_n)_{n \in \mathbb{N}}$  be a sequence of  $\mathcal{Y} \times \mathcal{Z} \times \mathcal{Z}$ -valued random variables, where each  $(Y_n, Z_n, \tilde{Z}_n)$  is defined on some probability space  $(\Omega_n, \mathcal{F}_n, \mathbf{P}_n)$ ,  $n \in \mathbb{N}$ .*

*Suppose that  $(Y_n, Z_n)$  converges in distribution to  $(Y, Z)$  as  $n \rightarrow \infty$  for some  $\mathcal{Y} \times \mathcal{Z}$ -valued random variable  $(Y, Z)$  defined on  $(\Omega, \mathcal{F}, \mathbf{P})$ , and that*

$$\mathbf{E}_n \left[ d_{\mathcal{Z}}(Z_n, \tilde{Z}_n) \right] \xrightarrow{n \rightarrow \infty} 0.$$

*Then  $(Y_n, \tilde{Z}_n)$  converges in distribution to  $(Y, Z)$  as  $n \rightarrow \infty$ .*

*Proof.* Let  $d_{\mathcal{Y}}$  be any metric compatible with the topology of  $\mathcal{Y}$ . Set

$$d((y, z), (\tilde{y}, \tilde{z})) \doteq d_{\mathcal{Y}}(y, \tilde{y}) + d_{\mathcal{Z}}(z, \tilde{z}), \quad (y, z), (\tilde{y}, \tilde{z}) \in \mathcal{Y} \times \mathcal{Z}.$$

Then  $d$  is a metric on  $\mathcal{Y} \times \mathcal{Z}$  compatible with the product topology. By hypothesis and Markov's inequality, we have that  $(d((Y_n, Z_n), (Y_n, \tilde{Z}_n)))_{n \in \mathbb{N}}$  converges to zero in probability. As the limit is a constant, this is equivalent to convergence in distribution, and the underlying probability spaces may depend on  $n \in \mathbb{N}$ . The assertion now follows from Theorem I.4.1 in Billingsley [1968, p. 25].  $\square$

For the next result, let  $\kappa_n$  be a regular conditional distribution of  $Y_n$  given the  $\sigma$ -algebra generated by  $Z_n$ , each  $n \in \mathbb{N}$ . Thus,  $\kappa_n$  is a mapping  $\Omega_n \times \mathcal{B}(\mathcal{Y}) \rightarrow [0, 1]$  that induces a  $\mathcal{P}(\mathcal{Y})$ -valued random variable and is such that, for every  $A \in \mathcal{B}(\mathcal{Y})$ ,

$$\mathbf{P}_n(Y_n \in A \mid Z_n) = \kappa_n(A) \quad \mathbf{P}_n\text{-almost surely.}$$

A regular conditional distribution of  $Y_n$  given  $Z_n$  exists since  $\mathcal{Y}$  is a Polish space, and it is uniquely determined with probability one when seen as a  $\mathcal{P}(\mathcal{Y})$ -valued random variable.

**Lemma A.3.** *Let  $\kappa_n$  be a regular conditional distribution of  $Y_n$  given  $\sigma(Z_n)$  as above. Suppose that  $(Y_n, Z_n, \kappa_n)$  converges in distribution to  $(Y, Z, \kappa)$  as  $n \rightarrow \infty$  for some  $\mathcal{Y} \times \mathcal{Z} \times \mathcal{P}(\mathcal{Y})$ -valued random variable  $(Y, Z, \kappa)$  defined on  $(\Omega, \mathcal{F}, \mathbf{P})$ . Then, for every  $A \in \mathcal{B}(\mathcal{Y})$ , every  $B \in \mathcal{B}(\mathcal{Z})$ ,*

$$\mathbf{P}(Y \in A, Z \in B) = \mathbf{E}[\mathbf{1}_B(Z) \cdot \kappa(A)].$$

*If, in addition,  $\kappa(A)$  is  $\sigma(Z)$ -measurable for every  $A \in \mathcal{B}(\mathcal{Y})$ , then  $\kappa$  is a regular conditional distribution of  $Y$  given  $\sigma(Z)$ .*

*Proof.* Let  $Q \in \mathcal{P}(\mathcal{Y} \times \mathcal{Z})$  be the joint law of  $Y$  and  $Z$ :  $Q \doteq \mathbf{P} \circ (Y, Z)^{-1}$ . Define another measure  $\tilde{Q} \in \mathcal{P}(\mathcal{Y} \times \mathcal{Z})$  by setting, for  $A \in \mathcal{B}(\mathcal{Y})$ ,  $B \in \mathcal{B}(\mathcal{Z})$ ,

$$\tilde{Q}(A \times B) \doteq \mathbf{E}[\mathbf{1}_B(Z) \cdot \kappa(A)].$$

Let  $g: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathbb{R}$  be bounded and measurable. Then, for every  $n \in \mathbb{N}$ ,

$$\mathbf{E}_n[g(Y_n, Z_n)] = \mathbf{E}_n \left[ \int_{\mathcal{Y}} g(y, Z_n) \kappa_n(dy) \right],$$

since  $\kappa_n$  is a version of the regular conditional distribution of  $Y_n$  given  $\sigma(Z_n)$  by hypothesis. If  $g$  is bounded and continuous, then, by convergence in distribution of  $(Y_n, Z_n)$  to  $(Y, Z)$ ,

$$\lim_{n \rightarrow \infty} \mathbf{E}_n[g(Y_n, Z_n)] = \mathbf{E}[g(Y, Z)],$$

but also, by convergence in distribution of  $(Z_n, \kappa_n)$  to  $(Z, \kappa)$ ,

$$\lim_{n \rightarrow \infty} \mathbf{E}_n \left[ \int_{\mathcal{Y}} g(y, Z_n) \kappa_n(dy) \right] = \mathbf{E} \left[ \int_{\mathcal{Y}} g(y, Z) \kappa(dy) \right]$$

since the mapping  $(z, m) \mapsto \int_{\mathcal{Y}} g(y, z) m(dy)$  is bounded and continuous on  $\mathcal{Z} \times \mathcal{P}(\mathcal{Y})$  if  $g$  is bounded and continuous on  $\mathcal{Y} \times \mathcal{Z}$ ; cf. Theorem I.5.5 in Billingsley [1968, p. 34].

Therefore, for every  $g: \mathcal{Y} \times \mathcal{Z} \rightarrow \mathbb{R}$  bounded and continuous,

$$\int_{\mathcal{Y} \times \mathcal{Z}} g dQ = \mathbf{E}[g(Y, Z)] = \mathbf{E} \left[ \int_{\mathcal{Y}} g(y, Z) \kappa(dy) \right] = \int_{\mathcal{Y} \times \mathcal{Z}} g d\tilde{Q}.$$

A measure on the Borel sets of a Polish space is uniquely determined by its integrals over all bounded continuous functions. It follows that  $Q = \tilde{Q}$ . This in turn implies that for all  $A \in \mathcal{B}(\mathcal{Y})$ , all  $B \in \mathcal{B}(\mathcal{Z})$ ,

$$\mathbf{E}[\mathbf{1}_A(Y) \cdot \mathbf{1}_B(Z)] = \mathbf{E}[\mathbf{1}_B(Z) \cdot \kappa(A)],$$

which yields the first part of the assertion. If, in addition,  $\kappa(A)$  is  $\sigma(Z)$ -measurable for every  $A \in \mathcal{B}(\mathcal{Y})$ , then

$$\mathbf{P}(Y \in A \mid Z) = \kappa(A) \quad \mathbf{P}\text{-almost surely}$$

by the above property and the definition of conditional expectation.  $\square$

**Lemma A.4.** *Let  $Y$  and  $\kappa$  be random variables on some probability space  $(\Omega, \mathcal{F}, \mathbf{P})$  with values in  $\mathcal{Y}$  and  $\mathcal{P}(\mathcal{Y})$ , respectively. Let  $\mathcal{C}, \mathcal{G}, \tilde{\mathcal{C}}$  be sub- $\sigma$ -algebras of  $\mathcal{F}$  such that  $\mathcal{C} \subseteq \mathcal{G} \subseteq \tilde{\mathcal{C}}$ .*

*If  $\kappa$  is a regular conditional distribution of  $Y$  given  $\mathcal{C}$  as well as given  $\tilde{\mathcal{C}}$ , then  $\kappa$  is also a regular conditional distribution of  $Y$  given  $\mathcal{G}$ .*

*Proof.* Suppose that  $\kappa$  is a regular conditional distribution of  $Y$  given  $\mathcal{C}$  as well as given  $\tilde{\mathcal{C}}$ . Let  $A \in \mathcal{B}(\mathcal{Y})$ . The first part of the assumption implies that  $\kappa(A)$  is  $\mathcal{C}$ -measurable, hence also  $\mathcal{G}$ -measurable (since  $\mathcal{C} \subseteq \mathcal{G}$ ). The second part of the assumption entails that

$$\mathbf{P}(\{Y \in A\} \cap C) = \mathbf{E}[\mathbf{1}_C \cdot \kappa(A)]$$

for all  $C \in \tilde{\mathcal{C}}$ , hence also for all  $C \in \mathcal{G}$  (since  $\mathcal{G} \subseteq \tilde{\mathcal{C}}$ ). This shows that  $\kappa$  is a regular conditional distribution of  $Y$  given  $\mathcal{G}$ .  $\square$

The next result recalls the conditional distribution of an element of a finite exchangeable sequence given the associated empirical measure.

**Lemma A.5.** *Let  $Y_1, \dots, Y_N$  be a finite exchangeable sequence of  $\mathcal{Y}$ -valued random variables on some probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ , and let*

$$\mu_\omega^N \doteq \frac{1}{N} \sum_{i=1}^N \delta_{Y_i(\omega)}, \quad \omega \in \Omega,$$

*be the associated empirical measure. Then, for every  $i \in \{1, \dots, N\}$ ,*

$$\mathbf{P}(Y_i \in \cdot \mid \mu^N) = \mu^N(\cdot)$$

*in the sense that  $\mu^N$  is a regular conditional distribution of  $Y_i$  given (the  $\sigma$ -algebra generated by)  $\mu^N$ .*

*Proof.* The assertion follows from Lemma 11.11 in Kallenberg [2001, p. 213].  $\square$

## References

- R.J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.
- R.J. Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55:1–18, 1987.
- I. Bárány. Fair distribution protocols or how the players replace fortune. *Mathematics of Operations Research*, 17.2:327–340, 1992.

- P. Billingsley. *Convergence of Probability Measures*. Wiley series in Probability and Statistics. John Wiley & Sons, New York, 1968.
- P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games*: (AMS-201). Vol. 201. Princeton University Press, 2019.
- R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM J. Control Optim.*, 51(4):2705–2734, 2013.
- R. Carmona and F. Delarue. Probabilistic Theory of Mean Field Games with Applications I-II. Springer Nature, 2018.
- F. Forges. Correlated equilibria and communication in games. In: R.A. Meyers (ed.), *Computational Complexity*. Springer, New York, NY, 2012.
- I. Gilboa and E. Zemel. Nash and correlated equilibria: Some complexity considerations. *Games and Economic Behavior*, 1(1):80–93, 1989.
- D. Gomes, J. Mohr, and R.R. Souza. Continuous time finite state mean field games. *Appl. Math. Optim.*, 68:99–143, 2013.
- A. D. Gottlieb. *Markov transitions and the propagation of chaos*. PhD thesis, Lawrence Berkeley National Laboratory, 1998.
- S. Hart. Adaptive heuristics. *Econometrica*, 73(5):1401–1430, 2005.
- S. Hart and D. Schmeidler. Existence of correlated equilibria. *Mathematics of Operations Research*, 14(1):18–25, 1989.
- M. Huang, R.P. Malhamé, and P.E. Caines. Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–252, 2006.
- O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, New York, 2nd edition, 2001.
- D. Lacker. On the convergence of closed-loop Nash equilibria to the mean field game limit. *Annals of Applied Probability*, 30(4):1693–1761, 2020.
- J.-M. Lasry and P.-L. Lions. Mean field games. *Japan. J. Math.*, 2(1):229–260, 2007.
- T. Roughgarden. Twenty lectures on algorithmic game theory. Cambridge University Press, 2016.

- E. Solan. Rationality and Extensive Form Correlated Equilibria in Stochastic Games. D.P. 1298. The Center for Mathematical Studies in Economics and Management Science, Northwestern University, 2000.
- E. Solan. Characterization of correlated equilibria in stochastic games. *International Journal of Game Theory*, 30(2):259–277, 2001.
- E. Solan and N. Vieille. Correlated equilibrium in stochastic games. *Games and Economic Behavior*, 38(2):362–399, 2002.