

A PROBABILISTIC APPROACH TO EXTENDED FINITE STATE MEAN FIELD GAMES

RENÉ CARMONA AND PEIQI WANG

ABSTRACT. We develop a probabilistic approach to continuous-time finite state mean field games. Based on an alternative description of continuous-time Markov chain by means of semimartingale and the weak formulation of stochastic optimal control, our approach not only allows us to tackle the mean field of states and the mean field of control in the same time, but also extend the strategy set of players from Markov strategies to closed-loop strategies. We show the existence and uniqueness of Nash equilibrium for the mean field game, as well as how the equilibrium of mean field game consists of an approximative Nash equilibrium for the game with finite number of players under different assumptions of structure and regularity on the cost functions and transition rate between states.

1. INTRODUCTION

Mean field game in which players' states belong to a finite space is first studied in [14]. The dynamics of each player's states is depicted by a continuous-time Markov chain, whose transition rate matrix is a function of the player's control and probability distribution of all players' states. By assuming that each player adopts a Markovian strategy, the Nash equilibrium can be characterized by a HJB equation corresponding to the optimal control of continuous-time Markov chain on the one hand, and a Kolmogorov equation on how probability distribution of player's states evolves on the other hand. Due to the finite nature of the state space, both equations turn out to be ordinary differential equations and existence of the solution to this forward-backward system can be obtained by a fixed point argument. Continuous-time finite state mean field games were applied to model socio-economic phenomena such as paradigm shift in a scientific community and consumer choice in [15]. In [19], the strategic aspect of cyber attack and defense is analyzed through a finite state mean field game model, in which the author introduces a major player - the hacker - whose action influences each minor player - the computer user - in terms of their payoff and dynamics. Theoretical aspects of finite state mean field games with major and minor players are investigated in [4] where existence of Nash equilibria and results on approximate Nash equilibrium for finite player game in small duration are obtained, along with the master equation characterizing the Nash equilibrium.

In this paper, we develop a probabilistic framework for continuous-time finite state mean field game. Our starting point is a semimartingale representation of continuous-time Markov chain introduced in [10]: Let $(X_t)_{0 \leq t \leq T}$ be a continuous-time Markov chain with m states which are identified with the m standard basis vectors in $\mathbb{R}^m$, then we can write:

$$X_t = X_0 + \int_{(0,t]} Q^*(t) \cdot X_{t-} dt + \mathcal{M}_t.$$

where $Q(t)$ is the transition rate matrix (also known as the Q-matrix) with $Q^*(t)$ being its transpose and $\mathcal{M}$ is a martingale. We immediately notice the analogy with diffusion processes and apply Girsanov Theorem to construct equivalent probability measures under which the process X admits a different transition rate process. This opens a pathway to formulating the optimal control problem of continuous-time Markov chain in a so-called *weak* fashion. Indeed, in the context of optimal control of diffusion processes, the weak formulation links the control of the drift to the control of the probability measure (as opposed to the control of the path) and identifies the value function of the control problem as the solution to a backward stochastic differential equation (BSDE). By the comparison principle of the BSDE, the optimality of the control problem can be obtained by optimizing the driver of the BSDE which coincides with the Hamiltonian function. It turns out that such procedure can be transplanted to the case of

optimal control of continuous-time Markov chain, thanks to the theory of BSDE driven by Markov chain developed in [6] and [7].

Once the optimal control problem can be characterized by a BSDE, our next step is to develop a probabilistic approach to the mean field game. Probabilistic approach to mean field game is first proposed in [2], where player's optimization problem is treated in the *strong* formulation. By applying Pontryagin's Maximum Principle, the optimality of player's control problem is characterized by a forward-backward stochastic differential equation (FBSDE). Later in [3], the authors consider the weak formulation of control problem and use the argument of change of measure which we briefly described above to obtain the BSDE characterizing the optimality. In both cases, the existence of Nash equilibria of the mean field game boils down to the well-posedness of a BSDE (or FBSDE) in which the probability distribution of the solution enters into the driver and the terminal condition of the equation. These are the so-called McKean-Vlasov type of BSDE (or FBSDE) for which the existence of the solution can be obtained by a fixed-point argument à la Schauder.

By developing the weak formulation, our contributions to finite state mean field game are three-fold. On the one hand, the flexibility of the probabilistic approach allows us to incorporate not only the mean field of state, but also the mean field of control into the dynamics and cost functionals of individual players. Mean field of control is known to be notoriously intractable via PDE method, due to the difficulties in deriving the equation obeyed by the flow of probability measure of the optimal control. Under the probabilistic framework however, the mean field of state and the mean field of control can be dealt with in similar manners, although the treatment of mean field of control is more involved in terms of the topological argument. On the other hand, using the weak formulation we are able to show Nash equilibria exist among all closed-loop strategies, including the strategies depending on the past history of player's states, whereas the PDE approach can only accommodate Markovian strategies.

Lastly, the weak formulation we develop for the finite state mean field game will serve as a launching pad to tackle the finite state mean field agent-principal problem. Such model is a form of Stackelberg game in which the principal fixes a contract first and a large population of agents reaches Nash equilibrium according to the contract proposed by the principal. By fixing a contract we actually mean that the principal chooses a control which enters into each agent's dynamics and cost functions. One meaningful direction in probing mean field agent-principal problems is to understand how the principal can choose the optimal contract so that its own cost function depending on agent's distribution is minimized. To the best of our knowledge, this type of problem is first investigated in [9] where the agent's dynamics is a diffusion. The main idea is to formulate the optimal contract problem as a McKean-Vlasov optimal control problem, in which the state process to be controlled is the McKean-Vlasov BSDE characterizing the Nash equilibrium in the weak formulation of the mean field game. With the help of the weak formulation we develop in this paper, we believe that the same technique can be applied to the case of finite state mean field agent-principal problem, which could lead to potential applications in epidemics and cyber security.

We would also like to mention a few literatures related to our paper. In [5] the authors proposed a probabilistic framework for finite state mean field game where the player's dynamics of states is represented by stochastic differential equations driven by Poisson random measures. By using Ky Fan's fixed point theorem, the authors obtained existence and uniqueness of the Nash equilibrium in relaxed open-loop as well as relaxed feedback controls. Then under stronger assumption that guarantees uniqueness of optimal non-relaxed feedback control, the authors deduced existence of Nash equilibria in non-relaxed feedback form. In [8], continuous-time mean field games with finite state space and finite action space were studied. The authors proved existence of Nash equilibrium among relaxed feedback controls. In [1] the authors investigated mean field games where each player's state follows a jump-diffusion process and the player controls the sizes of the jumps. The formulation is based on weak formulation of stochastic controls and martingale problems. Existence of Nash equilibrium among relaxed controls and Markovian controls is established.

The rest of the paper is organized as follows. In Section 2, we introduce the weak formulation of finite state mean field game, which is based on a semimartingale representation of continuous-time Markov chain and an argument of change of measure. We state the assumptions used throughout the paper and give the precise definition of the Nash equilibrium in the weak formulation. In Section 3, we analyze player's optimal control problem when facing a fixed mean field of state and control, by characterizing the value function and the optimal control using a BSDE driven by Markov chain. Section 4 is devoted to the existence and the uniqueness of the Nash equilibrium. Finally in Section 5, we formulate the game with finite number of players and show the Nash equilibrium of the mean field game is an approximate Nash equilibrium of the game with finite number of players.

2. THE WEAK FORMULATION FOR FINITE STATE MEAN FIELD GAMES

2.1. Notations. If M is a square real matrix, we denote by M^* its transpose and M^+ its Moore-Penrose pseudo inverse. For a column vector X , we denote by $\text{diag}(X)$ the square diagonal matrix whose diagonal elements are given by the entries of X . If γ is a random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we denote its law or its distribution, namely the push-forward of $\mathbb{P}$ by γ by $\mathbb{P}_{\# \gamma} := \mathbb{P} \circ \gamma^{-1}$.

For two square integrable martingales L, M , we denote by $[L, M]$ the quadratic covariation process of L and M . For two semimartingales L and M , we denote by $\langle L, M \rangle$ the predictable quadratic covariation process of L and M . For a semimartingale L such that $L_0 = 0$, we denote by $\mathcal{E}(L)$ the process of Doléans-Dade exponential of L . See Chapter II.6 in [21] for the definitions of these standard concepts.

2.2. Controlled probability measure. For the control of continuous-time finite state Markov chains we adopt the formalism first introduced in [10], and later developed in [6] and [7]. If $\mathbf{X} = (X_t)_{0 \leq t \leq T}$ is a continuous-time Markov chain with m states, we identify these states with the basis vectors e_i in $\mathbb{R}^m$ and we denote by E the resulting state space $E = \{e_1, \dots, e_m\}$. We assume that the sample paths $t \rightarrow X_t$ are *càdlàg*, i.e. right continuous with left limits, and continuous at T . In other words, we force $X_{T-} = X_T$.

We first construct a canonical probability space for $\mathbf{X}$. Let Ω be the space of *càdlàg* functions from $[0, T]$ to E which are continuous at T , and let $\mathbf{X}$ be the canonical process on Ω , that is $X_t(\omega) := \omega_t$. We denote by $\mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}$ with $\mathcal{F}_t := \sigma\{X_s, s \leq t\}$ the natural filtration generated by $\mathbf{X}$, and we set $\mathcal{F} := \mathcal{F}_T$. Throughout the rest of the paper, we fix a probability measure $\mathbf{p}^0$ on the set E . It will be used as the initial distribution of the process $\mathbf{X}$. On the filtered space $(\Omega, \mathbb{F}, \mathcal{F})$, we consider the probability measure $\mathbb{P}$ under which $\mathbf{X}$ is a continuous-time Markov chain with initial distribution $\mathbf{p}^0$ and transition rates between any two different states equal to 1. This means that for $i, j \in \{1, \dots, m\}$, $i \neq j$ and $\Delta t > 0$, we have $\mathbb{P}[X_{t+\Delta t} = e_j | \mathcal{F}_t] = \mathbb{P}[X_{t+\Delta t} = e_j | X_t]$ and $\mathbb{P}[X_{t+\Delta t} = e_j | X_t = e_i] = \Delta t + o(\Delta t)$. By Appendix B in [10], the process X has the representation:

$$(1) \quad X_t = X_0 + \int_{(0, t]} Q^0 \cdot X_{t-} dt + \mathcal{M}_t,$$

where Q^0 is the square matrix with diagonal elements all equal to $-(m-1)$ and off-diagonal elements all equal to 1, and

$bcM = (\mathcal{M}_t)_{t \geq 0}$ is a $\mathbb{R}^m$ -valued $\mathbb{P}$ -martingale. The multiplication $\cdot$ is understood as matrix multiplication. Indeed, Q^0 is the transition rate matrix of $\mathbf{X}$ under the probability measure $\mathbb{P}$.

Remark 2.1. The representation originally proposed in [10] is:

$$X_t = X_0 + \int_{(0, t]} Q^0 \cdot X_t dt + \mathcal{M}_t.$$

However since X_t is only discontinuous on a countable set, we can replace X_t by X_{t-} in the integral. The reason for this slight change of representation is to make the integrand a predictable process, which will be suitable for the change of measure argument in what follows.

We shall refer to the probability measure $\mathbb{P}$ as the *reference measure* on the sample space. The first step of the weak formulation of mean field game consists in depicting how each player's control as well as the mean field determine the probability measure of the sample path. We denote by $\mathcal{S}$ the m -dimensional simplex:

$$\mathcal{S} := \{p \in \mathbb{R}^m; \sum_{i=1}^m p_i = 1, p_i \geq 0\},$$

which we identify with the space of probability distributions on E . Let A be a compact subset of $\mathbb{R}^l$ from which the players can choose their controls. Denote by $\mathcal{P}(A)$ the space of probability measures on A . We introduce a function q :

$$[0, T] \times \{1, \dots, m\}^2 \times A \times \mathcal{S} \times \mathcal{P}(A) \rightarrow q(t, i, j, \alpha, p, \nu),$$

and we denote by $Q(t, \alpha, p, \nu)$ the matrix $[q(t, i, j, \alpha, p, \nu)]_{1 \leq i, j \leq m}$. Throughout the rest of the paper, we make the following assumption on q :

Assumption 2.2. (i) For all $(t, \alpha, p, \nu) \in [0, T] \times A \times \mathcal{S} \times \mathcal{P}(A)$, the matrix $Q(t, \alpha, p, \nu)$ is a Q -matrix.
(ii) There exist constants $C_1, C_2 > 0$ such that for all $(t, i, j, \alpha, p, \nu) \in [0, T] \times E^2 \times A \times \mathcal{S} \times \mathcal{P}(A)$ such that $i \neq j$, we have $0 < C_1 < q(t, i, j, \alpha, p, \nu) < C_2$.
(iii) There exists a constant $C > 0$ such that for all $(t, i, j) \in [0, T] \times E^2$, $\alpha, \alpha' \in A$, $p, p' \in \mathcal{S}$ and $\nu, \nu' \in \mathcal{P}(A)$, we have:

$$|q(t, i, j, \alpha, p, \nu) - q(t, i, j, \alpha', p', \nu')| \leq C(\|\alpha - \alpha'\| + \|p - p'\| + \mathcal{W}_1(\nu, \nu')).$$

where $\mathcal{W}_1$ denotes the 1-Wasserstein distance between probability measures on A .

Recall that a matrix $Q = [Q_{ij}]$ is called a Q -matrix if $Q_{ij} \geq 0$ for $i \neq j$ and

$$\sum_{j \neq i} Q_{ij} = -Q_{ii}, \quad \text{for all } i.$$

Remark 2.3. Assumption 2.2 is analog to the non-degeneracy condition in the diffusion-based mean field game models. It guarantees that the probability measure $\mathbb{Q}^{(\alpha, p, \nu)}$ defined in (7) below, is equivalent to the reference measure $\mathbb{P}$. In some applications of continuous-time Markov chain models, it happens that jumps from some states to others are forbidden, in which case the transition rate function q would satisfy $q(t, i, j, \alpha, p, \nu) \equiv 0$ for some couples (i, j) . For example, this is the case in the botnet defense model proposed by [19], as well as in the extended version of the model which includes an attacker studied in [4]. When that happens, we need to use a different reference probability measure $\mathbb{P}$: we set the transition rate to 1 for all the jumps, except for those that are forbidden, for which we set the transition rate to 0. Fortunately, this is the only modification we need to make in order to accommodate this kind of special case. The arguments presented in the following can be trivially extended to be compatible with this modified reference probability.

We state without proof a useful property of the martingale $\mathcal{M}$. The proof of this result can be found in [6]:

Lemma 2.4. The predictable quadratic variation of the martingale $\mathcal{M}$ under $\mathbb{P}$ is given by the formula:

$$(2) \quad \langle \mathcal{M}, \mathcal{M} \rangle_t = \int_0^t \psi_t dt,$$

where ψ_t is given by:

$$(3) \quad \psi_t := \text{diag}(Q^0 \cdot X_{t-}) - Q^0 \cdot \text{diag}(X_{t-}) - \text{diag}(X_{t-}) \cdot Q^0.$$

If we define for each i the matrix ψ^i by:

$$\psi^i := \text{diag}(Q^0 \cdot e_i) - Q^0 \cdot \text{diag}(e_i) - \text{diag}(e_i) \cdot Q^0,$$

then clearly we have $\psi_t = \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i) \psi^i$. Since each ψ^i is a semi-definite positive matrix, so is ψ_t . We define the corresponding (stochastic) seminorm $\|\cdot\|_{X_{t-}}$ on $\mathbb{R}^m$ by:

$$(4) \quad \|Z\|_{X_{t-}}^2 := Z^* \cdot \psi_t \cdot Z.$$

The semi-norm $\|\cdot\|_{X_{t-}}$ can be rewritten in a more explicit way. For $i \in \{1, \dots, m\}$, let us define the seminorm $\|\cdot\|_{e_i}$ on $\mathbb{R}^m$ by $\|Z\|_{e_i}^2 := Z^* \cdot \psi^i \cdot Z = \sum_{j \neq i} |Z_j - Z_i|^2$. Then it is easy to see that $\|Z\|_{X_{t-}} = \sum_{i=1}^m \mathbb{1}(X_{t-} = i) \|Z\|_{e_i}$.

Since ψ_t is symmetric, we have $(\psi_t^+)^* = \psi_t^+$. Recall that ψ_t^+ is the Moore-Penrose generalized inverse of the matrix ψ_t . On the other hand, it is straightforward to verify that for all $t \in [0, T]$ and $w \in \Omega$, the range of the matrix ψ_t (i.e. the linear space spanned by the columns of ψ_t) is the space $\{q \in \mathbb{R}^m; \sum_{i=1}^m q_i = 0\}$. Therefore for all $q \in \mathbb{R}^m$ with $\sum_{i=1}^m q_i = 0$, we have $\psi_t \cdot \psi_t^+ \cdot q = q$. This holds in particular for any row vector from any Q -matrix, or any vector of the form $(e_j - e_i)$.

In order for the paper to be as self-contained as possible, we also recall the following version of Girsanov Theorem on change of probability measure. See Theorem III.41 in [21] or Lemma 4.3 in [23].

Theorem 2.5. *Let $T > 0$ and $\mathbf{L} = (L_t)_{t \geq 0}$ be a martingale defined on $[0, T]$ with $\Delta L_t \geq -1$. Assume that the Doléans-Dade exponential $\mathcal{E}(\mathbf{L})$ of $\mathbf{L}$ is a uniformly integrable martingale and let $\mathbb{Q}$ be the probability measure having Radon-Nikodym derivative $\mathcal{E}(\mathbf{L})_T$ with respect to $\mathbb{P}$. If the quadratic covariation process $[\mathbf{M}, \mathbf{L}]$ is integrable under $\mathbb{P}$, then $\mathbf{M} - \langle \mathbf{M}, \mathbf{L} \rangle$ is a martingale under $\mathbb{Q}$, where the predictable quadratic covariation $\langle \mathbf{M}, \mathbf{L} \rangle$ is computed under the measure $\mathbb{P}$.*

We now describe how the control of a player and the mean field affect the probability law of $\mathbf{X}$. Let us define the player's strategy set $\mathbf{A}$ to be the collection of $\mathbb{F}$ -predictable processes $\alpha = (\alpha_t)_{t \in [0, T]}$ such that $\alpha_t \in A$ for $t \in [0, T]$. Given a flow of probability measures $\mathbf{p} = (p_t)_{t \in [0, T]}$ on E , and a flow of probability measures $\nu = (\nu_t)_{t \in [0, T]}$ on A , we define the scalar martingale $\mathbf{L}^{(\alpha, \mathbf{p}, \nu)}$ under $\mathbb{P}$ by:

$$(5) \quad L_t^{(\alpha, \mathbf{p}, \nu)} := \int_0^t X_{s-}^* \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \psi_s^+ \cdot d\mathcal{M}_s.$$

Clearly, the jumps of this are given by:

$$(6) \quad \Delta L_t^{(\alpha, \mathbf{p}, \nu)} = X_{t-}^* \cdot (Q(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot \psi_t^+ \cdot \Delta X_t.$$

One can easily check that $\psi_t^+ \cdot (e_j - X_{t-}) = \frac{m-1}{m} e_j - \sum_{i \neq j} \frac{1}{m} e_i$ when $X_{t-} = e_i \neq e_j$. Therefore when $X_{t-} = e_i \neq e_j = X_t$, we have:

$$\begin{aligned} \Delta L_t^{(\alpha, \mathbf{p}, \nu)} &= X_{t-}^* \cdot (Q(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot \psi_t^+ \cdot (e_j - X_{t-}) \\ &= e_i^* \cdot (Q(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot \left[\frac{m-1}{m} e_j - \sum_{k \neq j} \frac{1}{m} e_k \right] \\ &= \frac{m-1}{m} (q(t, i, j, \alpha_t, p_t, \nu_t) - q_{i,j}^0) - \frac{1}{m} \sum_{k \neq j} (q(t, i, k, \alpha_t, p_t, \nu_t) - q_{i,k}^0) \\ &= q(t, i, j, \alpha_t, p_t, \nu_t) - q_{i,j}^0 \\ &= q(t, i, j, \alpha_t, p_t, \nu_t) - 1, \end{aligned}$$

where the last equality is due to the fact that $\sum_{k=1}^m (q(t, i, k, \alpha_t, p_t, \nu_t) - q_{i,k}^0) = 0$. Therefore we have $\Delta L_t^{(\alpha, \mathbf{p}, \nu)} \geq -1$. By Theorem III.45 in [21] and the remark that follows, in order to show that $\mathcal{E}(\mathbf{L}^{(\alpha, \mathbf{p}, \nu)})$ is uniformly integrable, it suffices to show $\mathbb{E}[\exp(\langle \mathbf{L}^{(\alpha, \mathbf{p}, \nu)}, \mathbf{L}^{(\alpha, \mathbf{p}, \nu)} \rangle_T)] < \infty$. This is straightforward since we have:

$$\langle \mathbf{L}^{(\alpha, \mathbf{p}, \nu)}, \mathbf{L}^{(\alpha, \mathbf{p}, \nu)} \rangle_T$$

$$\begin{aligned}
&= \int_0^T X_{s-}^* \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \psi_s^+ \cdot \frac{d\langle \mathcal{M}, \mathcal{M} \rangle_s}{ds} \cdot (X_{s-}^* \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \psi_s^+)^* ds \\
&= \int_0^T X_{s-}^* \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \psi_s^+ \cdot (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} ds.
\end{aligned}$$

and the integrand is bounded by some constant by Assumption 2.2.

We now apply Girsanov's Theorem. It is straightforward to obtain that:

$$\begin{aligned}
\langle \mathcal{M}, \mathbf{L}^{(\alpha, \mathbf{p}, \nu)} \rangle_t &= \int_0^t d\langle \mathcal{M}, \mathcal{M} \rangle_s \cdot (\psi_s^+)^* \cdot (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} \\
&= \int_0^t \psi_s \cdot \psi_s^+ \cdot (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} ds \\
&= \int_0^t (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} ds.
\end{aligned}$$

In the last equality, we use the fact that $(Q^*(s, \alpha_s, p_s, \nu_s) - Q^0)X_s$ is the difference between two row vectors coming from Q -matrices, therefore is invariant by $\psi_s \cdot \psi_s^+$. Let us define the probability measure $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ by:

$$(7) \quad \frac{d\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}}{d\mathbb{P}} := \mathcal{E}(\mathbf{L}^{(\alpha, \mathbf{p}, \nu)})_T.$$

By Theorem 2.5, we know that the process $\mathcal{M}^{(\alpha, \mathbf{p}, \nu)}$, defined as:

$$(8) \quad \mathcal{M}_t^{(\alpha, \mathbf{p}, \nu)} := \mathcal{M}_t - \int_0^t (Q^*(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot X_{s-} ds,$$

is a $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ -martingale. Therefore the canonical decomposition (1) of X under $\mathbb{P}$ can be rewritten as:

$$(9) \quad X_t = X_0 + \int_0^t Q^*(s, \alpha_s, p_s, \nu_s) \cdot X_{s-} dt + \mathcal{M}_t^{(\alpha, \mathbf{p}, \nu)}.$$

This means that under the measure $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$, the stochastic intensity rate of $\mathbf{X}$ is given by $Q(t, \alpha_t, p_t, \nu_t)$. In addition, since $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ and $\mathbb{P}$ coincides on $\mathcal{F}_0$, the law of X_0 under $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ is the same as under the reference measure $\mathbb{P}$, which is $\mathbf{p}^\circ$. In particular, when α is a Markov control, i.e. of the form $\alpha_t = \phi(t, X_{t-})$ for some measurable function ϕ , $\mathbf{X}$ becomes a continuous-time Markov chain with intensity rate $q(t, i, j, \phi(t, i), p_t, \nu_t)$ under the measure $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$.

Remark 2.6. In the optimal control literature, admissible controls are often classified into the categories of open-loop controls and closed-loop controls. Open-loop controls are often referred to controls adapted to the underlying filtration, which is often generated by the noise process. Closed-loop controls, on the other hand, are controls that are adapted to the filtration generated by the history of the state process. In our set up, however, we see that the underlying filtration is indeed the one generated by the past path of the state process. Therefore this difference vanishes.

2.3. Weak formulation of mean field games. Let $f : [0, T] \times E \times A \times \mathcal{S} \times \mathcal{P}(A) \rightarrow \mathbb{R}$ and $g : E \times \mathcal{S} \rightarrow \mathbb{R}$ be respectively the running and terminal cost functions. In the rest of the paper, we make the following assumptions on the regularity of the cost functions.

Assumption 2.7. *There exists a constant $C > 0$ such that for all $(t, i, j) \in [0, T] \times E^2$, $\alpha, \alpha' \in A$, $p, p' \in \mathcal{S}$ and $\nu, \nu' \in \mathcal{P}(A)$, we have:*

$$(10) \quad |f(t, e_i, \alpha, p, \nu) - f(t, e_i, \alpha', p', \nu')| \leq C(\|\alpha - \alpha'\| + \|p - p'\| + \mathcal{W}_1(\nu, \nu')),$$

$$(11) \quad |g(e_i, p) - g(e_i, p')| \leq C\|p - p'\|.$$

When a player chooses a strategy $\alpha \in \mathbb{A}$ and the mean field is $(\mathbf{p}, \nu)$, its cost is:

$$(12) \quad J(\alpha, \mathbf{p}, \nu) := \mathbb{E}^{\mathbb{Q}(\alpha, \mathbf{p}, \nu)} \left[\int_0^T f(t, X_t, \alpha_t, p_t, \nu_t) dt + g(X_T, p_T) \right].$$

Each player aims at minimizing its cost, that is, it solves the optimization problem:

$$(13) \quad V(\mathbf{p}, \nu) := \inf_{\alpha \in \mathbb{A}} \mathbb{E}^{\mathbb{Q}(\alpha, \mathbf{p}, \nu)} \left[\int_0^T f(t, X_t, \alpha_t, p_t, \nu_t) dt + g(X_T, p_T) \right].$$

The key idea of the theory of mean field games lies in the limit scenario of having infinitely many players in the game, where a single player's strategy α does not alter the mean field $(\mathbf{p}, \nu)$. Therefore when each player solves its own optimization problem, it considers $(\mathbf{p}, \nu)$ as given. A Nash equilibrium is then achieved when the law of X_t under the player controlled probability law, along with the distribution of its control under the same probability law, coincide with $(\mathbf{p}, \nu)$. This justifies the following definition of a Nash equilibrium for the weak formulation of finite state mean field games.

Definition 2.8. Let $\mathbf{p}^* : [0, T] \rightarrow \mathcal{S}$, and $\nu^* : [0, T] \rightarrow \mathcal{P}(A)$ be two measurable functions and $\alpha^* \in \mathbb{A}$. We say that the tuple $(\alpha^*, \mathbf{p}^*, \nu^*)$ is a Nash equilibrium for the weak formulation of the mean field game if:

(i) α^* minimizes the cost when the mean field is given by $(\mathbf{p}^*, \nu^*)$:

$$(14) \quad \alpha^* \in \arg \inf_{\alpha \in \mathbb{A}} \mathbb{E}^{\mathbb{Q}(\alpha, \mathbf{p}^*, \nu^*)} \left[\int_0^T f(t, X_t, \alpha_t, p_t^*, \nu_t^*) dt + g(X_T, p_T^*) \right].$$

(ii) $(\alpha^*, \mathbf{p}^*, \nu^*)$ satisfies the consistency conditions whereby for each time $t \in [0, T]$ it holds:

$$(15) \quad p_t^* = \{\mathbb{Q}(\alpha^*, \mathbf{p}^*, \nu^*)[X_t = e_i]\}_{i=1, \dots, m},$$

$$(16) \quad \nu_t^* = \mathbb{Q}_{\# \alpha_t^*}^{\alpha^*, \mathbf{p}^*, \nu^*}.$$

3. INDIVIDUAL PLAYER'S OPTIMIZATION PROBLEM

Before introducing and solving the individual player optimization problem, we provide the necessary background on stochastic equations based on continuous time Markov chains.

3.1. BSDE driven by continuous-time Markov chain. We first recall some of the results on BSDEs driven by continuous-time Markov chains obtained in [6] and [7]. Recall that $\mathcal{M}$ is the $\mathbb{P}$ -martingale in the canonical decomposition of the Markov chain $\mathbf{X}$ in (1). We consider the following BSDE with unknown $(\mathbf{Y}, \mathbf{Z})$, where $\mathbf{Y}$ is an adapted and càdlàg process in $\mathbb{R}$, and $\mathbf{Z}$ is an adapted and left-continuous process in $\mathbb{R}^m$:

$$(17) \quad Y_t = \xi + \int_t^T F(w, s, Y_s, Z_s) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s.$$

Here ξ is a $\mathcal{F}_T$ -measurable $\mathbb{P}$ -square integrable random variable and F is the driver function, assumed to be such that the process $t \rightarrow F(w, t, y, z)$ is predictable for all y, z .

Recalling the definition (4) of the stochastic semi-norm $\|\cdot\|_{X_{t-}}$, we have the following existence and uniqueness result. See Theorem 1.1 in [7].

Lemma 3.1. Assume that there exists $C > 0$ such that $dt \otimes d\mathbb{P}$ -a.s., for all $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^m$ we have:

$$|F(w, t, y, z) - F(w, t, y', z')| \leq C(|y - y'| + \|z - z'\|_{X_{t-}}).$$

Then the BSDE (17) admits a solution $(\mathbf{Y}, \mathbf{Z})$ satisfying

$$\mathbb{E} \left[\int_0^T |Y_t|^2 dt \right] < +\infty, \quad \mathbb{E} \left[\int_0^T \|Z_t\|_{X_{t-}}^2 dt \right] < +\infty.$$

In addition, the solution is unique in the sense that if $(\mathbf{Y}^1, \mathbf{Z}^1)$ and $(\mathbf{Y}^2, \mathbf{Z}^2)$ are two solutions, then $\mathbf{Y}^1$ and $\mathbf{Y}^2$ are indistinguishable and we have $\mathbb{E}[\int_0^T \|Z_t^1 - Z_t^2\|_{X_{t-}}^2 dt] = 0$.

We also have the following stability property, which can be proved by mimicking the argument used in the proof of Theorem 2.1 in [16].

Lemma 3.2. *For $n \geq 0$, let $(\mathbf{Y}^n, \mathbf{Z}^n)$ be the solution to the BSDE (17) with driver F^n and terminal condition ξ^n . Assume that for each n , F^n satisfies the Lipschitz continuity assumption in Lemma 3.1 with the same constant. In addition, assume that the following conditions hold:*

- (i) $\lim_{n \rightarrow \infty} \mathbb{E}[|\xi^n - \xi^0|^2] = 0$.
 - (ii) For each $t \leq T$, $\lim_{n \rightarrow \infty} \mathbb{E}[(\int_t^T |F^n(w, s, Y_s^0, Z_s^0) - F^0(w, s, Y_s^0, Z_s^0)| ds)^2] = 0$.
 - (iii) There exists $C > 0$ such that $\mathbb{E}[(\int_t^T (F^n(w, s, Y_s^0, Z_s^0) - F^0(w, s, Y_s^0, Z_s^0)) ds)^2] \leq C$ for all $t \leq T$ and $n \geq 0$.
- Then we have:

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_t^T \|Z_s^n - Z_s^0\|_{X_{s-}}^2 ds \right] + \mathbb{E}[|Y_t^n - Y_t^0|^2] = 0.$$

Finally we state a crucial comparison result for linear BSDEs. See Theorem 3.16 in [7].

Lemma 3.3. *Let γ be a bounded predictable process in $\mathbb{R}^m$, β a bounded predictable process in $\mathbb{R}$, ϕ a non-negative predictable process in $\mathbb{R}$ such that $\mathbb{E}[\int_0^T \|\phi_t\|^2 dt] < +\infty$, and ξ a non-negative square-integrable $\mathcal{F}_T$ -measurable random variable in $\mathbb{R}$, and let us assume that $(\mathbf{Y}, \mathbf{Z})$ solves the linear BSDE:*

$$(18) \quad Y_t = \xi + \int_t^T (\phi_u + \beta_u Y_u + \gamma_u^* \cdot Z_u) du - \int_t^T Z_u^* \cdot d\mathcal{M}_u.$$

If for all $t \in (0, T]$ and j such that $e_j^* \cdot Q^0 \cdot X_{t-} > 0$, we have $1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) \geq 0$ where ψ_t^+ is the Moore-Penrose inverse of the matrix ψ_t defined in equation (3), then $\mathbf{Y}$ is nonnegative.

Later in the treatment of games with finitely many players, we will need to consider BSDEs driven by multiple independent continuous-time Markov chains. It turns out that all the results above regarding BSDEs driven by one single continuous-time Markov chain can be easily extended to this more general setting. For the sake of completeness, we state and prove these results in the appendix.

3.2. Hamiltonian. We define the Hamiltonian for the optimization problem of the individual player as the function H from $[0, T] \times E \times \mathbb{R}^m \times A \times \mathcal{S} \times \mathcal{P}(A)$ into $\mathbb{R}$ by:

$$(19) \quad H(t, x, z, \alpha, p, \nu) := f(t, x, \alpha, p, \nu) + x^* \cdot (Q(t, \alpha, p, \nu) - Q^0) \cdot z.$$

Since the process X takes value in the set $\{e_1, \dots, e_m\}$, it is more convenient to consider m Hamiltonian functions H_i defined for $i = 1, \dots, m$ by $H_i(t, z, \alpha, p, \nu) := H(t, e_i, z, \alpha, p, \nu)$. Clearly we have:

$$(20) \quad H_i(t, z, \alpha, p, \nu) = f(t, e_i, \alpha, p, \nu) + \sum_{j \neq i} (z_j - z_i)(q(t, i, j, \alpha, p, \nu) - 1).$$

We denote by $\hat{H}_i$ the corresponding minimized Hamiltonian:

$$\hat{H}_i(t, z, p, \nu) := \inf_{\alpha \in A} H_i(t, z, \alpha, p, \nu),$$

and to show the existence of Nash equilibria, we make the following assumption on the minimizer of the Hamiltonian.

Assumption 3.4. (i) For any $t \in [0, T]$, $i \in \{1, \dots, m\}$, $z \in \mathbb{R}^m$, $p \in \mathcal{S}$ and $\nu \in \mathcal{P}(A)$, the mapping $\alpha \rightarrow H_i(t, z, \alpha, p, \nu)$ admits a unique minimizer which does not depend on the mean field of control ν . We denote the minimizer by $\hat{\alpha}_i(t, z, p)$.

(ii) $\hat{\alpha}_i$ is measurable on $[0, T] \times \mathbb{R}^m \times \mathcal{S}$ and there exist constants $C_1 > 0$ and $C_2 \geq 0$ such that for all $i \in \{1, \dots, m\}$, $z, z' \in \mathbb{R}^m$, $p, p' \in \mathcal{S}$:

$$(21) \quad \|\hat{\alpha}_i(t, z, p) - \hat{\alpha}_i(t, z', p')\| \leq C_1 \|z - z'\|_{e_i} + (C_1 + C_2 \|z\|_{e_i}) \|p - p'\|.$$

Remark 3.5. For the sake of convenience, we choose to make the assumption directly on the uniqueness and the regularity of the minimizer of the Hamiltonian. One possible way to make sure Assumption 3.4 holds is to impose linearity on the transition rate function q , and strong convexity of the running cost function f . For example, the following set of conditions will guarantee that Assumption 3.4 holds:

Assumption 3.6. (i) A is a convex and compact subset of $\mathbb{R}^l$.

(ii) The transition rate function q takes the form $q(t, i, j, \alpha, p, \nu) = q_0(t, i, j, p, \nu) + q_1(t, i, j, p) \cdot \alpha$, where $q_0 : [0, T] \times E^2 \times \mathcal{S} \times \mathcal{P}(A) \rightarrow \mathbb{R}$ and $q_1 : [0, T] \times E^2 \times \mathcal{S} \rightarrow \mathbb{R}^l$ are two continuous mappings.

(iii) The running cost function f is of the form $f(t, x, \alpha, p, \nu) = f_0(t, x, \alpha, p) + f_1(t, x, p, \nu)$, where for each $i \in \{1, \dots, m\}$, the mapping $f_0(\cdot, e_i, \cdot, \cdot)$ (resp. $f_1(\cdot, e_i, \cdot, \cdot)$) is continuous on $[0, T] \times A \times \mathcal{P}$ (resp. $[0, T] \times \mathcal{S} \times \mathcal{P}(A)$).

(iv) For all $(t, e_i, p) \in [0, T] \times E \times \mathcal{S}$, the mapping $\alpha \rightarrow f_0(t, e_i, \alpha, p)$ is once continuously differentiable and there exists a constant $C > 0$ such that:

$$(22) \quad \|\nabla_\alpha f_0(t, e_i, \alpha, p) - \nabla_\alpha f_0(t, e_i, \alpha, p')\| \leq C\|p - p'\|.$$

(v) f_0 is γ -uniformly convex in α , i.e., for all $(t, e_i, p) \in [0, T] \times E \times \mathcal{S}$ and $\alpha, \alpha' \in A$, we have:

$$(23) \quad f_0(t, e_i, \alpha, p) - f_0(t, e_i, \alpha', p) - (\alpha - \alpha') \cdot \nabla_\alpha f_0(t, e_i, \alpha, p) \geq \gamma\|\alpha' - \alpha\|^2$$

We define the functions $\hat{H}$ and $\hat{a}$ by:

$$(24) \quad \hat{H}(t, x, z, p, \nu) := \sum_{i=1}^m \mathbb{1}(x = e_i) \hat{H}_i(t, z, p, \nu),$$

$$(25) \quad \hat{a}(t, x, z, p) := \sum_{i=1}^m \mathbb{1}(x = e_i) \hat{a}_i(t, z, p).$$

From item (i) of Assumption 3.4 and the definition of the reduced Hamiltonian H_i , it is clear that $\hat{a}(t, x, z, p)$ is the unique minimizer of the mapping $\alpha \rightarrow H(t, x, z, \alpha, p, \nu)$, and the minimum equals $\hat{H}(t, x, z, p, \nu)$. In addition, from Assumptions 2.2, 2.7, 3.4, and the definition of the stochastic semi-norm $\|\cdot\|_{X_{t-}}$, it is easy to deduce the regularity of the mappings $\hat{H}$ and $\hat{a}$.

Lemma 3.7. *There exists a constant $C > 0$ such that for all $(\omega, t) \in \Omega \times (0, T]$, $p, p' \in \mathcal{S}$, $\nu, \nu' \in \mathcal{P}(A)$ and $z, z' \in \mathbb{R}^m$, we have:*

$$(26) \quad |\hat{H}(t, X_{t-}, z, p, \nu) - \hat{H}(t, X_{t-}, z', p', \nu')| \leq C\|z - z'\|_{X_{t-}} + C(1 + \|z\|_{X_{t-}})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')),$$

$$(27) \quad |\hat{a}(t, X_{t-}, z, p) - \hat{a}(t, X_{t-}, z', p')| \leq C\|z - z'\|_{X_{t-}} + C(1 + \|z\|_{X_{t-}})\|p - p'\|.$$

Proof. Inequality (27) is an easy consequence of Assumption 3.4 and the definition of the stochastic seminorm $\|\cdot\|_{X_{t-}}$. We now deal with the regularity of $\hat{H}$. By Berge's maximum theorem, the continuity of H_i and the compactness of A imply the continuity of $\hat{H}_i$. Let $z, z' \in \mathbb{R}^m$, $p, p' \in \mathcal{S}$ and $\nu, \nu' \in \mathcal{P}(A)$. For any $\alpha \in A$, we have:

$$\begin{aligned} \hat{H}_i(t, z, p, \nu) - H_i(t, z', \alpha, p', \nu') &\leq H_i(t, z, \alpha, p, \nu) - H_i(t, z', \alpha, p', \nu') \\ &= f(t, e_i, \alpha, p, \nu) - f(t, e_i, \alpha, p', \nu') + \sum_{j \neq i} [(z_j - z_i) - (z'_j - z'_i)] q(t, i, j, \alpha, p', \nu') \\ &\quad + \sum_{j \neq i} (z_j - z_i) [q_0(t, i, j, p, \nu) - q_0(t, i, j, p', \nu')] + (z_j - z_i) [q_1(t, i, j, p) - q_1(t, i, j, p')] \cdot \alpha \\ &\leq C\|z - z'\|_{e_i} + C(1 + \|z\|_{e_i})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')), \end{aligned}$$

where we used the Lipschitz property of f and q , and the boundedness of A and q . Since the above is true for all $\alpha \in A$, taking supremum of the left-hand side, we obtain:

$$\hat{H}_i(t, z, p, \nu) - \hat{H}_i(t, z', p', \nu') \leq C\|z - z'\|_{e_i} + C(1 + \|z\|_{e_i})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')).$$

Exchanging the roles of z and z' , we obtain:

$$|\hat{H}_i(t, z, p, \nu) - \hat{H}_i(t, z', p', \nu')| \leq C\|z - z'\|_{e_i} + C(1 + \|z\|_{e_i})(\|p - p'\| + \mathcal{W}_1(\nu, \nu')),$$

and (26) follows immediately from the definition of the seminorm $\|\cdot\|_{X_{t-}}$. $\square$

3.3. Player's optimization problem. In this subsection, we show that the optimization problem of the player facing a given mean field of state and control can be characterized by a BSDE driven by the continuous-time Markov chain $\mathbf{X}$. Let us fix measurable flows $\mathbf{p} : [0, T] \rightarrow \mathcal{S}$ and $\nu : [0, T] \rightarrow \mathcal{P}(A)$, an admissible strategy $\alpha \in \mathbb{A}$, and let us consider the BSDE:

$$(28) \quad Y_t = g(X_T, p_T) + \int_t^T H(s, X_{s-}, Z_s, \alpha_s, p_s, \nu_s) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s.$$

Lemma 3.8. *The BSDE (28) admits a unique solution (Y, Z) and $J(\alpha, p, \nu) = \mathbb{E}^\mathbb{P}[Y_0]$.*

Proof. From the boundedness of the transition rate function q guaranteed by Assumption 2.2, it is easy to check that the driver function H of the BSDE (28) is Lipschitz in z with respect to the semi-norm $\|\cdot\|_{X_{t-}}$. Therefore by Lemma 3.1, it admits a unique solution (Y, Z) . Moreover, we have:

$$\begin{aligned} Y_0 &= g(X_T, p_T) + \int_0^T H(t, X_{t-}, Z_t, \alpha_t, p_t, \nu_t) dt - \int_0^T Z_t^* \cdot d\mathcal{M}_t \\ &= g(X_T, p_T) + \int_0^T f(t, X_{t-}, \alpha_t, p_t, \nu_t) dt - \int_0^T Z_t^* \cdot (d\mathcal{M}_t - (Q^*(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot X_{t-} dt) \\ &= g(X_T, p_T) + \int_0^T f(t, X_{t-}, \alpha_t, p_t, \nu_t) dt - \int_0^T Z_t^* \cdot d\mathcal{M}_t^{(\alpha, \mathbf{p}, \nu)}. \end{aligned}$$

Since $\mathcal{M}^{(\alpha, \mathbf{p}, \nu)}$ is a martingale under the measure $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$, we take expectation under $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ and obtain $J(\alpha, \mathbf{p}, \nu) = \mathbb{E}^{\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}}[Y_0]$. Now since Y_0 is $\mathcal{F}_0$ -measurable, and $\mathbb{Q}^{(\alpha, \mathbf{p}, \nu)}$ coincides with $\mathbb{P}$ on $\mathcal{F}_0$, we obtain $J(\alpha, p, \nu) = \mathbb{E}^\mathbb{P}[Y_0]$. $\square$

Now we consider the following BSDE:

$$(29) \quad Y_t = g(X_T, p_T) + \int_t^T \hat{H}(s, X_{s-}, Z_s, p_s, \nu_s) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s,$$

and we show that it characterizes the optimality of the control problem (13).

Proposition 3.9. *For any measurable function $\mathbf{p}$ from $[0, T]$ to $\mathcal{S}$ and any measurable function ν from $[0, T]$ to $\mathcal{P}(A)$, the BSDE (29) admits a unique solution (Y, Z) . The value function of the optimal control problem (13) is given by $V(\mathbf{p}, \nu) = \mathbb{E}^\mathbb{P}[Y_0]$ and the process $\hat{\alpha}^{(\mathbf{p}, \nu)}$ defined by:*

$$(30) \quad \hat{\alpha}_t^{(\mathbf{p}, \nu)} := \hat{a}(t, X_{t-}, Z_t, p_t)$$

is an optimal control. In addition, if $\alpha' \in \mathbb{A}$ is an optimal control, we have $\alpha'_t = \hat{\alpha}_t^{(\mathbf{p}, \nu)}$, $dt \otimes d\mathbb{P}$ -a.e.

Proof. The existence and uniqueness of the solution to (29) is easily verified by using the Lipschitz property of $\hat{H}$ provided by Lemma 3.7. Let (Y, Z) be this unique solution and define the process $\hat{\alpha}$ by $\hat{\alpha}_t := \hat{a}(t, X_{t-}, Z_t, p_t)$. Recall the definition of $\hat{a}$ in equation (25). We have:

$$\hat{a}(t, X_{t-}, Z_t, p_t) = \sum_{i=1}^m \mathbb{1}(X_{t-} = e_i) \hat{a}_i(t, Z_t, p_t) = X_{t-}^* \cdot \left(\sum_{i=1}^m \hat{a}_i(t, Z_t, p_t) e_i \right).$$

Since $\hat{a}_i$ is measurable for each $i \in E$, we see that $\hat{a}$ is a measurable mapping from $[0, T] \times \mathbb{R}^m \times \mathbb{R}^m \times \mathcal{S}$ to A . Since both the processes $t \rightarrow X_{t-}$ and Z are predictable, we conclude that $\hat{\alpha}$ is a predictable process and therefore an admissible control.

Now let us fix an arbitrary admissible control $\alpha \in \mathbb{A}$, and denote by $(\mathbf{Y}^\alpha, \mathbf{Z}^\alpha)$ the solution of the corresponding BSDE (28), and by $(\mathbf{Y}, \mathbf{Z})$ the unique solution of:

$$(31) \quad Y_t = \int_t^T H(s, X_{s-}, Z_s, \hat{\alpha}_s, p_s, \nu_s) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s.$$

Setting $\Delta \mathbf{Y} := \mathbf{Y}^\alpha - \mathbf{Y}$ and $\Delta \mathbf{Z} := \mathbf{Z}^\alpha - \mathbf{Z}$ and computing the difference of the two BSDEs, we notice that $\Delta \mathbf{Y}$ and $\Delta \mathbf{Z}$ solve the following BSDE:

$$\Delta Y_t = \int_t^T [H(s, X_{s-}, Z_s^\alpha, \alpha_s, p_s, \nu_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s, p_s, \nu_s)] ds - \int_t^T \Delta Z_s^* \cdot d\mathcal{M}_s.$$

We can further decompose the driver of the above BSDE as:

$$\begin{aligned} & H(s, X_{s-}, Z_s^\alpha, \alpha_s, p_s, \nu_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s, p_s, \nu_s) \\ &= H(s, X_{s-}, Z_s^\alpha, \alpha_s, p_s, \nu_s) - H(s, X_{s-}, Z_s, \alpha_s, p_s, \nu_s) + H(s, X_{s-}, Z_s, \alpha_s, p_s, \nu_s) \\ & \quad - H(s, X_{s-}, Z_s, \hat{\alpha}_s, p_s, \nu_s) \\ &= [H(s, X_{s-}, Z_s, \alpha_s, p_s, \nu_s) - H(s, X_{s-}, Z_s, \hat{\alpha}_s, p_s, \nu_s)] + X_{s-}^* \cdot (Q(s, \alpha_s, p_s, \nu_s) - Q^0) \cdot \Delta Z. \end{aligned}$$

Define the processes ψ and γ by $\psi_t := H(t, X_{t-}, Z_t, \alpha_t, p_t, \nu_t) - H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t)$ and $\gamma_t := (Q^*(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot X_{t-}$. Therefore $(\Delta \mathbf{Y}, \Delta \mathbf{Z})$ appears as the solution to a linear BSDE of the form (18) with ψ and γ defined previously and $\beta = 0$. Clearly ψ and γ are both predictable. Since $\hat{\alpha}_t$ minimizes the Hamiltonian, ψ is nonnegative. The boundedness of γ follows from the boundedness of the transition rate function q . It remains to check that $1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) \geq 0$.

When $X_{t-} = e_j$, the above inequality holds clearly. So we assume that $X_{t-} = e_i \neq e_j$. We have $\psi_t^+ \cdot (e_j - X_{t-}) = \frac{m-1}{m} e_j - \sum_{i \neq j} \frac{1}{m} e_i$. Therefore when $X_{t-} = e_i \neq e_j$, we have:

$$\begin{aligned} \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) &= X_{t-}^* \cdot (Q(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot \psi_t^+ \cdot (e_j - X_{t-}) \\ &= e_i^* \cdot (Q(t, \alpha_t, p_t, \nu_t) - Q^0) \cdot \left(\frac{m-1}{m} e_j - \sum_{k \neq j} \frac{1}{m} e_k \right) \\ &= \frac{m-1}{m} (q(t, i, j, \alpha_t, p_t, \nu_t) - q_{i,j}^0) - \frac{1}{m} \sum_{k \neq j} (q(t, i, k, \alpha_t, p_t, \nu_t) - q_{i,k}^0) \\ &= q(t, i, j, \alpha_t, p_t, \nu_t) - q_{i,j}^0, \end{aligned}$$

where the last equality is due to the fact that $\sum_k (q(t, i, k, \alpha_t, p_t, \nu_t) - q_{i,k}^0) = 0$. Therefore we have:

$$1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_j - X_{t-}) = 1 + q(t, i, j, \alpha_t, p_t, \nu_t) - q_{i,j}^0 = q(t, i, j, \alpha_t, p_t, \nu_t) \geq 0.$$

By Lemma 3.3, we conclude that $\Delta \mathbf{Y}$ is nonnegative and in particular $Y_0^\alpha \geq Y_0$. Since α is an arbitrary admissible control, in light of Lemma 3.8, this means that $\mathbb{E}^\mathbb{P}[Y_0] \leq \inf_{\alpha \in \mathbb{A}} J(\alpha, \mathbf{p}, \nu) = V(\mathbf{p}, \nu)$. Finally, we notice that Y_0 is the expected total cost when the control is $\hat{\alpha}$. We conclude that $\hat{\alpha}$ is an optimal control and $\mathbb{E}^\mathbb{P}[Y_0] = V(\mathbf{p}, \nu)$.

Now we show that $\hat{\alpha}$ is the unique optimal control. Let α' be another optimal control. We consider the solution $(\mathbf{Y}', \mathbf{Z}')$ to the following BSDE:

$$(32) \quad Y'_t = \int_t^T H(s, X_{s-}, Z'_s, \alpha'_s, p_s, \nu_s) ds - \int_t^T (Z'_s)^* \cdot d\mathcal{M}_s.$$

Since α' is optimal, we have $\mathbb{E}^\mathbb{P}[Y'_0] = J(\alpha', \mathbf{p}, \nu) = V(\mathbf{p}, \nu) = \mathbb{E}^\mathbb{P}[Y_0]$. Now taking the difference of the BSDE (31) and (32), we obtain:

$$Y_0 - Y'_0 = \int_0^T [H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t) - H(t, X_{t-}, Z'_t, \alpha'_t, p_t, \nu_t)] dt - \int_0^T (Z_t - Z'_t)^* \cdot d\mathcal{M}_t$$

$$\begin{aligned}
&= \int_0^T [X_{t-}^* \cdot (Q(t, \hat{\alpha}_t, p_t, \nu_t) - Q^0) \cdot Z_t - X_{t-}^* \cdot (Q(t, \alpha'_t, p_t, \nu_t) - Q^0) \cdot Z'_t] dt \\
&\quad + \int_0^T [f(t, X_{t-}, \hat{\alpha}_t, p_t, \nu_t) - f(t, X_{t-}, \alpha'_t, p_t, \nu_t)] dt - \int_0^T (Z_t - Z'_t)^* \cdot d\mathcal{M}_t \\
&= \int_0^T [f(t, X_{t-}, \hat{\alpha}_t, p_t, \nu_t) - f(t, X_{t-}, \alpha'_t, p_t, \nu_t) + X_{t-}^* \cdot (Q(t, \hat{\alpha}_t, p_t, \nu_t) - Q(t, \alpha'_t, p_t, \nu_t)) \cdot Z_t] dt \\
&\quad - \int_0^T (Z_t - Z'_t)^* \cdot [d\mathcal{M}_t - (Q^*(t, \alpha'_t, p_t, \nu_t) - Q^0) \cdot X_{t-} dt] \\
&= \int_0^T [H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t) - H(t, X_{t-}, Z_t, \alpha'_t, p_t, \nu_t)] dt - \int_0^T (Z_t - Z'_t)^* \cdot d\mathcal{M}_t^{(\alpha', \mathbf{p}, \nu)}.
\end{aligned}$$

Taking $\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}$ -expectations and using the fact that $\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}$ coincides with $\mathbb{P}$ in $\mathcal{F}_0$, we get:

$$\begin{aligned}
0 &= \mathbb{E}^\mathbb{P}[Y_0 - Y'_0] = \mathbb{E}^{\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}}[Y_0 - Y'_0] \\
&= \mathbb{E}^{\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}} \left[\int_0^T [H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t) - H(t, X_{t-}, Z_t, \alpha'_t, p_t, \nu_t)] dt \right] \leq 0,
\end{aligned}$$

where the last inequality is due to the fact that $\hat{\alpha}_t$ minimizes the Hamiltonian. In fact, we have $\hat{\alpha}_t = \alpha'_t$, $dt \otimes d\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}$ -a.e. If we assume otherwise, the last inequality would be strict, since the minimizer of the Hamiltonian is unique by Assumption 3.4. Since $\mathbb{P}$ is equivalent to $\mathbb{Q}^{(\alpha', \mathbf{p}, \nu)}$, we have $\hat{\alpha}_t = \alpha'_t$, $dt \otimes d\mathbb{P}$ -a.e. $\square$

4. EXISTENCE OF NASH EQUILIBRIA

We state the main result of this section:

Theorem 4.1. *Under Assumptions 2.2, 2.7 and 3.4, there exists a Nash equilibrium $(\alpha^*, \mathbf{p}^*, \nu^*)$ for the weak formulation of the finite state mean field game in the sense of Definition 2.8.*

The rest of this section is devoted to the proof of Theorem 4.1. As in the case of diffusion-based mean field games, we shall rely on a fixed point argument to show existence of Nash equilibria. We start from a measurable function $\mathbf{p} : [0, T] \rightarrow \mathcal{S}$ and a measurable function $\nu : [0, T] \rightarrow \mathcal{P}(A)$ where we recall that $\mathcal{S}$ is the m -dimensional simplex which we identify with the space of probability measures on E , while $\mathcal{P}(A)$ is the space of probability measures on A . We then solve the BSDE (29), and obtain the solution $(\mathbf{Y}^{(\mathbf{p}, \nu)}, \mathbf{Z}^{(\mathbf{p}, \nu)})$ as well as the optimal control $\hat{\alpha}^{(\mathbf{p}, \nu)}$ given by (30). Finally, we compute the probability measure $\hat{\mathbb{Q}}^{(\mathbf{p}, \nu)} := \mathbb{Q}^{(\hat{\alpha}^{(\mathbf{p}, \nu)}, \mathbf{p}, \nu)}$ as defined in (7), and consider the push-forward measures of $\hat{\mathbb{Q}}^{(\mathbf{p}, \nu)}$ by $(X_t, \hat{\alpha}_t^{(\mathbf{p}, \nu)})$. Clearly, we identified a Nash equilibrium if we find a fixed point for the mapping $(\mathbf{p}, \nu) \rightarrow \hat{\mathbb{Q}}_{\#(X_t, \hat{\alpha}_t^{(\mathbf{p}, \nu)})}^{(\mathbf{p}, \nu)}$.

In practice however, the implementation of the fixed-point argument mentioned above is prone to several difficulties. The foremost challenge lies in the lack of results allowing us to identify compact subsets of the spaces of measurable functions from $[0, T]$ to $\mathcal{S}$ or $\mathcal{P}(A)$. This makes it difficult to apply Schauder's theorem or similar versions of fixed point theorems. For this reason, we shall resort to different descriptions of the mean field for the state and the control. For the mean field of the state, since we have assumed from the very beginning that $\mathbf{X}$ is a càdlàg process, we will directly deal with its probability law on the space D of all càdlàg functions from $[0, T]$ to $E = \{e_1, \dots, e_m\}$ endowed with the Skorokhod topology. The space of probability measures on D and its topological properties have been studied thoroughly (see [18] for a detailed account), and a simple criterion for compactness is available.

Unfortunately, resolving the corresponding issue for the control is more involved. Here, we adopt the technique based on the *stable topology* used in [3]. Indeed, a measurable mapping from $[0, T]$ to $\mathcal{P}(A)$ can be viewed as a

random variable defined on the space $([0, T], \mathcal{B}([0, T]), \mathcal{L})$ taking values in $\mathcal{P}(A)$. Here, $\mathcal{B}([0, T])$ is the Borel σ -field of $[0, T]$, $\mathcal{L}$ is the uniform probability measure on $[0, T]$ and $\mathcal{P}(A)$ is endowed with the Wasserstein-1 distance. To obtain compactness, the idea is to use randomization. We consider the space of probability measures on $[0, T] \times \mathcal{P}(A)$, denoted by $\mathcal{P}([0, T] \times \mathcal{P}(A))$. Then for each measurable mapping ν from $[0, T]$ to $\mathcal{P}(A)$, we consider the measure η on $[0, T] \times \mathcal{P}(A)$ given by $\eta(dt, dm) := \mathcal{L}(dt) \times \delta_{\nu_t}(dm)$ where δ is the Dirac measure. We may endow the space $\mathcal{P}([0, T] \times \mathcal{P}(A))$ with the so-called *stable topology* introduced in [17], for which convenient results on compactness are readily available.

In the following, we detail the steps that lead to the existence of Nash equilibria. We start by specifying the topology we use for the space of mean fields on the state as well as the control. We then properly define the mapping compatible with the definition of Nash equilibrium, we show its continuity, and construct a stable compact. Once these ingredients are in place, we apply Schauder's fixed point theorem to conclude.

4.1. Topology for the space of mean fields. We first consider the mean field for the state by endowing the state space $E := \{e_1, \dots, e_m\}$ with the discrete metric $d_E(x, y) := \mathbb{1}(x \neq y)$. Then it is well known that (E, d_E) is a Polish space. Then, the Skorokhod space:

$$(33) \quad D := \{x : [0, T] \rightarrow E, x \text{ is càdlàg and left continuous on } T\}$$

is endowed with the J1 metric:

$$(34) \quad d_D(x, y) := \inf_{\lambda \in \Lambda} \max\{\sup_{t \leq T} |\lambda(t) - t|, \sup_{t \leq T} |y(\lambda(t)) - x(t)|\}$$

where Λ is the set of all strictly increasing, continuous bijections from $[0, T]$ to itself. It can be proved that d_D is a metric on D and the metric space (D, d_D) is a Polish space. Let us denote by $\mathcal{P}$ the collection of probability measures on (D, d_D) endowed with the weak topology. Recall that the reference measure $\mathbb{P}$ is an element of $\mathcal{P}$. Let $\mathcal{P}_0$ be the subset of $\mathcal{P}$ defined by:

$$(35) \quad \mathcal{P}_0 := \{\mathbb{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}} = L, \text{ with } \mathbb{E}^{\mathbb{P}}[L^2] \leq C_0\}.$$

where C_0 is a constant which we will specify later (see the proof of Proposition 4.10). We have the following result:

Proposition 4.2. *$\mathcal{P}_0$ is convex and relatively compact in $\mathcal{P}$.*

Proof. The convexity of $\mathcal{P}_0$ is trivial. Let us show that $\mathcal{P}_0$ is relatively compact. We proceed in three steps.

Step 1. For $K \in \mathbb{N}$ and $\delta > 0$, we define $D_{\delta, K}$ as the collection of paths in D which meet the following criteria: (a) the path has no more than K discontinuities, (b) the first jump time, if any, happens on or after δ , (c) the last jump happens on or before $T - \delta$, and (d) the amounts of time between jumps are greater or equal than δ . We now show that $D_{\delta, K}$ is compact in D . Since D is Polish space it is enough to show the sequential compactness. Let us fix a sequence x_n in $D_{\delta, K}$. For each x_n , we use the following notation: k_n is the number of its jumps, $\delta \leq t_n^1 < t_n^2 < \dots < t_n^{k_n} \leq T - \delta$ are the times of its jumps. $\Delta t_n^1 := t_n^1$ and $\Delta t_n^i := t_n^i - t_n^{i-1}$ for $i = 2, \dots, k_n$ are the time elapsed between consecutive jumps and $x_n^0, x_n^1, \dots, x_n^{k_n}$ are the value taken by x_n in each interval defined by the jumps. Then we can represent x_n using the vector y_n of dimension $2(K + 1)$:

$$y_n = [k_n, \Delta t_n^1, \Delta t_n^2, \dots, \Delta t_n^{k_n}, 0, \dots, 0, x_n^0, x_n^1, \dots, x_n^{k_n}, 0, \dots, 0].$$

In the above representation, the first coordinate of y_n is the number of jumps. Coordinate 2 to $K + 1$ are the times elapsed between jumps defined above, and if there are fewer than K jumps, we complete the vector by 0. Coordinates $K + 2$ to $2(K + 1)$ are the values taken by the path x and completed with 0. Clearly there is a bijection from x_n to y_n by this representation. By the definition of the set $D_{\delta, K}$, we have $\Delta t_n^i \in [\delta, T]$ for $i \leq k_n$ and $\sum_{i=1}^{k_n} \Delta t_n^i \leq T - \delta$, whereas the rest of the coordinates of y_n belongs to a finite set. This implies that y_n lives in a compact and therefore we can extract a converging subsequence which we still denote by y_n . Again, since k_n and the last $K + 1$ components can only take finitely many values by their definition, therefore there exists N_0

such that for $n \geq N_0$, we have $k_n = k$ and $x_n^i = x^i$ for all $i \leq k$. In addition we have Δt_n^i converges to Δt^i for all $i \leq k$, where $\Delta t^i \geq \delta$ for all $i \leq k$ and $\sum_{i=1}^k \Delta t^i \leq T - \epsilon$. We consider the path represented by the vector y :

$$y = [k, \Delta t^1, \Delta t^2, \dots, \Delta t^k, 0, \dots, 0, x^0, x^1, \dots, x^k, 0, \dots, 0].$$

Clearly x belongs to the set $D_{\delta, K}$ and it is straightforward to verify that x_n converge to x in J1 metric, where x_n is the path represented by the vector y_n . This implies that $D_{\delta, K}$ is compact.

Step 2. Now we show that for any $\epsilon > 0$, there exists $\delta > 0$ and $K \in \mathbb{N}$ such that $\mathbb{P}(D_{\delta, K}) \geq 1 - \epsilon$. Recall that $\mathbb{P}$ is the reference measure and under $\mathbb{P}$ the canonical process $\mathbf{X}$ is a continuous-time Markov chain with transition rate matrix Q^0 . Therefore the time of first jump, as well as the time between consecutive jumps thereafter, which we denote by $\Delta t_1, \Delta t_2, \dots$ are i.i.d. exponential random variables of parameter $(m-1)$ under the measure $\mathbb{P}$. We have:

$$\mathbb{P}(D_{\delta, K}) = \mathbb{P}[\Delta t_1 > T] + \sum_{k=1}^K \mathbb{P} \left[\{\Delta t_1 \geq \delta\} \cap \dots \cap \{\Delta t_k \geq \delta\} \cap \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^k \Delta t_i \leq T - \delta \right\} \right].$$

For each $k = 1, \dots, K$, we have:

$$\begin{aligned} & \mathbb{P} \left[\{\Delta t_1 \geq \delta\} \cap \dots \cap \{\Delta t_k \geq \delta\} \cap \left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^k \Delta t_i \leq T - \delta \right\} \right] \\ & \geq \mathbb{P}[\{\Delta t_1 \geq \delta\} \cap \dots \cap \{\Delta t_k \geq \delta\}] + \mathbb{P} \left[\left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^k \Delta t_i \leq T - \delta \right\} \right] - 1 \\ & = (\mathbb{P}[\Delta t_1 \geq \delta])^k + \mathbb{P} \left[\left\{ \sum_{i=1}^{k+1} \Delta t_i > T \right\} \cap \left\{ \sum_{i=1}^k \Delta t_i \leq T - \delta \right\} \right] - 1 \\ & = (\exp(-k(m-1)\delta) - 1) + \exp(-(m-1)T) \frac{(m-1)^k (T-\delta)^k}{k!}. \end{aligned}$$

It follows that:

$$\begin{aligned} \mathbb{P}(D_{\delta, K}) & \geq \sum_{k=1}^K (\exp(-k(m-1)\delta) - 1) + \exp(-(m-1)T) \sum_{k=0}^K \frac{(m-1)^k (T-\delta)^k}{k!} \\ & \geq \sum_{k=1}^K (\exp(-k(m-1)\delta) - 1) + \exp(-(m-1)T) \sum_{k=0}^K \frac{(m-1)^k T^k}{k!} - (1 - \exp(-(m-1)\delta)). \end{aligned}$$

We can first pick K greater enough such that $(\exp(-(m-1)T) \sum_{k=0}^K (m-1)^k T^k / k!)$ is greater than $(1 - \epsilon/2)$ and then pick δ small enough to make the rest of the terms greater than $-\epsilon/2$, which eventually makes $\mathbb{P}(D_{\delta, K})$ greater than $(1 - \epsilon)$.

Step 3. Finally we show that $\mathcal{P}_0$ is tight. For any $\epsilon > 0$, by Step 2, we can pick $\delta > 0$ and $K \in \mathbb{N}$ such that $\mathbb{P}(D \setminus D_{\delta, K}) \leq (\epsilon/C_0)^2$. For all $\mathbb{Q} \in \mathcal{P}_0$, we have $d\mathbb{Q}/d\mathbb{P} = L$ and $(\mathbb{E}^\mathbb{P}[L^2])^{1/2} \leq C_0$ and by Cauchy-Schwartz inequality we obtain:

$$\mathbb{Q}(D \setminus D_{\delta, K}) = \mathbb{E}^\mathbb{P}[L \cdot 1_{x \in D \setminus D_{\delta, K}}] \leq (\mathbb{E}^\mathbb{P}[L^2])^{1/2} \mathbb{P}(D \setminus D_{\delta, K})^{1/2} \leq \epsilon.$$

This implies the tightness of $\mathcal{P}_0$. Finally by Prokhorov's Theorem we conclude that $\mathcal{P}_0$ is relatively compact. $\square$

We now need to link the convergence of measures on path space to the convergence in $\mathcal{S}$, i.e. measures on state space. We define the function π by:

$$\pi : [0, T] \times \mathcal{P}(\mathbb{E}) \ni (t, \mu) \rightarrow [\mu_{\#X_t}(\{e_1\}), \mu_{\#X_t}(\{e_2\}), \dots, \mu_{\#X_t}(\{e_m\})] \in \mathcal{S}$$

and prove the following result:

Lemma 4.3. *If $\mu^n \mapsto \mu$ in $\mathcal{P}$, there exists a subset $\mathcal{D}(\mu)$ of $[0, T]$ at most countable such that for all $t \notin \mathcal{D}(\mu)$:*

$$(36) \quad \lim_{n \rightarrow +\infty} \pi(t, \mu^n) = \pi(t, \mu).$$

Proof. Define $\mathcal{D}(\mu) := \{0 \leq t \leq T; \mu(X_t - X_{t-} \neq 0) > 0\}$. By Lemma 3.12 in [18], the set $\mathcal{D}(\mu)$ is at most countable. In addition, we have $T \notin \mathcal{D}(\mu)$ since all the paths in D is left-continuous on T . In light of Proposition 3.14 in [18], we have $\mu_{\#X_t}^n$ converges to $\mu_{\#X_t}$ weakly for all $t \notin \mathcal{D}(\mu)$. To conclude, we use the fact that $\mu_{\#X_t}^n$ for all $t \in [0, T]$ and n are counting measure on the discrete set E . $\square$

We now turn to the mean field of control. Let $(\mathcal{P}(A), \mathcal{W}_1)$ be the space of probability measures on the compact set $A \subset \mathbb{R}^l$ endowed with the weak topology and metricized by the Wasserstein-1 distance. $(\mathcal{P}(A), \mathcal{W}_1)$ is a Polish space. Since A is compact, it is easy to show that $\mathcal{P}(A)$ is tight and therefore by Prokhorov's theorem $(\mathcal{P}(A), \mathcal{W}_1)$ is in fact compact. We endow $\mathcal{P}(A)$ with its Borel σ -algebra denoted by $\mathcal{B}(\mathcal{P}(A))$. We endow $[0, T]$ with its Borel σ -algebra $\mathcal{B}([0, T])$ and the (normalized) Lebesgue measure $\mathcal{L}(dt) := \frac{1}{T}dt$. Finally, we construct the product space $[0, T] \times \mathcal{P}(A)$ endowed with the σ -algebra $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathcal{P}(A))$. The space of probability measures on $[0, T] \times \mathcal{P}(A)$ can be viewed as a randomized version of the space of mean field of control. We introduce the stable topology on this space:

Definition 4.4. Let us denote by $\mathcal{R}$ the space of probability measures on $([0, T] \times \mathcal{P}(A), \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathcal{P}(A)))$. We call the stable topology of $\mathcal{R}$ the coarsest topology such that the mappings $\eta \rightarrow \int g(t, m)\eta(dt, dm)$ are continuous for all bounded and measurable mappings g defined on $[0, T] \times \mathcal{P}(A)$ such that $m \rightarrow g(t, m)$ is continuous for each fixed $t \in [0, T]$.

We collect a few useful results on the space $\mathcal{R}$ endowed with the stable topology.

Proposition 4.5. *The topology space $\mathcal{R}$ is compact, metrizable, and Polish.*

Proof. Notice that both $[0, T]$ and $\mathcal{P}(A)$ are Polish for their respective topologies. This implies that the σ -algebra $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathcal{P}(A))$ is separable. It follows from Proposition 2.10 in [17] that $\mathcal{R}$ is metrizable.

We now show that $\mathcal{R}$ is compact. Notice that for an element η in $\mathcal{R}$, its first marginal is a probability measure on $[0, T]$ and its second marginal is a probability measure on $\mathcal{P}(A)$. It is trivial to see that both the spaces of probability measures on $[0, T]$ and on $\mathcal{P}(A)$ are tight and therefore relatively compact by Prokhorov's theorem. We then apply Theorem 2.8 in [17] and obtain the compactness of $\mathcal{R}$.

Having showed that $\mathcal{R}$ is compact and metrizable, we see that $\mathcal{R}$ is separable. Compactness also leads to completeness. Therefore $\mathcal{R}$ is Polish space. Finally, we notice that $\mathcal{R}$ is also sequential compact since $\mathcal{R}$ is metrizable. $\square$

The following result provides a more convenient way to characterize the convergence in the stable topology.

Lemma 4.6. *Denote by $\mathcal{H}$ the collection of mappings f of the form $f(t, \nu) = 1_B(t) \cdot g(\nu)$ where B is a Borel subset of $[0, T]$ and $g : \mathcal{P}(A) \rightarrow \mathbb{R}$ is a bounded Lipschitz function (with respect to the Wasserstein-1 distance on $\mathcal{P}(A)$). Then the stable topology introduced in Definition 4.4 is the coarsest topology which makes the mappings $\eta \rightarrow \int_{[0, T] \times \mathcal{P}(A)} f(t, \nu)\eta(dt, d\nu)$ continuous for all $f \in \mathcal{H}$.*

Proof. Let $\mathcal{H}_0$ be the collection of mappings f of the form $f(t, \nu) = 1_B(t) \cdot g(\nu)$ where B is a Borel subset of $[0, T]$ and $g : \mathcal{P}(A) \rightarrow \mathbb{R}$ is a bounded and uniformly continuous function. Then clearly we have $\mathcal{H} \subset \mathcal{H}_0$. By Proposition 2.4 in [17], the stable topology is the coarsest topology under which the mappings $\eta \rightarrow \int_{[0, T] \times \mathcal{P}(A)} f(t, \nu)\eta(dt, d\nu)$ are continuous for all $f \in \mathcal{H}_0$. Therefore, we only need to show that if η^n is a sequence of elements in $\mathcal{R}$ such that $\int f(t, \nu)\eta^n(dt, d\nu) \rightarrow \int f(t, \nu)\eta^0(dt, d\nu)$ for all $f \in \mathcal{H}$, then we have $\int f(t, \nu)\eta^n(dt, d\nu) \rightarrow \int f(t, \nu)\eta^0(dt, d\nu)$ for all $f \in \mathcal{H}_0$ as well.

Now let us fix $f \in \mathcal{H}_0$ with $f(t, \nu) = 1_B(t) \cdot g(\nu)$. Note that $\mathcal{P}(A)$ is a compact metric space and g is a bounded, uniformly continuous and real-valued function. A famous result from [13] (see also [20]) shows that

g can be approximated uniformly by bounded Lipschitz continuous function. That is, for all $\epsilon > 0$, we can find $g_\epsilon \in \mathcal{H}$ such that $\sup_{\nu \in \mathcal{P}(A)} |g_\epsilon(\nu) - g(\nu)| \leq \epsilon/3$. By our assumption we have $\int 1_B(t)g_\epsilon(\nu)\eta^n(dt, d\nu) \rightarrow \int 1_B(t)g_\epsilon(\nu)\eta^0(dt, d\nu)$. Therefore there exists N_0 such that $|\int 1_B(t)g_\epsilon(\nu)\eta^n(dt, d\nu) - \int 1_B(t)g_\epsilon(\nu)\eta^0(dt, d\nu)| \leq \epsilon/3$ for all $n \geq N_0$. Combining these facts we have, for $n \geq N_0$:

$$\begin{aligned} |\int 1_B(t)g(\nu)\eta^n(dt, d\nu) - \int 1_B(t)g(\nu)\eta^0(dt, d\nu)| &\leq |\int 1_B(t)g_\epsilon(\nu)\eta^n(dt, d\nu) - \int 1_B(t)g_\epsilon(\nu)\eta^0(dt, d\nu)| \\ &\quad + \int 1_B(t)|g_\epsilon(\nu) - g(\nu)|\eta^n(dt, d\nu) + \int 1_B(t)|g_\epsilon(\nu) - g(\nu)|\eta^0(dt, d\nu) \\ &\leq \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon, \end{aligned}$$

which shows that $\int f(t, \nu)\eta^n(dt, d\nu) \rightarrow \int f(t, \nu)\eta^0(dt, d\nu)$. $\square$

Now we consider the following subset of $\mathcal{R}$:

$$\mathcal{R}_0 := \{\eta \in \mathcal{R}; \text{ the marginal distribution of } \eta \text{ on } [0, T] \text{ is } \mathcal{L}\}.$$

We have the following result:

Lemma 4.7. $\mathcal{R}_0$ is a convex and compact subset of $\mathcal{R}$.

Proof. We apply Theorem 2.8 in [17]. In particular, we verify without difficulty that $\{\eta^{[0, T]}; \eta \in \mathcal{R}_0\} = \{\mathcal{L}\}$ is compact and $\{\eta^{\mathcal{P}(A)}; \eta \in \mathcal{R}_0\}$ is a subset of $\mathcal{P}(\mathcal{P}(A))$, which is relatively compact as well. $\square$

For any $\eta \in \mathcal{R}_0$, since its first marginal is $\mathcal{L}$, by disintegration we can write $\eta(dt, dm) = \mathcal{L}(dt) \times \eta_t(dm)$ where the mapping $[0, T] \ni t \rightarrow \eta_t(\cdot) \in \mathcal{P}(\mathcal{P}(A))$ is a measurable mapping and the decomposition is unique up to almost everywhere equality. On the other hand, for any measurable function $\nu : [0, T] \rightarrow \mathcal{P}(A)$, we may construct an element $\Psi(\nu)$ in $\mathcal{R}_0$ by:

$$(37) \quad \Psi(\nu)(dt, dm) := \mathcal{L}(dt) \times \delta_{\nu_t}(dm).$$

Since we have changed the way we represent the mean field of control, we need to modify accordingly the definition of transition rate matrix as well as the cost functionals in order to make them compatible with the randomization procedure. For any function $F : \mathcal{P}(A) \rightarrow \mathbb{R}$ possibly containing other arguments, we denote $\underline{F} : \mathcal{P}(\mathcal{P}(A)) \rightarrow \mathbb{R}$ by $\underline{F}(m) := \int_{\nu \in \mathcal{P}(A)} F(\nu)m(d\nu)$, which we call the randomized version of F . Obviously we have $\underline{F}(\delta_\nu) = F(\nu)$. In this way, we define without any ambiguity the randomized version $\underline{q}$ of the rate function q , as well as its matrix representation $\underline{Q}$. We also define $\underline{f}$ as the randomized version of cost functional f . Since the terminal cost g does not depend on the mean field of control, we do not need to consider its randomized version.

Recall from Assumption 3.4 that the minimizer $\hat{a}_i$ of the reduced Hamiltonian is only a function of t, z and p . Consequently, for $\underline{H}$, $\underline{H}_i$, $\hat{\underline{H}}$ and $\hat{\underline{H}}^i$, which are the randomized version of H , H_i , $\hat{H}$ and $\hat{H}_i$ respectively, we still have:

$$\begin{aligned} \hat{\underline{H}}(t, x, z, p, m) &= \inf_{\alpha \in A} \underline{H}(t, x, z, \alpha, p, m), \\ \hat{a}(t, x, z, p) &= \arg \inf_{\alpha \in A} \underline{H}(t, x, z, \alpha, p, m). \end{aligned}$$

In addition, we have the following result on the Lipschitz property of $\hat{\underline{H}}$ and $\hat{a}$:

Lemma 4.8. *There exists a constant $C > 0$ such that for all $(\omega, t) \in \Omega \times (0, T]$, $p, p' \in \mathcal{S}$, $\alpha, \alpha' \in A$, $z, z' \in \mathbb{R}^m$ and $m, m' \in \mathcal{P}(\mathcal{P}(A))$, we have:*

$$(38) \quad |\underline{q}(t, i, j, \alpha, p, m) - \underline{q}(t, i, j, \alpha', p', m')| \leq C(\|\alpha - \alpha'\| + \|p - p'\| + \bar{\mathcal{W}}_1(m, m')),$$

and

$$(39) \quad |\hat{\underline{H}}(t, X_{t-}, z, p, m) - \hat{\underline{H}}(t, X_{t-}, z', p', m')| \leq C\|z - z'\|_{X_{t-}} + C(1 + \|z\|_{X_{t-}})(\|p - p'\| + \bar{\mathcal{W}}_1(m, m')).$$

Proof. We have:

$$\begin{aligned}
& |\hat{\underline{H}}(t, X_{t-}, z, p, m) - \hat{\underline{H}}(t, X_{t-}, z', p', m')| \\
& \leq |\hat{\underline{H}}(t, X_{t-}, z, p, m) - \hat{\underline{H}}(t, X_{t-}, z, p, m')| + |\hat{\underline{H}}(t, X_{t-}, z, p, m') - \hat{\underline{H}}(t, X_{t-}, z', p', m')| \\
& \leq \left| \int_{\nu \in \mathcal{P}(A)} \hat{H}(t, X_{t-}, z, p, \nu) (m(d\nu) - m'(d\nu)) \right| \\
& \quad + \int_{\nu \in \mathcal{P}(A)} |\hat{H}(t, X_{t-}, z, p, \nu) - \hat{H}(t, X_{t-}, z', p', \nu)| m'(d\nu) \\
& \leq \left| \int_{\nu \in \mathcal{P}(A)} \hat{H}(t, X_{t-}, z, p, \nu) (m(d\nu) - m'(d\nu)) \right| + C(1 + \|z\|_{X_{t-}}) \|p - p'\| + C\|z - z'\|_{X_{t-}}.
\end{aligned}$$

Since the space $\mathcal{P}(A)$ is compact and the mapping $\nu \rightarrow \hat{H}(t, X_{t-}, z, p, \nu)$ is Lipschitz, with Lipschitz constant equal to $C(1 + \|z\|_{X_{t-}})$, Kantorovich-Rubinstein duality theory implies:

$$\left| \int_{\nu \in \mathcal{P}(A)} \hat{H}(t, X_{t-}, z, p, \nu) (m(d\nu) - m'(d\nu)) \right| \leq C(1 + \|z\|_{X_{t-}}) \bar{\mathcal{W}}_1(m, m'),$$

where $\bar{\mathcal{W}}_1(m, m')$ is the Wasserstein-1 distance on the space of probability measure on $\mathcal{P}(A)$ whose definition we recall for the sake of definiteness:

$$(40) \quad \bar{\mathcal{W}}_1(m, m') := \inf_{\pi \in \mathcal{P}(\mathcal{P}(A) \times \mathcal{P}(A))} \int_{\mathcal{P}(A) \times \mathcal{P}(A)} \mathcal{W}_1(\nu, \nu') \pi(d\nu, d\nu').$$

Combined with the estimation above, we obtain the desired inequality for $\hat{\underline{H}}$. The Lipschitz property for $\underline{g}$ can be proved in the same way. $\square$

4.2. Mapping fixed points. We now define the mapping whose fixed points characterize the Nash equilibria of the mean field game in its weak formulation. For any $(\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0$, where η has the disintegration $\eta(dt, dm) = \mathcal{L}(dt) \times \eta_t(dm)$, we consider the solution $(\mathbf{Y}^{(\mu, \eta)}, \mathbf{Z}^{(\mu, \eta)})$ to the BSDE:

$$(41) \quad Y_t = g(X_T, p_T) + \int_t^T \hat{\underline{H}}(s, X_{s-}, Z_s, \pi(s, \mu), \eta_s) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s.$$

Denote by $\hat{\alpha}^{(\mu, \eta)}$ the predictable process $t \rightarrow \hat{a}(t, X_{t-}, Z_t^{(\mu, \eta)}, \pi(t, \mu))$, which is the optimal control of the player faced with the mean field $(\mu, \eta) \in \mathcal{P}(E) \times \mathcal{R}_0$. Next, we consider the scalar martingale $\mathbf{L}^{(\mu, \eta)}$ defined by:

$$(42) \quad L_t^{(\mu, \eta)} := \int_0^t X_{s-}^* \cdot (\underline{Q}(s, \hat{\alpha}_s^{(\mu, \eta)}, \pi(s, \mu), \eta_s) - Q^0) \cdot d\mathcal{M}_s.$$

Define the probability measure $\hat{\mathbb{Q}}^{(\mu, \eta)}$ by:

$$(43) \quad \frac{d\hat{\mathbb{Q}}^{(\mu, \eta)}}{d\mathbb{P}} := \mathcal{E}(\mathbf{L}^{(\mu, \eta)})_T,$$

where $\mathcal{E}(\mathbf{L}^{(\mu, \eta)})$ is the Doléans-Dade exponential of the martingale $\mathbf{L}^{(\mu, \eta)}$. Finally we define the mappings Φ^μ , Φ^η and Φ respectively by:

$$(44) \quad \Phi^\mu : \mathcal{P} \times \mathcal{R}_0 \ni (\mu, \eta) \rightarrow \hat{\mathbb{Q}}^{(\mu, \eta)} \in \mathcal{P}$$

$$(45) \quad \Phi^\eta : \mathcal{P} \times \mathcal{R}_0 \ni (\mu, \eta) \rightarrow \mathcal{L}(dt) \times \delta_{\hat{\alpha}_t^{(\mu, \eta)}}(d\nu) \in \mathcal{R}_0$$

$$(46) \quad \Phi : \mathcal{P} \times \mathcal{R}_0 \ni (\mu, \eta) \rightarrow (\Phi^\mu(\mu, \eta), \Phi^\eta(\mu, \eta)) \in \mathcal{P} \times \mathcal{R}_0.$$

Remark 4.9. Before delving into its properties of Φ , we first need to show that the mapping Φ is well-defined. More specifically, we need to show that given $(\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0$, the outputs $\hat{\mathbb{Q}}^{(\mu, \eta)}$ and $\mathcal{L}(dt) \times \delta_{\hat{\mathbb{Q}}^{(\mu, \eta)}_{\# \hat{\alpha}_t^{(\mu, \eta)}}}(d\nu)$ does

not depend on which solution to the BSDE (41) we use to construct $\hat{\alpha}^{(\mu, \eta)}$, $\mathbf{L}^{(\mu, \eta)}$ and $\mathcal{E}(L^{(\mu, \eta)})$. To this end, let us consider $(\mathbf{Y}, \mathbf{Z})$ and $(\mathbf{Y}', \mathbf{Z}')$ two solutions to BSDE (41), $\hat{\alpha}$ and $\hat{\alpha}'$ the corresponding optimal controls, $\mathbf{L}$ and $\mathbf{L}'$ the corresponding martingales defined in (42), and $\mathbb{Q}$ and $\mathbb{Q}'$ the resulting probability measures defined in (43). By uniqueness of solution to (41), we have $\mathbb{E} \left[\int_0^T \|Z'_t - Z_t\|_{X_{t-}}^2 dt \right] = 0$. Using the Lipschitz continuity of $\hat{a}$ and q , it is straightforward to show $\mathbb{E} \left[\int_0^T \|\alpha'_t - \alpha_t\|^2 dt \right] = 0$ and eventually $\mathbb{Q} = \mathbb{Q}'$.

Proposition 4.10. *Let us denote by $\bar{\mathcal{P}}_0$ the closure of the set $\mathcal{P}_0$ defined in (35). Then the set $\bar{\mathcal{P}}_0 \times \mathcal{R}_0$ is stable for the mapping Φ .*

Proof. It suffices to show that for all $(\mu, \eta) \in \mathcal{P} \times \mathcal{R}_0$, we have $\Phi^\mu(\mu, \eta) \in \mathcal{P}_0$. By the definition of $\mathcal{P}_0$ in (35), this boils down to showing that there exists a constant $C_0 > 0$ such that for all (μ, η) , we have:

$$\mathbb{E}^\mathbb{P}[(\mathcal{E}(\mathbf{L}^{(\mu, \eta)}))_T]^2 \leq C_0.$$

Let us denote $W_t := \mathcal{E}(\mathbf{L}^{(\mu, \eta)})_t$. By Itô's lemma we have:

$$d(W_t^2) = 2W_{t-}dW_t + d[W, W]_t$$

since $dL_t^{(\mu, \eta)} = X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(\mu, \eta)}, \pi(t, \mu), \eta_t) - Q^0) \cdot \psi_t^+ \cdot d\mathcal{M}_t$ and $dW_t = W_{t-}dL_t^{(\mu, \eta)}$, denoting $I_t := \psi_t^+ \cdot (\underline{Q}^*(t, \hat{\alpha}_t^{(\mu, \eta)}, \pi(t, \mu), \eta_t) - Q^0) \cdot X_{t-}$ we have:

$$d(W_t^2) = 2W_{t-}^2 dL_t^{(\mu, \eta)} + W_{t-}^2 I_t^* \cdot d[\mathcal{M}, \mathcal{M}]_t \cdot I_t.$$

We know that the optional quadratic variation of $\mathcal{M}$ can be decomposed as:

$$[\mathcal{M}, \mathcal{M}]_t = G_t + \langle \mathcal{M}, \mathcal{M} \rangle_t = G_t + \int_0^t \psi_s ds,$$

where G is a martingale. Therefore we have:

$$d(W_t^2) = 2W_{t-}^2 dL_t^{(\mu, \eta)} + W_{t-}^2 I_t^* \cdot dG_t \cdot I_t + W_{t-}^2 I_t^* \cdot \psi_t \cdot I_t dt.$$

Let T_n be a sequence of stopping time converging to $+\infty$ which localizes both the local martingales $\int_0^t W_{s-}^2 dL_s^{(\mu, \eta)}$ and $\int_0^t W_{s-}^2 I_s^* \cdot dG_s \cdot I_s$. Then integrating the above SDE between 0 and $T \wedge T_n$ and taking the expectation under $\mathbb{P}$ we obtain:

$$\begin{aligned} \mathbb{E}^\mathbb{P}[W_{T \wedge T_n}^2] &= 1 + \mathbb{E}^\mathbb{P} \left[\int_0^{T \wedge T_n} W_{t-}^2 I_t^* \cdot \psi_t \cdot I_t dt \right] = 1 + \mathbb{E}^\mathbb{P} \left[\int_0^{T \wedge T_n} W_{t-}^2 I_t^* \cdot \psi_t \cdot I_t dt \right] \\ &\leq 1 + \int_0^T \mathbb{E}^\mathbb{P} [W_{t \wedge T_n}^2 I_{t \wedge T_n}^* \cdot \psi_{t \wedge T_n} \cdot I_{t \wedge T_n}] dt \leq 1 + C_0 \int_0^T \mathbb{E}^\mathbb{P}[W_{t \wedge T_n}^2] dt. \end{aligned}$$

Here we have used Tonelli's theorem as well as the fact that $I_s^* \cdot \psi_s \cdot I_s$ is bounded by a constant C_0 independent of μ, η and n , which is a consequence of the boundedness of the transition rate function q . Now applying Gronwall's lemma we obtain:

$$\mathbb{E}^\mathbb{P}[W_{T \wedge T_n}^2] \leq C_0.$$

where the constant C_0 does not depend on n, μ or η . Notice that $W_{T \wedge T_n}^2$ converges to W_T^2 almost surely, we apply Fatou's lemma and obtain $\mathbb{E}^\mathbb{P}[W_T^2] \leq C_0$. $\square$

4.3. Existence of Nash equilibria. The last missing piece in applying Schauder's fixed point theorem is to show the continuity of the mapping Φ on $\mathcal{P} \times \mathcal{R}_0$ for the product topology. To this end, we show the continuity of the mappings Φ^μ and Φ^η . Notice that both $\mathcal{P}$ and $\mathcal{R}_0$ are metrizable, so we only need to show sequential continuity.

Let us fix a sequence $(\mu^{(n)}, \eta^{(n)})_{n \geq 1}$ converging to $(\mu^{(0)}, \eta^{(0)})$ in $\mathcal{P} \times \mathcal{R}_0$, with the decomposition $\eta^{(n)}(dt, d\nu) = \mathcal{L}(dt) \times \eta_t^{(n)}(d\nu)$. To simplify the notation we denote $\mathbf{Y}^{(\mu^{(n)}, \eta^{(n)})}$, $\mathbf{Z}^{(\mu^{(n)}, \eta^{(n)})}$, $\hat{\alpha}^{(\mu^{(n)}, \eta^{(n)})}$, $\mathbf{L}^{(\mu^{(n)}, \eta^{(n)})}$, $\hat{\mathbb{Q}}^{(\mu^{(n)}, \eta^{(n)})}$ respectively by $\mathbf{Y}^{(n)}$, $\mathbf{Z}^{(n)}$, $\hat{\alpha}^{(n)}$, $\mathbf{L}^{(n)}$ and $\mathbb{Q}^{(n)}$ for $n \geq 0$. We also denote by $\mathbb{E}^{(n)}$ the expectation under $\mathbb{Q}^{(n)}$ and $p_t^{(n)} = \pi(t, \mu^{(n)})$, whereas $\mathbb{E}$ still denotes the expectation under the reference measure $\mathbb{P}$.

We start by proving the continuity of Φ^μ , or equivalently the convergence of $\mathbb{Q}^{(n)}$ toward $\mathbb{Q}^{(0)}$. We divide the proof into several intermediary results.

Lemma 4.11. *Without any loss of generality, we may assume that there exists a constant C such that $\|Z_t^{(0)}\|_{X_{t-}} \leq C$ for all $(\omega, t) \in \Omega \times [0, T]$.*

Proof. We consider the following ODE of unknown $V_t = [V_1(t), \dots, V_m(t)] \in \mathbb{R}^m$:

$$(47) \quad \begin{aligned} 0 &= \frac{dV_i(t)}{dt} + \hat{H}_i(t, V(t), p_t^{(0)}, \eta_t^{(0)}) + \sum_{j \neq i} [V_j(t) - V_i(t)], \\ V_i(T) &= g(e_i, p_T^{(0)}), \quad i = 1, \dots, m. \end{aligned}$$

Set $\zeta : [0, T] \times \mathbb{R}^m \ni (t, v) \rightarrow [\zeta_1(t, v), \dots, \zeta_m(t, v)] \in \mathbb{R}^m$ where $\zeta_i(t, v) := \hat{H}_i(t, v, p_t^{(0)}, \eta_t^{(0)}) + \sum_{j \neq i} [v_j - v_i]$. By Lemma 4.8, we see that $t \rightarrow \zeta(t, v)$ is measurable for all $v \in \mathbb{R}^m$ and $v \rightarrow \zeta(t, v)$ is Lipschitz in v uniformly in t . By Theorem 1 and Theorem 2 in [11], the ODE (47) admits a unique solution on the interval $[0, T]$, which is absolutely continuous. Now we define $Y_t = \sum_{i=1}^m \mathbb{1}(X_t = e_i) V_i(t)$ and $Z_t = V_t$. By continuity of V , we have $\Delta Y_t := Y_t - Y_{t-} = V_t^*(X_t - X_{t-}) = Z_t^* \cdot \Delta X_t$. Applying Ito's formula to Y , we obtain:

$$\begin{aligned} Y_t &= Y_T - \int_t^T \sum_{i=1}^m \mathbb{1}(X_s = e_i) \dot{V}_i(s) ds - \sum_{t < s \leq T} \Delta Y_s \\ &= g(X_T, p_T^{(0)}) + \int_t^T \sum_{i=1}^m \mathbb{1}(X_s = e_i) \hat{H}_i(s, V(s), p_s^{(0)}, \eta_s^{(0)}) ds \\ &\quad + \int_t^T \sum_{i=1}^m \mathbb{1}(X_s = e_i) \sum_{j \neq i} [V_j(s) - V_i(s)] ds - \int_t^T Z_s^* \cdot dX_s \\ &= g(X_T, p_T^{(0)}) + \int_t^T \hat{H}(s, X_s, Z_s, p_s^{(0)}, \eta_s^{(0)}) ds - \int_t^T Z_s^* \cdot d\mathcal{M}_s, \end{aligned}$$

where in the last equality we used the fact that $dX_s = Q^0 \cdot X_{s-} ds + d\mathcal{M}_s$ and $V_t = Z_t$. Therefore (Y, Z) and $(Y^{(0)}, Z^{(0)})$ solves the same BSDE. As we have discussed in Remark 4.9, we may assume that $Z^{(0)} = Z$. Therefore $Z_t^{(0)} = V(t)$. It follows from the continuity of $t \rightarrow V(t)$ that $\|Z_t^{(0)}\|_{X_{t-}}$ is bounded for all $\omega \in \Omega$ and $t \in [0, T]$ by a uniform constant. $\square$

Now we show that $Z^{(n)}$ converges toward $Z^{(0)}$.

Proposition 4.12. *We have:*

$$(48) \quad \lim_{n \rightarrow \infty} \mathbb{E} \left[\int_0^T \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}}^2 dt \right] = 0.$$

Proof. By Lemma 3.2, it suffices to check that:

$$I_n(t) := \mathbb{E} \left[\left(\int_t^T \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(n)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)}) ds \right)^2 \right]$$

converges to 0 for all $t \leq T$, and that $I_n(t)$ is bounded by C uniformly in t and n . We also need to check $J_n := \mathbb{E}[|g(X_T, p_T^{(n)}) - g(X_T, p_T^{(0)})|^2]$ converges to 0. By Lipschitz property of the cost functional g and Lemma 4.3, we have:

$$J_n \leq C\|p_T^{(n)} - p_T^{(0)}\|^2 = C\|\pi(\mu^{(n)}, T) - \pi(\mu^{(0)}, T)\|^2 \rightarrow 0,$$

as $n \rightarrow +\infty$. To check the uniform boundedness of $I_n(t)$, we recall from Lemma 3.7 that:

$$|\hat{H}(t, X_{t-}, Z_t^{(0)}, p_t^{(n)}, \eta_t^{(n)}) - \hat{H}(t, X_{t-}, Z_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})| \leq C(1 + \|Z_t^{(0)}\|_{X_{t-}})(\|p_t^{(n)} - p_t^{(0)}\| + \bar{\mathcal{W}}_1(\eta_t^{(n)}, \eta_t^{(0)})),$$

where $\bar{\mathcal{W}}_1$ is the Wasserstein distance on the space $\mathcal{P}(\mathcal{P}(A))$. Clearly $\|p_t^{(n)} - p_t^{(0)}\|$ can be bounded by a constant since $p_t^{(n)}$ is in the simplex $\mathcal{S}$. On the other hand, we have:

$$\bar{\mathcal{W}}_1(\eta_t^{(n)}, \eta_t^{(0)}) \leq \int_{(\nu_1, \nu_2) \in \mathcal{P}(A)^2} \mathcal{W}_1(\nu_1, \nu_2) \eta_t^{(n)}(d\nu_1) \eta_t^{(0)}(d\nu_2).$$

Since A is compact, $\mathcal{W}_1(\nu_1, \nu_2)$ for $(\nu_1, \nu_2) \in \mathcal{P}(A)^2$ is bounded, which implies that $\bar{\mathcal{W}}_1(\eta_t^{(n)}, \eta_t^{(0)})$ is also bounded by a constant uniformly in n and t . This implies:

$$I_n(t) \leq C\mathbb{E}\left[\int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}}) ds\right] \leq C(1 + (\mathbb{E}[\int_0^T \|Z_s^{(0)}\|_{X_{s-}}^2 ds])^{1/2}) < +\infty,$$

which means that $I_n(t)$ is uniformly bounded in n and t . To show that $I_n(t)$ converges to 0, we write:

$$\begin{aligned} I_n(t) &\leq 2\mathbb{E}\left[\left(\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(n)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)})) dt\right)^2\right] \\ &\quad + 2\mathbb{E}\left[\left(\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) dt\right)^2\right] \\ &\leq 2C\mathbb{E}\left[\int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}})^2 \|p_s^{(n)} - p_s^{(0)}\|^2 ds\right] \\ &\quad + 2\mathbb{E}\left[\left(\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) dt\right)^2\right]. \end{aligned}$$

By Lemma 4.3, we have $(1 + \|Z_s^{(0)}\|_{X_{s-}})^2 \|p_s^{(n)} - p_s^{(0)}\|^2 \rightarrow 0$, $dt \otimes d\mathbb{P}$ -a.e. On the other hand, we have:

$$(1 + \|Z_s^{(0)}\|_{X_{s-}})^2 \|p_s^{(n)} - p_s^{(0)}\|^2 \leq C(1 + \|Z_s^{(0)}\|_{X_{s-}})^2,$$

where the right hand side is $ds \otimes d\mathbb{P}$ -integrable. Therefore by the dominated convergence theorem, we obtain:

$$\mathbb{E}\left[\int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}})^2 \|p_s^{(n)} - p_s^{(0)}\|^2 ds\right] \rightarrow 0,$$

as $n \rightarrow +\infty$. It remains to show that:

$$K_n := \mathbb{E}\left[\left(\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) dt\right)^2\right]$$

converges to 0. For a fix $w \in \Omega$ and $s \leq T$, we have:

$$\int_t^T (\hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) ds$$

$$\begin{aligned}
&= \int_t^T \int_{\nu \in \mathcal{P}(A)} \hat{H}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu) (\eta_s^{(n)} - \eta_s^{(0)}) (d\nu) ds \\
&= \int_{[0, T] \times \mathcal{P}(A)} \kappa(s, \nu) \eta^{(n)}(ds, d\nu) - \int_{[0, T] \times \mathcal{P}(A)} \kappa(s, \nu) \eta^{(0)}(ds, d\nu),
\end{aligned}$$

where we defined $\kappa(s, \nu) := 1_{t \leq s \leq T} H(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)$. Clearly κ is continuous in ν for all s . On the other hand, by inequality (26) in Lemma 3.7, for all $t \leq s \leq T$ and $\nu \in \mathcal{P}(A)$ we have:

$$|H(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)| \leq |H(s, X_{s-}, 0, 0, 0)| + C \|Z_s^{(0)}\|_{X_{s-}} + C(1 + \|Z_s^{(0)}\|_{X_{s-}})(\|p_s^{(0)}\| + \mathcal{W}_1(\nu, 0)).$$

Therefore by Lemma 4.11 and the boundedness of $\mathcal{P}(A)$, we conclude that the mapping $(s, \nu) \rightarrow \kappa(s, \nu)$ is bounded. It follows from the definition of stable topology and $\eta^{(n)} \rightarrow \eta^{(0)}$ that:

$$\lim_{n \rightarrow +\infty} \int_t^T (\hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) ds = 0,$$

for all $w \in \Omega$. In addition, we have:

$$\begin{aligned}
&\left(\int_t^T (\hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})) ds \right)^2 \\
&\leq (T-t) \int_t^T |\hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(n)}) - \hat{\underline{H}}(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \eta_s^{(0)})|^2 ds \\
&\leq C \int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}})^2 (\bar{\mathcal{W}}_1(\eta_s^{(n)}, \eta_s^{(0)}))^2 ds \leq C \int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}})^2 ds
\end{aligned}$$

and $\int_t^T (1 + \|Z_s^{(0)}\|_{X_{s-}})^2 ds$ is integrable. Apply once again the dominated convergence theorem, we conclude that K_n converges to 0. This completes the proof. $\square$

We will also need a result on a more convenient representation of the Doléans-Dade exponential of $\mathbf{L}^{(n)}$.

Lemma 4.13. *Denote by $\mathbf{W}^{(n)}$ the Doléans-Dade exponential of $\mathbf{L}^{(n)}$. Then the Itô differential of $\log(\mathbf{W}^{(n)})$ satisfies:*

$$d[\log(W_t^{(n)})] = X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0 + \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \cdot Q^0) \cdot X_{t-} dt + X_{t-}^* \cdot \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \cdot d\mathcal{M}_t,$$

where $\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})$ is the matrix with $\log(\underline{q}(t, i, j, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}))$ as off-diagonal elements and zeros on the diagonal.

Proof. Since $\mathbf{W}^{(n)}$ is the Doléans-Dade exponential of $\mathbf{L}^{(n)}$, $\mathbf{W}^{(n)}$ satisfies the SDE $dW_t^{(n)} = W_{t-}^{(n)} dL_t^{(n)}$. Applying Ito's formula and noticing that the continuous martingale part of $\mathbf{L}^{(n)}$ is zero, we have:

$$d \log(W_t^{(n)}) = dL_t^{(n)} - \Delta L_t^{(n)} + \log(1 + \Delta L_t^{(n)}).$$

Then using $dL_t^{(n)} = X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0) \cdot \psi_t^+ \cdot d\mathcal{M}_t$ and noticing that the jumps of $\mathbf{L}^{(n)}$ are driven by the jumps of $\mathcal{M}$, and hence $\bar{X}$, we obtain:

$$\begin{aligned}
d \log(W_t^{(n)}) &= -X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0) \cdot \psi_t^+ \cdot Q^0 \cdot X_{t-} dt + \log(1 + \Delta L_t^{(n)}) \\
&= X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - Q^0) \cdot X_{t-} dt + \log(1 + \Delta L_t^{(n)}),
\end{aligned}$$

where we have used the fact that for all q -matrices A , we have $X_{t-}^* \cdot A \cdot \psi_t^+ \cdot Q^0 \cdot X_{t-} = -X_{t-}^* \cdot A \cdot X_{t-}$. Piggybacking on the derivation following equation (6), for $X_{t-} = e_i$ and $X_t = e_j$ we have:

$$\log(1 + \Delta L_t^{(n)}) = \log(\underline{q}(t, i, j, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})).$$

Using matrix notation and recalling the definition of $\underline{Q}$ in the statement of Lemma 4.13, we may write:

$$\log(1 + \Delta L_t^{(n)}) = X_{t-}^* \cdot \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) \cdot \Delta X_t.$$

Using again the equality $\Delta X_t = dX_t = Q^0 \cdot X_{t-} dt + d\mathcal{M}_t$, we arrive at the desired representation of the differential of $\log(W_t)$. $\square$

We now show the first component of the mapping Φ is sequentially continuous.

Proposition 4.14. $\mathbb{Q}^{(n)}$ converges to $\mathbb{Q}^{(0)}$ in $\mathcal{P}$.

Proof. For two probability measures $\mathbb{Q}$ and $\mathbb{Q}'$ in $\mathcal{P}$, the total variation distance d between $\mathbb{Q}$ and $\mathbb{Q}'$ is:

$$(49) \quad d_{TV}(\mathbb{Q}, \mathbb{Q}') := \sup\{|\mathbb{Q}(A) - \mathbb{Q}'(A)|, A \in \mathcal{B}(D)\}.$$

It is well-known that convergence in total variation implies weak convergence, hence convergence in the topological space $\mathcal{P}$. Therefore our aim is to show that $d_{TV}(\mathbb{Q}^{(n)}, \mathbb{Q}^{(0)}) \rightarrow 0$ as $n \rightarrow +\infty$. By Pinsker's inequality, we have:

$$d_{TV}^2(\mathbb{Q}^{(0)}, \mathbb{Q}^{(n)}) \leq \frac{1}{2} \mathbb{E}^{(0)} \left[\log \left(\frac{d\mathbb{Q}^{(0)}}{d\mathbb{Q}^{(n)}} \right) \right].$$

Since $\frac{d\mathbb{Q}^{(n)}}{d\mathbb{P}} = \mathcal{E}(\mathbf{L}^{(n)})_T$, we have:

$$d_{TV}^2(\mathbb{Q}^{(0)}, \mathbb{Q}^{(n)}) \leq \mathbb{E}^{(0)} [\log(\mathcal{E}(L^{(0)})_T) - \log(\mathcal{E}(L^{(n)})_T)].$$

Using Lemma 4.13, we have:

$$\begin{aligned} & \mathbb{E}^{(0)} [\log(\mathcal{E}(\mathbf{L}^{(0)})_T) - \log(\mathcal{E}(\mathbf{L}^{(n)})_T)] \\ &= \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)}) - \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \cdot X_{t-} dt \right] \\ &+ \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)}) - \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \cdot Q^0 \cdot X_{t-} dt \right] \\ &+ \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)}) - \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \cdot d\mathcal{M}_t \right]. \end{aligned}$$

By Assumption 2.2, the process $t \rightarrow \int_0^t X_{s-}^* \cdot (O(s, \hat{\alpha}_s^{(0)}, p_s^{(0)}, \nu_s^{(n)}) - O(s, \hat{\alpha}_s^{(n)}, p_s^{(n)}, \nu_s^{(n)})) \cdot d\mathcal{M}_s$ is a true martingale therefore have zero expectation. We now deal with the convergence of the term $\mathbb{E}^0[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot X_{t-} dt]$, whereas the term $\mathbb{E}^0[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot Q^0 \cdot X_{t-} dt]$ can be dealt with in the exact the same way. Using the Lipschitz property of $\hat{\alpha}$ and $\underline{Q}$ in Lemma 3.7 and Lemma 4.8, we obtain:

$$\begin{aligned} & \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)}) - \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \cdot X_{t-} dt \right] \\ & \leq \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})) \cdot X_{t-} dt \right] \\ & + \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(n)})) \cdot X_{t-} dt \right] \\ & \leq \mathbb{E}^{(0)} \left[\int_0^T C(\|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\| + \|p_t^{(n)} - p_t^{(0)}\|) dt \right] \end{aligned}$$

$$\begin{aligned}
& + \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot X_{t-} dt \right] \\
& \leq \mathbb{E}^{(0)} \left[\int_0^T C \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}} dt \right] + \mathbb{E}^{(0)} \left[\int_0^T C(1 + \|Z_t^{(0)}\|_{X_{t-}}) \|p_t^{(n)} - p_t^{(0)}\| dt \right] \\
& + \mathbb{E}^{(0)} \left[\int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot X_{t-} dt \right].
\end{aligned}$$

We deal with these terms separately. For the first expectation, by Cauchy-Schwartz inequality, we have:

$$\begin{aligned}
\left(\mathbb{E}^{(0)} \left[\int_0^T C \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}} dt \right] \right)^2 &= \left(\mathbb{E} \left[W_T^{(0)} \int_0^T C \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}} dt \right] \right)^2 \\
&\leq \mathbb{E}[(W_T^{(0)})^2] \mathbb{E} \left[\left(\int_0^T C \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}} dt \right)^2 \right] \\
&\leq C \mathbb{E}[(W_T^{(0)})^2] \mathbb{E} \left[\int_0^T \|Z_t^{(n)} - Z_t^{(0)}\|_{X_{t-}}^2 dt \right].
\end{aligned}$$

This converges to 0 by Proposition 4.12. For the second expectation, we notice from Lemma 4.11 that $\|Z_t^{(0)}\|_{X_{t-}}$ is bounded by a constant for all $(\omega, t) \in \Omega \times [0, T]$. Therefore we have:

$$\mathbb{E}^{(0)} \left[\int_0^T C(1 + \|Z_t^{(0)}\|_{X_{t-}}) \|p_t^{(n)} - p_t^{(0)}\| dt \right] \leq C \int_0^T C \|p_t^{(n)} - p_t^{(0)}\| dt,$$

where the right-hand side converges to 0 by dominated convergence theorem. Finally for the third expectation, we rewrite the integrand as:

$$\begin{aligned}
& \int_0^T X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(n)}) - \underline{Q}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot X_{t-} dt \\
&= \int_{[0, T] \times \mathcal{P}(A)} X_{t-}^* \cdot Q(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \nu) \cdot X_{t-} (\eta^{(n)}(dt, d\nu) - \eta^{(0)}(dt, d\nu)).
\end{aligned}$$

This converges to 0, since $\eta^{(n)}$ converges to $\eta^{(0)}$ in stable topology and the mapping $\nu \rightarrow Q(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \nu)$ is continuous for all t . Notice also that the integrand is bounded by a constant, since q is bounded according to Assumption 2.2. Then by dominated convergence theorem the third expectation converges to 0 as well. This completes the proof. $\square$

To show the continuity of Φ^η , we need the following lemma.

Lemma 4.15. *Let $(\nu_t^{(n)})_{t \leq T}$ be a sequence of measurable functions from $[0, T]$ to $\mathcal{P}(A)$ such that $\int_0^T \mathcal{W}_1(\nu_t^{(n)}, \nu_t^{(0)}) \rightarrow 0$. Then $\mathcal{L}(dt) \times \delta_{\nu_t^{(n)}}(d\nu)$ converges to $\mathcal{L}(dt) \times \delta_{\nu_t^{(0)}}(d\nu)$ in $\mathcal{R}_0$ in the sense of the stable topology.*

Proof. Set $\lambda^{(n)}(dt, d\nu) := \mathcal{L}(dt) \times \delta_{\nu_t^{(n)}}(d\nu)$ for $n \geq 0$, let $f : [0, T] \times \mathcal{P}(A) \rightarrow \mathbb{R}$ be a mapping of the form $f(t, \nu) = 1_{t \in B} \cdot g(\nu)$ where B is measurable subset of $[0, T]$ and g is a bounded Lipschitz function on $\mathcal{P}(A)$. We then have:

$$\left| \int_{[0, T] \times \mathcal{P}(A)} f(t, \nu) \lambda^{(n)}(dt, d\nu) - \int_{[0, T] \times \mathcal{P}(A)} f(t, \nu) \lambda^{(0)}(dt, d\nu) \right| \leq \int_{t \in B} |g(\nu_t^{(n)}) - g(\nu_t^{(0)})| dt \leq C \int_0^T \mathcal{W}_1(\nu_t^{(n)}, \nu_t^{(0)}) dt.$$

By Lemma 4.6, we conclude that $\lambda^{(n)}$ converges to $\lambda^{(0)}$ for the stable topology. $\square$

Proposition 4.16. $\mathcal{L}(\cdot) \times \delta_{\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}}(\cdot)$ converges to $\mathcal{L}(\cdot) \times \delta_{\mathbb{Q}_{\# \hat{\alpha}_t^{(0)}}}(\cdot)$ in $\mathcal{R}_0$ in the sense of the stable topology.

Proof. By Lemma 4.15, we only need to show $\int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(0)}}^{(0)}) dt$ converges to 0. Notice that:

$$\int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(0)}}^{(0)}) dt \leq \int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}) dt + \int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(0)}}^{(0)}) dt.$$

By the very definition of the total variation distance (recall equation 49), we have clearly:

$$d_{TV}(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}) \leq d_{TV}(\mathbb{Q}^{(n)}, \mathbb{Q}^{(0)}),$$

which converges to 0 according to the proof of Proposition 4.14. By Theorem 6.16 in [24], since A is bounded and $\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)} \in \mathcal{P}(A)$, there exists a constant C such that:

$$\mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}) \leq C \cdot d_{TV}(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}).$$

This shows that $\mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)})$ converges to 0. In addition, it is also bounded since A is bounded. The dominated convergence theorem then implies that:

$$\int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(n)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}) dt \rightarrow 0, n \rightarrow +\infty.$$

Now for the other term, we have:

$$\begin{aligned} \int_0^T \mathcal{W}_1(\mathbb{Q}_{\# \hat{\alpha}_t^{(n)}}^{(0)}, \mathbb{Q}_{\# \hat{\alpha}_t^{(0)}}^{(0)}) dt &\leq \int_0^T \mathbb{E}^{(0)}[\|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\|] dt \\ &= \mathbb{E} \left[W_T^{(0)} \int_0^T \|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\| dt \right] \\ &\leq (\mathbb{E}[(W_T^{(0)})^2])^{1/2} \left(\mathbb{E} \left[T \int_0^T \|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\|^2 dt \right] \right)^{1/2}. \end{aligned}$$

The Lipschitz property of $\hat{\alpha}$ (see Lemma 3.7) and Proposition 4.12 imply that $\mathbb{E} \left[T \int_0^T \|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\|^2 dt \right] \rightarrow 0$. $\square$

We are now ready to show the existence of Nash equilibria.

Proof. (of Theorem 4.1) Consider the product space $\Gamma := \mathcal{P} \times \mathcal{R}$ endowed with the product topology of the weak topology on $\mathcal{P}$ and the stable topology on $\mathcal{R}$. By Proposition 4.5, Γ is a Polish space. By Proposition 4.2 and Lemma 4.7, $\Gamma_0 := \bar{\mathcal{P}}_0 \times \mathcal{R}_0$ is a compact and convex subset of Γ and it is stable by the mapping Φ defined in (46). In addition, we see from Proposition 4.14 and Proposition 4.16 that Φ is continuous. Therefore applying Schauder's fixed point theorem, we conclude that Φ admits a fixed point $(\mu^*, \eta^*) \in \bar{\mathcal{P}}_0 \times \mathcal{R}_0$.

Now let us define $p_t^* := \pi(t, \mu^*) \in \mathcal{S}$ and $\alpha_t^* := \hat{\alpha}(t, X_t, Z_t^*, \pi(t, \mu^*))$ where $(\mathbf{Y}^*, \mathbf{Z}^*)$ is the solution to the BSDE (41) with $\mu = \mu^*$ and $\eta = \eta^*$. We then define $\mathbb{P}^* := \mathbb{P}^{\mu^*, \eta^*}$ and $\nu_t^* := \mathbb{P}_{\# \alpha_t^*}^*$. Since (μ^*, η^*) is the fixed point of the mapping Φ , we have $\eta_t^* = \delta_{\nu_t^*}$ and $\mu^* = \mathbb{P}^*$. It follows that $p_t^* = \pi(t, \mathbb{P}^*) = [\mathbb{P}^*(X_t = e_i)]_{1 \leq i \leq m}$. By Proposition 3.9, we see that α^* is the solution to the optimal control problem (13) when the mean field of state is $\mathbf{p}^*$ and the mean field of control is ν^* . This implies that $(\alpha^*, \mathbf{p}^*, \nu^*)$ is a Nash equilibrium. $\square$

5. UNIQUENESS OF NASH EQUILIBRIUM

Uniqueness of Nash equilibria will be proven under the following conditions.

Assumption 5.1. (i) The transition rate function q does depend neither on the mean field of state $\mathbf{p}$ nor on the mean field of control $\boldsymbol{\nu}$. The cost functional f is separable in the sense that it is of the form:

$$(50) \quad f(t, x, \alpha, p, \nu) = f_0(t, x, \alpha) + f_1(t, X, p) + f_2(t, p, \nu).$$

(ii) For all $t \in [0, T]$, $i \in \{1, \dots, m\}$, $z \in \mathbb{R}^m$, $p \in \mathcal{S}$ and $\nu \in \mathcal{P}(A)$, the mapping $\alpha \rightarrow H_i(t, z, \alpha, p, \nu)$ admits a unique minimizer, which because of assumption (i), only depends on t and z . We denote it by $\hat{a}_i(t, z)$. In addition, we assume that $\hat{a}_i$ is a measurable from $[0, T] \times \mathbb{R}^m$ into A , and that there exists a constant $C > 0$ such that for all $i \in \{1, \dots, m\}$, and $z, z' \in \mathbb{R}^m$:

$$(51) \quad \|\hat{a}_i(t, z) - \hat{a}_i(t, z')\| \leq C \|z - z'\|_{e_i}.$$

(iii) For all $p, p' \in \mathcal{S}$ and $t \in [0, T]$, we have:

$$(52) \quad \sum_{i=1}^m (g(e_i, p) - g(e_i, p'))(p_i - p'_i) \geq 0,$$

$$(53) \quad \sum_{i=1}^m (f_1(t, e_i, p) - f_1(t, e_i, p'))(p_i - p'_i) \geq 0.$$

Remark 5.2. Item (ii) of Assumption 5.1 holds if we impose additional conditions of linearity and strong convexity on the transition rate function and the cost function, for example:

Assumption 5.3. (i) The transition rate function q takes the form $q(t, i, j, \alpha) = q_0(t, i, j) + q_1(t, i, j) \cdot \alpha$, where the mappings $q_0 : [0, T] \times E^2 \rightarrow \mathbb{R}$ and $q_1 : [0, T] \times E^2 \rightarrow \mathbb{R}^l$ are continuous.

(ii) f_0 is γ -strongly convex in α , i.e., for all $(t, i) \in [0, T] \times E$ and $\alpha, \alpha' \in A$, we have:

$$(54) \quad f_0(t, e_i, \alpha) - f_0(t, e_i, \alpha') - (\alpha - \alpha') \cdot \nabla_{\alpha} f_0(t, e_i, \alpha) \geq \gamma \|\alpha' - \alpha\|^2$$

Theorem 5.4. Under Assumptions 2.2, 2.7 and 5.1, there exists at most one Nash equilibrium for the weak formulation of the finite state mean field game.

Proof. Let $(\boldsymbol{\alpha}^{(1)}, \mathbf{p}^{(1)}, \boldsymbol{\nu}^{(1)})$ and $(\boldsymbol{\alpha}^{(2)}, \mathbf{p}^{(2)}, \boldsymbol{\nu}^{(2)})$ be two Nash equilibria of the mean field game. For $i = 1, 2$, we denote by $(\mathbf{Y}^{(i)}, \mathbf{Z}^{(i)})$ the solution to the BSDE (29) with $\mathbf{p} = \mathbf{p}^{(i)}$, $\boldsymbol{\nu} = \boldsymbol{\nu}^{(i)}$, which is written as:

$$Y_0^{(i)} = g(X_T, p_T^{(i)}) + \int_0^T \hat{H}(t, X_{t-}, Z_t^{(i)}, p_t^{(i)}, \nu_t^{(i)}) dt - \int_0^T (Z_t^{(i)})^* \cdot d\mathcal{M}_t$$

we have $\hat{\alpha}_t^{(i)} := \hat{a}(t, Z_t^{(i)})$, $d\mathbb{P} \otimes dt$ -a.e. Let us denote $\mathbb{Q}^{(i)} := \mathbb{Q}^{(\hat{\alpha}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\nu}^{(i)})}$, the controlled probability measure defined in (7), under which $\mathcal{M}^{(i)} := \mathcal{M}^{(\hat{\alpha}^{(i)}, \mathbf{p}^{(i)}, \boldsymbol{\nu}^{(i)})}$ is a martingale. In addition, we use the abbreviation $f_t^{(i)} := f(t, X_{t-}, \hat{\alpha}_t^{(i)}, p_t^{(i)}, \nu_t^{(i)})$, $g^{(i)} = g(X_T, p_T^{(i)})$ and $Q_t^{(i)} := Q(t, \hat{\alpha}_t^{(i)})$. Taking the difference of the BSDEs we obtain:

$$\begin{aligned} Y_0^{(1)} - Y_0^{(2)} &= g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)} - X_{t-}^* \cdot (Q_t^{(2)} - Q_t^{(0)}) \cdot Z_t^{(2)}) dt \\ &\quad + \int_0^T (Z_t^{(1)} - Z_t^{(2)})^* \cdot d\mathcal{M}_t \\ &= g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) dt + \int_0^T (Z_t^{(1)} - Z_t^{(2)})^* \cdot d\mathcal{M}_t^{(2)} \\ &= g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)}) dt + \int_0^T (Z_t^{(1)} - Z_t^{(2)})^* \cdot d\mathcal{M}_t^{(1)}. \end{aligned}$$

Taking expectations with respect to $\mathbb{Q}^{(1)}$ and $\mathbb{Q}^{(2)}$ and using the fact that

$$\mathbb{E}^{\mathbb{P}}[Y_0^{(1)} - Y_0^{(2)}] = \mathbb{E}^{\mathbb{Q}^{(1)}}[Y_0^{(1)} - Y_0^{(2)}] = \mathbb{E}^{\mathbb{Q}^{(2)}}[Y_0^{(1)} - Y_0^{(2)}],$$

we obtain the following equality:

$$(55) \quad \begin{aligned} & \mathbb{E}^{\mathbb{Q}^{(1)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)}) dt \right] \\ &= \mathbb{E}^{\mathbb{Q}^{(2)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) dt \right]. \end{aligned}$$

Next we notice that:

$$\begin{aligned} f_t^1 + X_{t-}^* \cdot Q_t^{(1)} \cdot Z_t^2 &= f(t, X_{t-}, \hat{\alpha}_t^{(1)}, p_t^{(1)}, \nu_t^{(1)}) + X_{t-}^* \cdot Q(t, \hat{\alpha}_t^{(1)}) \cdot Z_t^{(2)} \\ &= H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) \geq H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) \\ &= H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(2)}, \nu_t^{(2)}) + (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) \\ &\quad + (f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})) \end{aligned}$$

and using the inequality:

$$H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) \geq H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$$

which is due to the fact that $\hat{\alpha}_t^{(2)}$ minimizes the Hamiltonian $\alpha \rightarrow H(t, X_{t-}, \alpha, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$ and Assumption 5.1 that the minimizer does not depend on the mean field terms, we get:

$$f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)} \geq (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) + (f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})).$$

Interchanging the indices we obtain:

$$f_t^{(2)} - f_t^{(1)} + X_{t-}^* \cdot (Q_t^{(2)} - Q_t^{(1)}) \cdot Z_t^{(1)} \geq (f_1(t, X_{t-}, p_t^{(2)}) - f_1(t, X_{t-}, p_t^{(1)})) + (f_2(t, p_t^{(2)}, \nu_t^{(2)}) - f_2(t, p_t^{(1)}, \nu_t^{(1)})).$$

Injecting these inequalities into equation (55) we have:

$$\begin{aligned} 0 &= \mathbb{E}^{\mathbb{Q}^{(1)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(2)}) dt \right] \\ &\quad - \mathbb{E}^{\mathbb{Q}^{(2)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_t^{(1)} - f_t^{(2)} + X_{t-}^* \cdot (Q_t^{(1)} - Q_t^{(2)}) \cdot Z_t^{(1)}) dt \right] \\ &\geq \mathbb{E}^{\mathbb{Q}^{(1)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)}) + f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})) dt \right] \\ &\quad - \mathbb{E}^{\mathbb{Q}^{(2)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)}) + f_2(t, p_t^{(1)}, \nu_t^{(1)}) - f_2(t, p_t^{(2)}, \nu_t^{(2)})) dt \right] \\ &= \mathbb{E}^{\mathbb{Q}^{(1)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) dt \right] \\ &\quad - \mathbb{E}^{\mathbb{Q}^{(2)}} \left[g^{(1)} - g^{(2)} + \int_0^T (f_1(t, X_{t-}, p_t^{(1)}) - f_1(t, X_{t-}, p_t^{(2)})) dt \right] \end{aligned}$$

where the last equality is due to the fact that $[f_2(t, p_t^{(2)}, \nu_t^{(2)}) - f_2(t, p_t^{(1)}, \nu_t^{(1)})]$ is deterministic.

From Proposition 3.9, since $\alpha^{(i)}$ is the optimal control with regard to the mean field $\mathbf{p}^{(i)}$ and $\nu^{(i)}$, we have $\alpha_t^{(i)} = \hat{\alpha}_t^{(i)}$, $dt \otimes d\mathbb{P}$ -a.e. This implies that $\mathbb{Q}^{(i)}[X_{t-} = e_k] = \mathbb{Q}^{(\alpha^{(i)}, \mathbf{p}^{(i)}, \nu^{(i)})}[X_{t-} = e_k]$ for all $i = 1, 2$ and

$k = 1, \dots, m$. Since $(\alpha^{(i)}, \mathbf{p}^{(i)}, \nu^{(i)})$ is a Nash equilibrium, we have $\mathbb{Q}^{(\alpha^{(i)}, \mathbf{p}^{(i)}, \nu^{(i)})}[X_{t-} = e_k] = [p_t^{(i)}]_k$. Therefore we obtain $\mathbb{Q}^{(i)}[X_{t-} = e_k] = [p_t^{(i)}]_k$ for all $i = 1, 2$ and $k = 1, \dots, m$. Now using item (iii) of Assumption 5.1, we have:

$$(56) \quad \begin{aligned} 0 &\geq \sum_{i=1}^m (g(e_i, p_T^{(1)}) - g(e_i, p_T^{(2)}))([p_T^{(1)}]_i - [p_T^{(2)}]_i) \\ &\quad + \int_0^T \sum_{i=1}^m (f_0(t, e_i, p_t^{(1)}) - f_0(t, e_i, p_t^{(2)}))([p_t^{(1)}]_i - [p_t^{(2)}]_i) dt \geq 0 \end{aligned}$$

Assume that there exists a measurable subset N of $[0, T] \times \Omega$ with strictly positive $dt \otimes d\mathbb{Q}^{(1)}$ measure, such that $\hat{\alpha}_t^{(1)} \neq \hat{\alpha}_t^{(2)}$ on N . By Assumption 5.1, the mapping $\alpha \rightarrow H(t, X_{t-}, \alpha, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$ admits a unique minimizer and therefore for all $(t, w) \in N$, we have:

$$H(t, X_{t-}, \hat{\alpha}_t^{(1)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)}) > H(t, X_{t-}, \hat{\alpha}_t^{(2)}, Z_t^{(2)}, p_t^{(1)}, \nu_t^{(1)})$$

Piggybacking on the argument laid out above, we see that the first inequality is strict in (56) which leads to a contradiction. Therefore we have $\hat{\alpha}_t^{(1)} = \hat{\alpha}_t^{(2)}$, $dt \otimes d\mathbb{Q}^{(1)}$ -a.e., and $dt \otimes d\mathbb{P}$ -a.e., since $\mathbb{P}$ is equivalent to $\mathbb{Q}^{(1)}$. It follows that $\alpha_t^{(1)} = \alpha_t^{(2)}$, $dt \otimes d\mathbb{P}$ -a.e. Finally, using the same type of argument as in the proof of Proposition 4.14, we obtain $\mathbb{Q}^{(1)} = \mathbb{Q}^{(2)}$ which finally leads to $(\mathbf{p}^{(1)}, \nu^{(1)}) = (\mathbf{p}^{(2)}, \nu^{(2)})$. This completes the proof of the uniqueness. $\square$

6. APPROXIMATE NASH EQUILIBRIUM FOR GAMES WITH FINITE MANY PLAYERS

In this section we show that the solution of a mean field game can be used to construct approximate Nash equilibria for games with finitely many players. We first set the stage for the weak formulation of the game with N players in finite state spaces. Recall that Ω is the space of càdlàg mappings from $[0, T]$ to $E = \{e_1, \dots, e_M\}$ which are continuous on T , $t \rightarrow X_t$ is the canonical process and $\mathbb{F} := (\mathcal{F}_t)_{t \leq T}$ is the natural filtration generated by $\mathbf{X}$. Let us fix $p^\circ \in \mathcal{S}$ a probability distribution on the state space E . Let $\mathbb{P}$ be the probability on $(\Omega, \mathcal{F}_T)$ under which $\mathbf{X}$ is a continuous-time Markov chain with transition rate matrix Q^0 and initial distribution p° . Let Ω^N be the product space of N copies of Ω , and $\mathbb{P}^N$ be the product probability measure of N identical copies of $\mathbb{P}$. For $n = 1, \dots, N$, define the process $X_t^n(w) := w_t^n$ of which the natural filtration is denoted by $\mathbb{F}^{n,N} := (\mathcal{F}_t^{n,N})_{t \in [0, T]}$. We also denote by $\mathbb{F}^N := (\mathcal{F}_t^N)_{t \in [0, T]}$ the natural filtration generated by the process $(\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^N)$. Denote $\mathcal{M}_t^n := X_t^n - X_t^0 - \int_0^t Q^0 \cdot X_s^n ds$. It is clear that under $\mathbb{P}^N$, $\mathbf{X}^1, \dots, \mathbf{X}^N$ are N independent continuous-time Markov chains with initial distribution $\mathbf{p}^\circ$ and Q^0 as the transition rate matrix, and $\mathcal{M}^1, \dots, \mathcal{M}^N$ are independent $\mathcal{F}^N$ -martingales. For later use, for $i = 1, \dots, N$, we define the matrix ψ_t^n by $\psi_t^n := \text{diag}(Q^0 \cdot X_{t-}^n) - Q^0 \cdot \text{diag}(X_{t-}^n) - \text{diag}(X_{t-}^n) \cdot Q^0$.

Throughout this section, we let Assumptions 2.2, 2.7 and 3.4 hold. In addition, we adopt the following assumption:

Assumption 6.1. *The transition rate function q does not depend on the mean field of state, nor the mean field of control.*

We assume that each player can observe the entire past history of every player's state. We denote by $\mathbb{A}^N$ the collection of $\mathbb{F}^N$ -predictable processes taking values in A . Each player n chooses a strategy $\alpha^n \in \mathbb{A}^N$. We define the martingale $L(\alpha^1, \dots, \alpha^N)$ by:

$$(57) \quad L_t(\alpha^1, \dots, \alpha^N) := \int_0^t \sum_{n=1}^N (X_{s-}^n)^* \cdot (Q(s, \alpha_s^n) - Q^0) \cdot (\psi_s^n)^+ \cdot d\mathcal{M}_s^n,$$

and the probability measure $\mathbb{Q}(\alpha^1, \dots, \alpha^N)$ by:

$$(58) \quad \frac{d\mathbb{Q}(\alpha^1, \dots, \alpha^N)}{d\mathbb{P}^N} = \mathcal{E}_T^{\alpha^1, \dots, \alpha^N},$$

where we denote by $\mathcal{E}(\alpha^1, \dots, \alpha^N)$ the Doléans-Dade exponential of $L(\alpha^1, \dots, \alpha^N)$. Finally we introduce the empirical distribution of the states:

$$(59) \quad p_t^N := \frac{1}{N} \left[\sum_{n=1}^N \mathbb{1}(X_t^n = e_1), \sum_{n=1}^N \mathbb{1}(X_t^n = e_2), \dots, \sum_{n=1}^N \mathbb{1}(X_t^n = e_m) \right] \in \mathcal{S},$$

as well as the empirical distribution of the controls:

$$(60) \quad \nu(\alpha_t^1, \dots, \alpha_t^N) := \frac{1}{N} \sum_{n=1}^N \delta_{\alpha_t^n} \in \mathcal{P}(A),$$

where $\delta_a(\cdot)$ is the Dirac measure on a . The total expected cost of player n in the game with N players, denoted by $J^{n,N}(\alpha^1, \dots, \alpha^N)$, is defined as:

$$(61) \quad J^{n,N}(\alpha^1, \dots, \alpha^N) := \mathbb{E}^{\mathbb{Q}(\alpha^1, \dots, \alpha^N)} \left[\int_0^T f(t, X_t^n, \alpha_t^n, p_t^N, \nu(\alpha_t^1, \dots, \alpha_t^N)) dt + g(X_T^n, p_T^N) \right].$$

Now let us consider a Nash equilibrium $(\alpha^*, \mathbf{p}^*, \nu^*)$ of the mean field game in the sense of Definition 2.8. Recall that

α^* is a predictable process with respect to the natural filtration generated by the canonical process $\mathbf{X}$. For each $n = 1, \dots, N$, we may define the control $\hat{\alpha}^n$ of player n by:

$$(62) \quad \hat{\alpha}^n(w^1, \dots, w^N) := \alpha^*(w^n).$$

Clearly, $\hat{\alpha}^n$ is $\mathcal{F}^{n,N}$ -predictable. In other words, it only depends on the observation of player n 's own path. Therefore the strategy profile $\hat{\alpha}^{(N)} := (\hat{\alpha}^1, \dots, \hat{\alpha}^N)$ is a distributed strategy profile, which means that every player's strategy is only based on the observation of its own path.

In the following, we will show that $\hat{\alpha}^{(N)}$ is an approximate Nash equilibrium in a sense to be made clear later on. To this end, we first give a result on the propagation of chaos, which compare players n 's total expected cost in the mean field game versus its total expected cost in the finite player game. To simplify the notations, we use the abbreviation $(\beta, \hat{\alpha}^{-n,N})$ for $(\hat{\alpha}^1, \dots, \hat{\alpha}^{n-1}, \beta, \hat{\alpha}^{n+1}, \dots, \hat{\alpha}^N)$, $\hat{\mathbb{Q}}^{(N)}$ for $\mathbb{Q}^{\hat{\alpha}^{(N)}}$, $\hat{\mathbb{E}}^{(N)}$ for $\mathbb{E}^{\hat{\mathbb{Q}}^{(N)}}$, and finally $\hat{\mathcal{E}}^{(N)}$ for $\mathcal{E}^{\hat{\alpha}^{(N)}}$. We start from the following lemmas:

Lemma 6.2. *There exists a sequence $(\delta_N)_{N \geq 0}$ such that $\delta_N \rightarrow 0$ as $N \rightarrow +\infty$, and such that for all $N \geq 1$, $n \leq N$ and $t \leq T$ we have:*

$$(63) \quad \max \{ \hat{\mathbb{E}}^{(N)} [\mathcal{W}_1^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)], \hat{\mathbb{E}}^N [\|p_t^N - p_t^*\|^2] \} \leq \delta_N.$$

Proof. Since $\hat{\mathbb{Q}}^{(N)} = \mathbb{Q}^{\hat{\alpha}^{(N)}}$ and the fact that $(\alpha^*, \mathbf{p}^*, \nu^*)$ is an equilibrium of the mean field game, we deduce that under the measure $\hat{\mathbb{Q}}^{(N)}$, the states $X_t^1, \dots, X_t^N$ are independent and have the same distribution characterized by p_t^* , and that the controls $\alpha_t^1, \dots, \alpha_t^N$ are independent and have the same distribution ν_t^* . Therefore, for $i \in \{1, \dots, M\}$, we have:

$$\hat{\mathbb{E}}^{(N)} \left[\left(\frac{1}{N} \sum_{n=1}^N \mathbb{1}(X_t^n = e_i) - \hat{\mathbb{Q}}^N[X_t^1 = e_i] \right)^2 \right] = \frac{1}{N} (\hat{\mathbb{Q}}^{(N)}[X_t^1 = e_i] - (\hat{\mathbb{Q}}^{(N)}[X_t^1 = e_i])^2) \leq \frac{1}{4N},$$

which leads to:

$$\hat{\mathbb{E}}^{(N)} [\|p_t^N - p_t^*\|^2] = \sum_{i=1}^M \hat{\mathbb{E}}^{(N)} \left[\left(\frac{1}{N} \sum_{n=1}^N \mathbb{1}(X_t^n = e_i) - \hat{\mathbb{Q}}^N[X_t^1 = e_i] \right)^2 \right] \leq \frac{M}{4N}.$$

On the other hand, $\nu(\beta_t, \hat{\alpha}_t^{-1,N})$ and ν_t^* are in $\mathcal{P}(A)$ with A being a compact subset of $\mathbb{R}^d$. We have:

$$\hat{\mathbb{E}}^{(N)} [\mathcal{W}_1^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)] \leq C \hat{\mathbb{E}}^{(N)} [\mathcal{W}_1(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)]$$

$$\begin{aligned}
&\leq C \hat{\mathbb{E}}^{(N)}[\mathcal{W}_1(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu(\hat{\alpha}_t^{(N)})) + \mathcal{W}_1(\nu(\hat{\alpha}_t^{(N)}), \nu_t^*)] \\
&\leq C(\hat{\mathbb{E}}^{(N)}[\frac{1}{N} \|\beta_t - \hat{\alpha}_t^{1,N}\|] + \hat{\mathbb{E}}^{(N)}[\mathcal{W}_1(\nu(\hat{\alpha}_t^{(N)}), \nu_t^*)]) \\
&\leq C(\frac{1}{N} + \hat{\mathbb{E}}^{(N)}[\mathcal{W}_1(\nu(\hat{\alpha}_t^{(N)}), \nu_t^*)]),
\end{aligned}$$

where C is a constant only depending on $\sup_{a \in A} \|a\|$ which changes its value from line to line. Now applying Theorem 1 in [12], we have:

$$\hat{\mathbb{E}}^{(N)}[\mathcal{W}_1(\nu(\hat{\alpha}_t^{(N)}), \nu_t^*)] \leq \sup_{a \in A} \|a\| \cdot [\mathbb{1}(d \leq 2)(N^{-1/2} \log(1+N) + N^{-2/3}) + \mathbb{1}(d > 2)(N^{-1/d} + N^{-1/2})].$$

Combining with the estimates previously shown, we obtain the desired result. $\square$

Lemma 6.3. *There exists a constant C which only depends on the bound of the transition rate q , such that for all $N > 0$ and $\beta \in \mathbb{A}$ we have:*

$$(64) \quad \hat{\mathbb{E}}^{(N)} \left[\left(\frac{\mathcal{E}_T^{(\beta, \hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_T^{(N)}} \right)^2 \right] \leq C.$$

Proof. Let us denote $W_t := \mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})} / \hat{\mathcal{E}}_t^{(N)}$. By Ito's formula we have:

$$\begin{aligned}
dW_t &= \frac{d\mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})} - \Delta \mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_{t-}^{(N)}} - \frac{\mathcal{E}_{t-}^{(\beta, \hat{\alpha}^{-1,N})} (d\hat{\mathcal{E}}_t^{(N)} - \Delta \hat{\mathcal{E}}_t^{(N)})}{(\hat{\mathcal{E}}_{t-}^{(N)})^2} + \Delta W_t \\
&= W_{t-} \left(\frac{d\mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})} - \Delta \mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})}}{\mathcal{E}_{t-}^{(\beta, \hat{\alpha}^{-1,N})}} - \frac{d\hat{\mathcal{E}}_t^{(N)} - \Delta \hat{\mathcal{E}}_t^{(N)}}{\hat{\mathcal{E}}_{t-}^{(N)}} \right) + \Delta W_t
\end{aligned}$$

Recall that:

$$\begin{aligned}
\frac{d\hat{\mathcal{E}}_t^{(N)}}{\hat{\mathcal{E}}_{t-}^{(N)}} &= \sum_{n=1}^N (X_{t-}^n)^* \cdot (Q(t, \hat{\alpha}_t^n) - Q^0) \cdot (\psi_t^n)^+ \cdot d\mathcal{M}_t^n, \\
\frac{d\mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})}}{\mathcal{E}_{t-}^{(\beta, \hat{\alpha}^{-1,N})}} &= (X_{t-}^1)^* \cdot (Q(t, \beta_t) - Q^0) \cdot (\psi_t^1)^+ \cdot d\mathcal{M}_t^1 + \sum_{n=2}^N (X_{t-}^n)^* \cdot (Q(t, \hat{\alpha}_t^n) - Q^0) \cdot (\psi_t^n)^+ \cdot d\mathcal{M}_t^n,
\end{aligned}$$

and $d\mathcal{M}_t^n = \Delta \mathcal{M}_t^n - Q^0 X_{t-}^n dt$. Noticing that for $n \neq 1$, the jumps of $\mathcal{M}_t^n$ do not result in the jumps of W_t , we obtain:

$$\begin{aligned}
\Delta W_t &= \Delta \left(\frac{\mathcal{E}_t^{(\beta, \hat{\alpha}^{-1,N})}}{\mathcal{E}_t^{\hat{\alpha}}} \right) = \frac{\mathcal{E}_{t-}^{(\beta, \hat{\alpha}^{-1,N})}}{\mathcal{E}_{t-}^{\hat{\alpha}}} \cdot \left(\frac{1 + (X_{t-}^1)^* \cdot (Q(t, \beta_t) - Q^0) \cdot (\psi_t^1)^+ \cdot \Delta \mathcal{M}_t^1}{1 + (X_{t-}^1)^* \cdot (Q(t, \hat{\alpha}_t^1) - Q^0) \cdot (\psi_t^1)^+ \cdot \Delta \mathcal{M}_t^1} - 1 \right) \\
&= W_{t-} \frac{(X_{t-}^1)^* \cdot (Q(t, \beta_t) - Q(t, \hat{\alpha}_t^1)) \cdot (\psi_t^1)^+ \cdot \Delta \mathcal{M}_t^1}{1 + (X_{t-}^1)^* \cdot (Q(t, \hat{\alpha}_t^1) - Q^0) \cdot (\psi_t^1)^+ \cdot \Delta \mathcal{M}_t^1}.
\end{aligned}$$

Piggybacking on the computation in equation (6), we see that when $X_{t-}^1 = e_i \neq e_j = X_t$, we have $\Delta \mathcal{M}_t^1 = \Delta X_t^1 = e_j - e_i$ and:

$$\Delta W_t = W_{t-} \frac{q(t, i, j, \beta_t) - q(t, i, j, \hat{\alpha}_t^1)}{q(t, i, j, \hat{\alpha}_t^1)}.$$

Let us define Ξ_t^β to be an m by m matrix where the diagonal elements are 0 and the element on the i -th row and the j -th column is $\frac{q(t, i, j, \beta_t) - q(t, i, j, \hat{\alpha}_t^1)}{q(t, i, j, \hat{\alpha}_t^1)}$. Then it is clear that $\Delta W_t = e_i^* \cdot \Xi_t^\beta \cdot (e_j - e_i)$. It follows that:

$$\Delta W_t = W_{t-} \cdot (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot \Delta \mathcal{M}_t^1.$$

Injecting the above equation into the Itô decomposition of W_t , we obtain:

$$\begin{aligned} dW_t &= W_{t-}[(Q(t, \hat{\alpha}_t^1) - Q(t, \beta_t)) \cdot (\psi_t^1)^+ \cdot Q^0 \cdot X_{t-}^1 dt + (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot \Delta \mathcal{M}_t^1] \\ &= W_{t-}[(Q(t, \hat{\alpha}_t^1) - Q(t, \beta_t)) \cdot (\psi_t^1)^+ \cdot Q^0 \cdot X_{t-}^1 dt + (X_{t-}^1)^* \cdot \Xi_t^\beta \cdot (d\hat{\mathcal{M}}_t^1 + Q^*(t, \hat{\alpha}_t^1) \cdot X_{t-}^1 dt)]. \end{aligned}$$

In the second equality, we use the fact that under the measure $\hat{\mathbb{Q}}^{(N)}$, the state process X_t^1 has the canonical decomposition $dX_t^1 = d\hat{\mathcal{M}}_t^1 + Q^*(t, \hat{\alpha}_t^1) \cdot X_t^1 dt$ where $\hat{\mathcal{M}}^1$ is a $\hat{\mathbb{Q}}^{(N)}$ -martingale. We also use the equality $\Delta \mathcal{M}_t^1 = \Delta X_t^1 = dX_t^1$. In addition, by replacing X_t^1 with e_i for $i = 1, \dots, M$, it is plain to check the following equality:

$$(Q(t, \hat{\alpha}_t^1) - Q(t, \beta_t)) \cdot (\psi_t^1)^+ \cdot Q^0 \cdot X_t^1 + (X_t^1)^* \cdot \Xi_t^\beta \cdot Q^*(t, \hat{\alpha}_t^1) \cdot X_t^1 = 0.$$

This leads to the following representation of W_t :

$$dW_t = W_{t-} (X_t^1)^* \cdot \Xi_t^\beta \cdot d\hat{\mathcal{M}}_t^1,$$

which is a local martingale under the measure $\hat{\mathbb{Q}}^{(N)}$. At this stage, the rest of the proof is exactly the same as the proof of Proposition 4.10. In particular, we make use of Assumption 2.2, that is the transition rate q being bounded uniformly with regard to the controls. $\square$

We are now ready to prove the form of the propagation of chaos result which we need.

Proposition 6.4. *There exists a sequence $(\epsilon_N)_{N \geq 0}$ such that $\epsilon_N \rightarrow 0$ as $N \rightarrow +\infty$ and such that for all $N \geq 0$, $n \leq N$ and $\beta \in \mathbb{A}$:*

$$(65) \quad \left| J^{n,N}(\beta, \hat{\alpha}^{-n,N}) - \mathbb{E}^{\mathbb{Q}^{(\beta, \hat{\alpha}^{-n,N})}} \left[\int_0^T f(t, X_t^n, \beta_t, p_t^*, \nu_t^*) dt + g(X_T^n, p_T^*) \right] \right| \leq \epsilon_N.$$

Proof. Due to symmetry, we only need to show the claim for $n = 1$. Let $N > 0$ and $\beta \in \mathbb{A}$. Using successively Cauchy-Schwartz inequality, Assumption 2.7, Lemma 6.2 and Lemma 6.3, we have:

$$\begin{aligned} & \left| J^{n,N}(\beta, \hat{\alpha}^{-1,N}) - \mathbb{E}^{\mathbb{Q}^{(\beta, \hat{\alpha}^{-1,N})}} \left[\int_0^T f(t, X_t^1, \beta_t, p_t^*, \nu_t^*) dt + g(X_T^1, p_T^*) \right] \right| \\ & \leq \mathbb{E}^{\mathbb{Q}^{(\beta, \hat{\alpha}^{-1,N})}} \left[\int_0^T |f(t, X_t^1, \beta_t, p_t^*, \nu_t^*) - f(t, X_t^1, \beta_t, p_t^N, \nu(\beta_t, \hat{\alpha}_t^{-1,N}))| dt + |g(X_T^1, p_T^*) - g(X_T^1, p_T^N)| \right] \\ & = \hat{\mathbb{E}}^{(N)} \left[\frac{\mathcal{E}_T^{(\beta, \hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_T^{(N)}} \int_0^T |f(t, X_t^1, \beta_t, p_t^*, \nu_t^*) - f(t, X_t^1, \beta_t, p_t^N, \nu(\beta_t, \hat{\alpha}_t^{-1,N}))| dt + |g(X_T^1, p_T^*) - g(X_T^1, p_T^N)| \right] \\ & \leq \hat{\mathbb{E}}^{(N)} \left[\left(\frac{\mathcal{E}_T^{(\beta, \hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_T^{(N)}} \right)^2 \right]^{1/2} \left[\hat{\mathbb{E}}^{(N)} \left[\left(\int_0^T |f(t, X_t^1, \beta_t, p_t^*, \nu_t^*) - f(t, X_t^1, \beta_t, p_t^N, \nu(\beta_t, \hat{\alpha}_t^{-1,N}))| dt + |g(X_T^1, p_T^*) - g(X_T^1, p_T^N)| \right)^2 \right] \right]^{1/2} \\ & \leq C \hat{\mathbb{E}}^{(N)} \left[\left(\frac{\mathcal{E}_T^{(\beta, \hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_T^{(N)}} \right)^2 \right]^{1/2} \left[\int_0^T (\hat{\mathbb{E}}^{(N)}[\|p_t^N - p_t^*\|^2] + \hat{\mathbb{E}}^{(N)}[\mathcal{W}_1^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)]) dt + \hat{\mathbb{E}}^{(N)}[\|p_T^N - p_T^*\|^2] \right]^{1/2} \\ & \leq C \sqrt{\delta_N}, \end{aligned}$$

where δ_N is as appeared in Lemma 6.2, and C is a constant only depending on T , the Lipschitz constant of f and g and the constant appearing in Lemma 6.3. this gives us the desired inequality. $\square$

As a direct consequence of the above result on the propagation of chaos, we show that the Nash equilibrium of the mean field game consists of an approximate Nash equilibrium for the game with finite many players.

Theorem 6.5. *There exists a sequence ϵ_N converging to 0 such that for all $N > 0$, $\beta \in \mathbb{A}$ and $n \leq N$, we have:*

$$J^{n,N}(\beta, \hat{\alpha}^{-n,N}) \leq J^{n,N}(\hat{\alpha}^{(N)}) + \epsilon_N.$$

Proof. Recall that the strategy profile is $\hat{\alpha}^{(N)} = (\hat{\alpha}^1, \dots, \hat{\alpha}^N)$ is defined as:

$$\hat{\alpha}^n(w^1, w^2, \dots, w^N) := \alpha^*(w^n),$$

where α^* is the strategy of the mean field game equilibrium, together with $\mathbf{p}^*$ as the mean field of states and ν^* as the mean field of control. For a strategy profile $(\alpha^1, \dots, \alpha^N)$ we use the notation:

$$K^{n,N}(\alpha^1, \dots, \alpha^N) := \mathbb{E}^{\mathbb{Q}(\alpha^1, \dots, \alpha^N)} \left[\int_0^T f(t, X_t^n, \alpha_t^n, p_t^*, \nu_t^*) dt + g(X_T^n, p_T^*) \right].$$

Now taking $n = 1$, we observe that $K^{1,N}(\hat{\alpha}^{(N)}) = \mathbb{E}^{\mathbb{P}^N} [Y_0^{(\hat{\alpha}^{(N)})}]$, where $Y_0^{(\hat{\alpha}^{(N)})}$ is the solution (at time $t = 0$) of the following BSDE:

$$(66) \quad Y_t = g(X_T^1, p_T^*) + \int_t^T H(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*) ds - \int_t^T (Z_s^1)^* \cdot d\mathcal{M}_s^1.$$

By the optimality of the equilibrium, we know that for all $t \in [0, T]$, $\hat{\alpha}_t^1$ minimizes the mapping $\alpha \rightarrow H(t, X_{t-}^1, Z_t^1, \alpha, p_t^*, \nu_t^*)$. Clearly, the solution of the above BSDE (66) is also the unique solution to the following BSDE:

$$(67) \quad Y_t = g(X_T^1, p_T^*) + \int_t^T \left[H(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*) + \sum_{n=2}^N (X_t^n)^* \cdot (Q(s, \hat{\alpha}_s^n) - Q^0) \cdot Z_s^n \right] ds - \int_t^T \sum_{n=1}^N \int_t^T (Z_s^n)^* \cdot d\mathcal{M}_s^n,$$

with $Z_t^n = 0$ for $n = 2, \dots, N$. Indeed, the existence and uniqueness of the BSDE (67) can be checked easily by applying Theorem 6.7. On the other hand, by following exactly the same argument as in the proof of Lemma 3.8, we can show that $K^{1,N}(\beta, \hat{\alpha}^{-1,N}) = \mathbb{E}^{\mathbb{P}^N} [Y_0^{(\beta, \hat{\alpha}^{-1,N})}]$, where $Y_0^{(\beta, \hat{\alpha}^{-1,N})}$ is the solution (at time $t = 0$) of:

$$(68) \quad Y_t = g(X_T^1, p_T^*) + \int_t^T \left[H(s, X_{s-}^1, Z_s^1, \beta_s, p_s^*, \nu_s^*) + \sum_{n=2}^N (X_t^n)^* \cdot (Q(s, \hat{\alpha}_s^n) - Q^0) \cdot Z_s^n \right] ds - \int_t^T \sum_{n=1}^N \int_t^T (Z_s^n)^* \cdot d\mathcal{M}_s^n.$$

Notice that $H(s, X_{s-}^1, Z_s^1, \alpha, p_s^*, \nu_s^*) = f(s, X_{s-}^1, \alpha, p_s^*, \nu_s^*) + (X_{s-}^1)^* \cdot (Q(s, \alpha) - Q^0) \cdot Z_s^1$, and $H(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*) \geq H(s, X_{s-}^1, Z_s^1, \beta_s, p_s^*, \nu_s^*)$. Applying the comparison principle as stated in Theorem 6.8 to the BSDEs (67) and (68), we conclude that $K^{1,N}(\beta, \hat{\alpha}^{-1,N}) \leq K^{1,N}(\hat{\alpha}^{(N)})$ for all $\beta \in \mathbb{A}$. Now thanks to symmetry, we have $K^{n,N}(\beta, \hat{\alpha}^{-n,N}) \leq K^{1,N}(\hat{\alpha}^{(N)})$ for all $\beta \in \mathbb{A}$ and $n = 1, \dots, N$. The desired results immediately follows by applying Proposition 6.4. $\square$

APPENDIX: BSDEs DRIVEN BY MULTIPLE INDEPENDENT CONTINUOUS-TIME MARKOV CHAINS

Let us consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ supporting N independent continuous-time Markov chains $\mathbf{X}^1, \dots, \mathbf{X}^N$. For each $n = 1, \dots, N$, we assume that $\mathbf{X}^n$ takes only m_n states, which are represented by the basis vectors of the space $\mathbb{R}^{m_n}$. We assume that under $\mathbb{P}$, the transition rate matrix of $\mathbf{X}^n$ is $Q^{0,n}$, which is an $m_n \times m_n$ matrix where all the diagonal elements equal $-(m_n - 1)$ and all the off-diagonal elements equal 1. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$ the natural filtration generated by $(\mathbf{X}^1, \dots, \mathbf{X}^N)$. It is clear that for each n , we can decompose the Markov chain $\mathbf{X}^n$ as $X_t^n = X_0^n + \int_0^t Q^{0,n} \cdot X_{s-}^n ds + d\mathcal{M}_t^n$, where $\mathcal{M}^n$ is an $\mathbb{F}$ -martingale. In addition, due to the independence of the Markov chains, for all $n_1 \neq n_2$ and $t \leq T$, $\mathbb{P}$ -almost surely we have $\Delta X_t^{n_1} = 0$ or $\Delta X_t^{n_2} = 0$. In other words, any two Markov chains cannot jump simultaneously.

Let us consider the process $\tilde{\mathbf{X}}$ defined by $\tilde{X}_t := X_t^1 \otimes X_t^2 \otimes \dots \otimes X_t^N$ where $\otimes$ stands for the Kronecker product. Indeed, $\tilde{\mathbf{X}}$ is a Markov chain encoding the joint states of the N independent Markov chains, and $\tilde{\mathbf{X}}$ only takes values among the unit vectors of the space $\mathbb{R}^{m_1 \times \dots \times m_N}$. We have the following result on the decomposition of $\tilde{\mathbf{X}}$.

Lemma 6.6. $\tilde{\mathbf{X}}$ is a continuous-time Markov chain with transition rate matrix $\tilde{Q}^0$ given by:

$$(69) \quad \tilde{Q}^0 := \sum_{n=1}^N I_{m_1} \otimes \cdots \otimes I_{m_{n-1}} \otimes Q^{0,n} \otimes I_{m_{n+1}} \otimes \cdots \otimes I_{m_N}.$$

In addition it has the canonical decomposition:

$$(70) \quad d\tilde{X}_t = \tilde{Q}^0 \cdot \tilde{X}_{t-} dt + d\tilde{\mathcal{M}}_t,$$

where $\tilde{\mathcal{M}}$ is a $\mathbb{F}$ -martingale which satisfies:

$$(71) \quad d\tilde{\mathcal{M}}_t = \sum_{n=1}^N (X_{t-}^1 \otimes \cdots \otimes X_{t-}^{n-1} \otimes I_{m_n} \otimes X_{t-}^{n+1} \otimes \cdots \otimes X_{t-}^N) \cdot d\mathcal{M}_t^n.$$

Proof. In order to keep the notation to a reasonable level of complexity, we only argue the proof for $N = 2$. Applying Itô's formula to $X_t^1 \otimes X_t^2$ and noticing that X_t^1 and X_t^2 have no simultaneous jumps, we obtain:

$$\begin{aligned} d(X_t^1 \otimes X_t^2) &= dX_t^1 \otimes X_{t-}^2 + X_{t-}^1 \otimes dX_t^2 \\ &= (Q^{0,1} \cdot X_{t-}^1) \otimes X_{t-}^2 dt + d\mathcal{M}_t^1 \otimes X_{t-}^2 + X_{t-}^1 \otimes (Q^{0,2} \cdot X_{t-}^2) dt + X_{t-}^1 \otimes d\mathcal{M}_t^2. \end{aligned}$$

Using the properties of the Kronecker product, we have:

$$\begin{aligned} (Q^{0,1} \cdot X_{t-}^1) \otimes X_{t-}^2 &= (Q^{0,1} \cdot X_{t-}^1) \otimes (I_{m_2} \cdot X_{t-}^2) = (Q^{0,1} \otimes I_{m_2}) \cdot (X_{t-}^1 \otimes X_{t-}^2) \\ d\mathcal{M}_t^1 \otimes X_{t-}^2 &= (I_{m_1} \cdot d\mathcal{M}_t^1) \otimes (X_{t-}^2 \cdot 1) \\ &= (I_{m_1} \otimes X_{t-}^2) \cdot (d\mathcal{M}_t^1 \otimes 1) = (I_{m_1} \otimes X_{t-}^2) \cdot d\mathcal{M}_t^1 \\ X_{t-}^1 \otimes (Q^{0,2} \cdot X_{t-}^2) &= (I_{m_1} \cdot X_{t-}^1) \otimes (Q^{0,2} \cdot X_{t-}^2) = (I_{m_1} \otimes Q^{0,2}) \cdot (X_{t-}^1 \otimes X_{t-}^2) \\ X_{t-}^1 \otimes d\mathcal{M}_t^2 &= (X_{t-}^1 \cdot 1) \otimes (I_{m_2} \cdot d\mathcal{M}_t^2) \\ &= (X_{t-}^1 \otimes I_{m_2}) \cdot (1 \otimes d\mathcal{M}_t^2) = (X_{t-}^1 \otimes I_{m_2}) \cdot d\mathcal{M}_t^2. \end{aligned}$$

Plugging the above equalities into the Itô decomposition yields the desired result for $N = 2$. The case $N > 2$ can be treated by applying a simple argument of induction, which we will not detail here. $\square$

As in the case of a single Markov chain, we define the stochastic matrix $\psi_t^n := \text{diag}(Q^{0,n} \cdot X_{t-}^n) - Q^{0,n} \cdot \text{diag}(X_{t-}^n) - \text{diag}(X_{t-}^n) \cdot Q^{0,n}$ for $n = 1, \dots, N$ as well as $\tilde{\psi}_t := \text{diag}(\tilde{Q}^0 \cdot \tilde{X}_{t-}) - \tilde{Q}^0 \cdot \text{diag}(\tilde{X}_{t-}) - \text{diag}(\tilde{X}_{t-}) \cdot \tilde{Q}^0$. For $n = 1, \dots, N$, we define the stochastic seminorm $\|\cdot\|_{X_{t-}^n}$ by $\|Z\|_{X_{t-}^n}^2 := Z^* \cdot \psi_t^n \cdot Z$ where $Z \in \mathbb{R}^{m_n}$. We then define the stochastic seminorm $\|\cdot\|_{\tilde{X}_{t-}}$ by $\|\tilde{Z}\|_{\tilde{X}_{t-}}^2 := \tilde{Z}^* \cdot \tilde{\psi}_t \cdot \tilde{Z}$ where $\tilde{Z} \in \mathbb{R}^{m_1 \times \cdots \times m_N}$. Our objective is to show existence and uniqueness of the following BSDE:

$$(72) \quad Y_t = \xi + \int_t^T F(w, s, Y_s, Z_s^1, \dots, Z_s^N) ds - \sum_{n=1}^N \int_t^T (Z_s^n)^* \cdot d\mathcal{M}_s^n.$$

Here ξ is a $\mathcal{F}_T$ -measurable $\mathbb{P}$ -square integrable random variable and the driver $F : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N} \rightarrow \mathbb{R}$ is a function such that the process $t \rightarrow F(w, t, y, z_1, \dots, z_N)$ is predictable for all $y, z_1, \dots, z_N \in \mathbb{R} \times \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$. The unknowns of the equation are a càdlàg process $\mathbf{Y}$ taking values in $\mathbb{R}$ and predictable processes $\mathbf{Z}^1, \dots, \mathbf{Z}^N$ taking values in $\mathbb{R}^{m_1}, \dots, \mathbb{R}^{m_N}$ respectively.

Theorem 6.7. Assume that there exists a constant $C > 0$ such that $dt \times \mathbb{P}$ -a.s., we have:

$$(73) \quad |F(w, t, y, z_1, \dots, z_N) - F(w, t, \tilde{y}, \tilde{z}_1, \dots, \tilde{z}_N)| \leq C \left(|y - \tilde{y}| + \sum_{n=1}^N \|z_n - \tilde{z}_n\|_{X_{t-}^n} \right).$$

Then the BSDE (72) admits a solution $(\mathbf{Y}, \mathbf{Z}^1, \dots, \mathbf{Z}^N)$ satisfying:

$$\mathbb{E} \left[\int_0^T |Y_t|^2 dt \right] < +\infty, \quad \mathbb{E} \left[\sum_{n=1}^N \int_0^T \|Z_t^n\|_{X_{t-}^n}^2 dt \right] < +\infty$$

Moreover, the solution is unique in the sense that if $(\mathbf{Y}^{(1)}, \mathbf{Z}^{(1),1}, \dots, \mathbf{Z}^{(1),N})$ and $(\mathbf{Y}^{(2)}, \mathbf{Z}^{(2),1}, \dots, \mathbf{Z}^{(2),N})$ are two solutions, then $\mathbf{Y}^{(1)}$ and $\mathbf{Y}^{(2)}$ are indistinguishable and we have $\mathbb{E}[\int_0^T \|\tilde{Z}_t^{(1)} - \tilde{Z}_t^{(2)}\|_{\tilde{X}_{t-}}^2 dt] = 0$.

Proof. For simplicity of the presentation, we give the proof for $N = 2$. It can be easily generalized to any $N > 2$. Our first step is to show that the following equality holds for all $\tilde{Z} \in \mathbb{R}^{m_1 \times m_2}$:

$$(74) \quad \|\tilde{Z}\|_{\tilde{X}_{t-}}^2 = \|(I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z}\|_{X_{t-}^1}^2 + \|((X_{t-}^1)^* \otimes I_{m_2}) \cdot \tilde{Z}\|_{X_{t-}^2}^2$$

By the definition of the semi-norm $\|\cdot\|_{X_{t-}^1}$, we have:

$$\begin{aligned} & \|(I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z}\|_{X_{t-}^1}^2 = \mathbf{Z}^* \cdot (I_{m_1} \otimes X_{t-}^2) \cdot \psi_t^1 \cdot (I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z} \\ &= \mathbf{Z}^* \cdot (I_{m_1} \otimes X_{t-}^2) \cdot (\psi_t^1 \otimes 1) \cdot (I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z} = \mathbf{Z}^* \cdot (\psi_t^1 \otimes X_{t-}^2) \cdot (I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z} \\ &= \mathbf{Z}^* \cdot [\psi_t^1 \otimes (X_{t-}^2 \cdot (X_{t-}^2)^*)] \cdot \tilde{Z} = \mathbf{Z}^* \cdot (\psi_t^1 \otimes \text{diag}(X_{t-}^2)) \cdot \tilde{Z}. \end{aligned}$$

Similarly we have $\|((X_{t-}^1)^* \otimes I_{m_2}) \cdot \tilde{Z}\|_{X_{t-}^2}^2 = \mathbf{Z}^* \cdot (\text{diag}(X_{t-}^1) \otimes \psi_t^2) \cdot \tilde{Z}$. Now by the definition of $\tilde{\psi}_t$, we have:

$$\begin{aligned} \tilde{\psi}_t &= \text{diag}(\tilde{Q}^0 \cdot \tilde{X}_{t-}) - \tilde{Q}^0 \cdot \text{diag}(\tilde{X}_{t-}) - \text{diag}(\tilde{X}_{t-}) \cdot \tilde{Q}^0 \\ &= \text{diag}((I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2}) \cdot (X_{t-}^1 \otimes X_{t-}^2)) - (I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2}) \cdot \text{diag}(X_{t-}^1 \otimes X_{t-}^2) \\ &\quad - \text{diag}(X_{t-}^1 \otimes X_{t-}^2) \cdot (I_{m_1} \otimes Q^{0,2} + Q^{0,1} \otimes I_{m_2}) \\ &= \text{diag}(X_{t-}^1 \otimes (Q^{0,2} \cdot X_{t-}^2)) + \text{diag}((Q^{0,1} \cdot X_{t-}^1) \otimes X_{t-}^2) \\ &\quad - \text{diag}(X_{t-}^1) \otimes (Q^{0,2} \cdot \text{diag}(X_{t-}^2)) - (Q^{0,1} \cdot \text{diag}(X_{t-}^1)) \otimes \text{diag}(X_{t-}^2) \\ &\quad - \text{diag}(X_{t-}^1) \otimes (\text{diag}(X_{t-}^2) \cdot Q^{0,2}) - (\text{diag}(X_{t-}^1) \cdot Q^{0,1}) \otimes \text{diag}(X_{t-}^2) \\ &= \psi_t^1 \otimes \text{diag}(X_{t-}^2) + \text{diag}(X_{t-}^1) \otimes \psi_t^2, \end{aligned}$$

where we have used the fact that for any two vectors X^1, X^2 we have $\text{diag}(X^1 \otimes X^2) = \text{diag}(X^1) \otimes \text{diag}(X^2)$. This immediately leads to the equality (74). Now we consider the BSDE driven by the continuous-time Markov chain $\tilde{X}$ with terminal condition ξ and the driver function $\tilde{F}$ defined by:

$$\tilde{F}(w, t, Y, \tilde{Z}) := F(w, t, Y, (I_{m_1} \otimes (X_{t-}^2)^*) \cdot \tilde{Z}, ((X_{t-}^1)^* \otimes I_{m_2}) \cdot \tilde{Z}).$$

By equality (74) and the assumption on the regularity of F , we have:

$$\begin{aligned} & |\tilde{F}(w, t, Y_1, \tilde{Z}_1) - \tilde{F}(w, t, Y_2, \tilde{Z}_2)| \\ & \leq C(|Y_1 - Y_2| + \|(I_{m_1} \otimes (X_{t-}^2)^*) \cdot (\tilde{Z}_1 - \tilde{Z}_2)\|_{X_{t-}^1} + \|((X_{t-}^1)^* \otimes I_{m_2}) \cdot (\tilde{Z}_1 - \tilde{Z}_2)\|_{X_{t-}^2}) \\ & \leq C(|Y_1 - Y_2| + \sqrt{2}\|\tilde{Z}_1 - \tilde{Z}_2\|_{\tilde{X}_{t-}}). \end{aligned}$$

Applying Lemma 3.1 we obtain the existence of the solution to the BSDE:

$$Y_t = \xi + \int_t^T \tilde{F}(s, Y_s, \tilde{Z}_s) ds + \int_t^T \tilde{Z}_s^* \cdot d\tilde{\mathcal{M}}_s$$

Now we set $Z_t^1 := (I_{m_1} \otimes (X_{s-}^2)^*) \cdot \tilde{Z}_s$ and $Z_t^2 := ((X_{s-}^1)^* \otimes I_{m_2}) \cdot \tilde{Z}_s$. From the definition of the driver $\tilde{F}$ and $\tilde{\mathcal{M}}$ in equation (71), we see that:

$$Y_t = \xi + \int_t^T F(w, s, Y_s, (I_{m_1} \otimes (X_{s-}^2)^*) \cdot \tilde{Z}_s, ((X_{s-}^1)^* \otimes I_{m_2}) \cdot \tilde{Z}_s) ds$$

$$\begin{aligned}
& + \int_t^T \tilde{Z}_s^* \cdot [(I_{m_1} \otimes X_{t-}^2) \cdot dM_t^1 + (X_{t-}^1 \otimes I_{m_2}) \cdot dM_t^2] \\
& = \xi + \int_t^T F(w, s, Y_s, Z_s^1, Z_s^2) ds + \int_t^T (Z_s^1)^* \cdot dM_t^1 + \int_t^T (Z_s^2)^* \cdot dM_t^2
\end{aligned}$$

This shows that (Y, Z^1, Z^2) is a solution to BSDE (72). $\square$

We also state a comparison principle for linear BSDEs driven by multiple independent Markov chains.

Theorem 6.8. *For each $n \in \{1, \dots, N\}$, let γ^n be a bounded predictable process in $\mathbb{R}^{m_n}$ such that $\sum_{i=1}^{m_n} [\gamma_t^n]_i = 0$ for all $t \in [0, T]$, and β a bounded predictable process in $\mathbb{R}$. Let ϕ be a non-negative predictable process in $\mathbb{R}$ such that $\mathbb{E}[\int_0^T \|\phi_t\|^2 dt] < +\infty$ and ξ a non-negative square-integrable $\mathcal{F}_T$ measurable random variable in $\mathbb{R}$. Let (Y, Z) be the solution of the linear BSDE:*

$$(75) \quad Y_t = \xi + \int_t^T (\phi_s + \beta_s Y_s + \sum_{n=1}^N (\gamma_s^n)^* \cdot Z_s^n) ds - \sum_{n=1}^N \int_t^T (Z_s^n)^* \cdot d\mathcal{M}_s^n.$$

Assume that for all $n = 1, \dots, N$, $t \in (0, T]$ and j such that $(e_j^n)^* \cdot Q^{0,n} \cdot X_{t-}^n > 0$, we have $1 + (\gamma_t^n)^* \cdot (\psi_t^n)^+ \cdot (e_j^n - X_{t-}^n) \geq 0$ where $(\psi_t^n)^+$ is the Moore-Penrose inverse of the matrix ψ_t^n . Then Y is nonnegative.

Proof. As before we treat the case for $N = 2$, for which the argument can be trivially generalized to any $N > 2$. Since γ^n and β are bounded processes and $\sum_{i=1}^{m_n} [\gamma_t^n]_i = 0$ for all $t \leq T$ and $n \leq N$, we easily verify that the Lipschitz condition (73) stated in Theorem 6.7 is satisfied and therefore the BSDE (76) admits a unique solution. Now consider the following BSDE driven by $\mathcal{M}$:

$$(76) \quad Y_t = \xi + \int_t^T (\phi_s + \beta_s Y_s + \gamma_s^* \cdot Z_s) ds - \sum_{n=1}^2 \int_t^T Z_s^* \cdot d\mathcal{M}_s,$$

where $\gamma_t := (\gamma_t^1 \otimes X_{t-}^2) + (X_{t-}^1 \otimes \gamma_t^2)$. It is easy to verify the BSDE (76) admits a unique solution (Y, Z) and following the same argument as in the proof of Theorem 6.7, we verify that $(Y_t, Z_t^1, Z_t^2) := (Y_t, (I_{m_1} \otimes (X_{s-}^2)^*) \cdot Z_s, ((X_{s-}^1)^* \otimes I_{m_2}) \cdot Z_s)$ solves the BSDE (18), which is also its unique solution. Therefore we only need to show that the solution Y to BSDE (18) is nonnegative. To this ends, we need to apply the comparison principal for the case of a single Markov chain, as is stated in Lemma 3.3. Note that X^1 and X^2 do not jump simultaneously and $X_t = X_t^1 \otimes X_t^2$. For the jump of X resulting from the jump of X^1 , we need to show that for $k = 1, \dots, m_1$:

$$(77) \quad 1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_k^1 \otimes X_{t-}^2 - X_{t-}^1 \otimes X_{t-}^2) \geq 0.$$

Let us assume that $X_{t-}^1 = e_i^1$, $X_{t-}^2 = e_j^2$. If $k = i$, the above equality is trivial. In the following, we consider the case $k \neq i$. Then by the assumption of the theorem, we have:

$$(78) \quad 1 + (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_i^1) \geq 0.$$

It can be easily verified that:

$$\begin{aligned}
& (diag(e_i^1) \otimes \psi_t^2 + \psi_t^1 \otimes diag(e_j^2)) \cdot \left[(m_1 + m_2 - 2)e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right] \\
& = e_k^1 \otimes e_j^2 - e_i^1 \otimes e_j^2,
\end{aligned}$$

so that we have:

$$\begin{aligned}
& \psi_t^+ \cdot (e_k^1 \otimes X_{t-}^2 - X_{t-}^1 \otimes X_{t-}^2) \\
& = \frac{1}{m_1 + m_2 - 1} \left[(m_1 + m_2 - 2)e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right].
\end{aligned}$$

It follows that:

$$\begin{aligned}
& \gamma_t^* \cdot \psi_t^+ \cdot (e_k^1 \otimes X_{t-}^2 - X_{t-}^1 \otimes X_{t-}^2) \\
&= \frac{1}{m_1 + m_2 - 1} (\gamma_t^1 \otimes e_j^2 + e_i^1 \otimes \gamma_t^2)^* \cdot \left[(m_1 + m_2 - 2) e_k^1 \otimes e_j^2 - \sum_{k_0 \neq k} e_{k_0}^1 \otimes e_j^2 - \sum_{j_0 \neq j} e_i^1 \otimes e_{j_0}^2 \right] \\
&= \frac{1}{m_1 + m_2 - 1} \left[(m_1 + m_2 - 2) (e_k^1)^* \cdot \gamma_t^1 - \sum_{k_0 \neq k} (e_{k_0}^1)^* \cdot \gamma_t^1 - (e_j^2)^* \cdot \gamma_t^2 - \sum_{j_0 \neq j} (e_{j_0}^2)^* \cdot \gamma_t^2 \right] \\
&= \frac{1}{m_1 + m_2 - 1} \left[(m_1 + m_2 - 1) (e_k^1)^* \cdot \gamma_t^1 - \sum_{k_0} (e_{k_0}^1)^* \cdot \gamma_t^1 - (e_j^2)^* \cdot \gamma_t^2 - \sum_{j_0} (e_{j_0}^2)^* \cdot \gamma_t^2 \right] \\
&= (e_k^1)^* \cdot \gamma_t^1,
\end{aligned}$$

where in the last equality we used the assumption that $\sum_{i=1}^{m_n} [\gamma_t^n]_i = 0$ for $n = 1, 2$. Now noticing that $(e_k^1)^* \cdot \gamma_t^1 = (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_i^1)$, we obtain:

$$1 + \gamma_t^* \cdot \psi_t^+ \cdot (e_k^1 \otimes X_{t-}^2 - X_{t-}^1 \otimes X_{t-}^2) = 1 + (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_i^1).$$

Combining this with the inequality (78), we obtain the inequality (77). Proceeding in a similar way we can also show that for $k = 1, \dots, m_2$:

$$1 + \gamma_t^* \cdot \psi_t^+ \cdot (X_{t-}^1 \otimes e_k^2 - X_{t-}^1 \otimes X_{t-}^2) \geq 0.$$

Applying Lemma 3.3 to the BSDE (76), we obtain the desired result. $\square$

REFERENCES

- [1] C. BENAZZOLI, L. CAMPI, AND L. D. PERSIO, *Mean-field games with controlled jumps*, (2017).
- [2] R. CARMONA AND F. DELARUE, *Probabilistic analysis of mean-field games*, SIAM Journal on Control and Optimization, 51 (2013), pp. 2705–2734.
- [3] R. CARMONA AND D. LACKER, *A probabilistic weak formulation of mean field games and applications*, The Annals of Applied Probability, 25 (2015), pp. 1189–1231.
- [4] R. CARMONA AND P. WANG, *Finite state mean field games with major and minor players*, arXiv preprint [arXiv:1610.05408](https://arxiv.org/abs/1610.05408), (2016).
- [5] A. CECCHIN AND M. FISCHER, *Probabilistic approach to finite state mean field games*, arXiv preprint [arXiv:1704.00984](https://arxiv.org/abs/1704.00984), (2017).
- [6] S. N. COHEN AND R. J. ELLIOTT, *Solutions of backward stochastic differential equations on markov chains*, Commun. Stoch. Anal., (2008), pp. 251–262.
- [7] ———, *Comparisons for backward stochastic differential equations on markov chains and related no-arbitrage conditions*, Ann. Appl. Probab., 20 (2010), pp. 267–311.
- [8] J. DONCEL, N. GAST, AND B. GAUJAL, *Mean-field games with explicit interactions*. Feb. 2016.
- [9] R. ELIE, T. MASTROLIA, AND D. POSSAMAÏ, *A tale of a principal and many many agents*, arXiv preprint [arXiv:1608.05226](https://arxiv.org/abs/1608.05226), (2016).
- [10] R. J. ELLIOTT, L. AGGOUN, AND J. B. MOORE, *Hidden Markov Models: Estimation and Control*, no. 29 in Applications of Mathematics, Springer, New York, 1995.
- [11] A. F. FILIPPOV, *Differential equations with discontinuous righthand sides: control systems*, vol. 18, Springer Science & Business Media, 2013.
- [12] N. FOURNIER AND A. GUILLIN, *On the rate of convergence in wasserstein distance of the empirical measure*, Probability Theory and Related Fields, 162 (2015), pp. 707–738.
- [13] G. GEORGANOPOULOS, *Sur l'approximation des fonctions continues par des fonctions lipschitziennes*, C. R. Acad. Sci. Paris, 264 (1967), pp. 319–321.
- [14] D. A. GOMES, J. MOHR, AND R. R. SOUZA, *Continuous time finite state mean field games*, Applied Mathematics & Optimization, 68 (2013), pp. 99–143.
- [15] D. A. GOMES, R. M. VELHO, AND M.-T. WOLFRAM, *Socio-economic applications of finite state mean field games*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372 (2014).
- [16] Y. HU AND S. PENG, *A stability theorem of backward stochastic differential equations and its application*, Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 324 (1997), pp. 1059–1064.
- [17] J. JACOD AND J. MÉMIN, *Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité*, in Séminaire de Probabilités XV 1979/80, Springer, 1981, pp. 529–546.

- [18] J. JACOD AND A. N. SHIRYAEV, *Skorokhod topology and convergence of processes*, in Limit theorems for stochastic processes, Springer, 1987, pp. 288–347.
- [19] V. KOLOKOLTSOV AND A. BENSOUSSAN, *Mean-field-game model for botnet defense in cyber-security*, Applied Mathematics & Optimization, 74 (2016), pp. 669–692.
- [20] R. MICULESCU, *Approximation of continuous functions by lipschitz functions*, Real Analysis Exchange, 26 (2000), pp. 449–452.
- [21] P. E. PROTTER, *Stochastic differential equations*, in Stochastic Integration and Differential Equations, Springer, 2005, pp. 249–361.
- [22] M. ROYER, *Backward stochastic differential equations with jumps and related non-linear expectations*, Stochastic processes and their applications, 116 (2006), pp. 1358–1376.
- [23] A. SOKOL AND N. R. HANSEN, *Exponential martingales and changes of measure for counting processes*, Stochastic analysis and applications, 33 (2015), pp. 823–843.
- [24] C. VILLANI, *Optimal transport: old and new*, vol. 338, Springer Science & Business Media, 2008.

DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING, PRINCETON UNIVERSITY