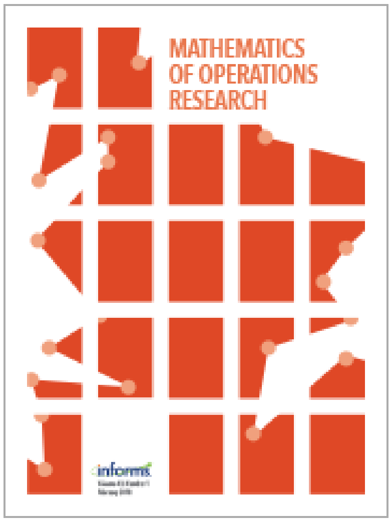

Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Probabilistic Approach to Extended Finite State Mean Field Games

René Carmona , Peiqi Wang

informs
http://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 46, No. 2, May 2021, pp. 471–502
ISSN 0364-765X (print), ISSN 1526-5471 (online)

A Probabilistic Approach to Extended Finite State Mean Field Games

René Carmona,[a] Peiqi Wang[b]

[a] Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey 08544; [b] Bank of America, New York, New York 10281

Contact: rcarmona@princeton.edu, https://orcid.org/0000-0003-1359-711X (RC); peiqiw@alumni.princeton.edu (PW)

Received: August 24, 2018
Revised: November 13, 2019
Accepted: January 12, 2020
Published Online in Articles in Advance: March 3, 2021

MSC2000 Subject Classification: Primary: 49N80; secondary: 90C40
OR/MS Subject Classification: Primary: Games/group decisions: stochastic; secondary: dynamic programming/optimal control: Markov: finite state

https://doi.org/10.1287/moor.2020.1071

Abstract. We develop a probabilistic approach to continuous-time finite state mean field games. Based on an alternative description of continuous-time Markov chains by means of semimartingales 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 at the same time, but also extends 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 a mean field game consists of an approximative Nash equilibrium for the game with a finite number of players under different assumptions of structure and regularity on the cost functions and transition rate between states.

Funding: This work was supported by the Army Research Office [Grant W911NF-17-1-0578] and the Directorate for Mathematical and Physical Sciences [Grant DMS-1716673].

Keywords: mean field games • finite state space • weak formulation of optimal control • approximative Nash equilibrium • McKean–Vlasov BSDE

1. Introduction

Mean field games in which players' states belong to a finite space were first studied in Gomes et al. [21]. The dynamics of each player's state is modeled as a continuous-time Markov chain whose transition rate matrix is a function of the player's control and the probability distribution of all the players' states. By assuming that each player adopts a Markovian strategy, the Nash equilibrium can be characterized by the Hamilton–Jacobi–Bellman equation corresponding to the optimal control of continuous-time Markov chains on the one hand and, on the other hand, a Kolmogorov equation on how the probability distribution of the players' states evolves. Because of the finite nature of the state space, both equations turn out to be ordinary differential equations, and the existence of the solution to this forward–backward system can be obtained by a fixed point argument. Basna et al. [1] study finite state mean field games in which the underlying dynamics follow a time-inhomogeneous nonlinear Feller process. Theoretical aspects of finite state mean field games with major and minor players are investigated in Carmona and Wang [8] in which the existence of Nash equilibria and results on approximate Nash equilibrium for a finite player game of small duration are obtained along with the master equation characterizing the Nash equilibrium. Cecchin and Pelino [11] and Bayraktar and Cohen [2] study the well-posedness and regularity of the solutions to a master equation of finite state mean field games. Continuous-time finite state mean field games are used to model socioeconomic phenomena, such as paradigm shift in a scientific community and consumer choice in Gomes et al. [22]. In Kolokoltsov and Bensoussan [26], the strategic aspect of cyberattack 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.

In this paper, we develop a probabilistic framework for a continuous-time finite state mean field game. Our starting point is a semimartingale representation of continuous-time Markov chains introduced in Elliott et al. [16]: Let $(X_t)_{0\le t\le T}$ be a continuous-time Markov chain with m states that are identified with the m standard basis vectors in $\mathbb{R}^m$ and then we can write

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

Here, $Q(t)$ 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 the 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 a continuous-time Markov chain in a so-called *weak formulation*. 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 theory of BSDEs, the optimality of the control problem can be obtained by optimizing the driver of the BSDE, which coincides with the Hamiltonian function (see Carmona [5, chapter 4, section 1]). It turns out that such a procedure can be transplanted to the case of optimal control of continuous-time Markov chains, thanks to the theory of BSDEs driven by Markov chains developed in Cohen and Elliott [12, 13].

Once the optimal control problem can be characterized by a BSDE, the next step is to develop a probabilistic approach to mean field games. The probabilistic approach to mean field games based on diffusion processes is first proposed in Carmona and Delarue [6], in which the player optimization problem is treated in the *strong* formulation. By applying Pontryagin's maximum principle, optimality is characterized by a forward–backward stochastic differential equation (FBSDE). In Carmona and Lacker [7], the authors later consider the weak formulation of the control of diffusion processes and use the change of measure argument that we briefly described earlier 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 (or mean field) BSDEs (or FBSDEs) for which it was shown that existence of solutions can be obtained by a fixed-point argument à la Schauder.

By developing the weak formulation, our contributions to finite state mean field games are threefold. First, the flexibility of the probabilistic approach allows us to incorporate not only the mean field of the states, but also the mean field of the controls into the dynamics and cost functionals of individual players. Mean field games in which interactions are realized through both states and controls are sometimes referred to as extended mean field games. Gomes and Voskanyan [20] study the system of partial differential equations (PDEs) characterizing an extended mean field game in which the dynamics of players are deterministic. Cardaliaguet and Lehalle [4] model the optimal liquidation within a crowd of investors as an extended mean field game and provide a closed-form solution when the investors' preference is characterized by a trade-off between terminal wealth and quadratic inventory costs. Generally speaking, the analysis of mean field games of control is known to be notoriously intractable via the PDE method because of the difficulties in deriving the equation obeyed by the flow of probability measures of the optimal strategies. However, in the probabilistic framework, the mean field of states and the mean field of controls can be dealt with in similar manners although the treatment of the mean field of controls is more involved in terms of the topological arguments needed to handle the appropriate spaces of controls.

Second, using the weak formulation, we are able to nuance the information structures of the agents' admissible strategies. In the literature of stochastic optimal control and stochastic differential games, admissible controls are traditionally classified into the categories of open- and closed-loop controls. Open-loop controls are controls adapted to the underlying filtration, which is often generated by the noise processes. Closed-loop controls, on the other hand, are controls that are adapted to the filtration generated by the history of the state process. In particular, Markovian controls form a subset of closed-loop controls in which the players are only allowed to observe their current state. Most of the past work on mean field games considers open-loop strategies for which the identification the optimal strategies can be tackled by a form of the Pontryagin maximum principle as in Carmona and Delarue [6] or Markovian strategies via the dynamic programming principle as in Lasry and Lions [27]. For mean field games with homogeneous populations of players, the choice of the information structure does not seem to matter. However, when a major player is introduced into the game, Carmona and Wang [9] show that different information structures lead to different equilibrium outcomes. Nevertheless, a more flexible information structure is always desirable from a practical modeling perspective. For finite state mean field games, the existing analytical approach only allows Markovian strategies. In the weak formulation, which we are about to introduce, the underlying information structure is generated by the state process itself; therefore, the admissible strategies can accommodate closed-loop strategies, including the strategies depending on the past history of player's states.

Finally, the weak formulation we develop for the finite state mean field game serves as a launching pad to tackle the finite state mean field agent–principal problem. Such a 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 that 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 the agent's distribution is minimized. To the best of our knowledge, this type of problem is first investigated in Elie et al. [15], in which the agent's dynamics are 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 the finite-state mean field agent–principal problem, which could lead to potential applications in epidemics and cybersecurity.

We would also like to mention a few existing works related to our paper. In Cecchin and Fischer [10], the authors propose a probabilistic framework for finite-state mean field games in which the dynamics of the states of the players are given by stochastic differential equations driven by Poisson random measures. By using Ky Fan's fixed-point theorem, the authors obtain the existence and uniqueness of the Nash equilibrium in relaxed open-loop as well as relaxed feedback controls. Then, under additional assumptions that guarantee the uniqueness of optimal nonrelaxed feedback controls, the authors deduce the existence of Nash equilibria in nonrelaxed feedback form. In Doncel et al. [14], continuous-time mean field games with finite state space and finite action space are studied. The authors prove the existence of a Nash equilibrium among relaxed feedback controls. In Benazzoli et al. [3], the authors investigate mean field games in which each player's state follows a jump-diffusion process, and the player controls the sizes of the jumps. The approach is based on the weak formulation of stochastic control and martingale problems. The 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 games. It is based on a semimartingale representation of continuous-time Markov chains and an argument of change of measure. We state the assumptions used throughout the paper and give the precise definition of a Nash equilibrium in the weak formulation. In Section 3, we analyze the optimization problem of a representative player when it faces a fixed mean field of states and controls. We characterize its value function and the optimal control using a BSDE driven by a Markov chain. Section 4 is devoted to the existence of Nash equilibria. The proof is largely inspired by Carmona and Lacker [7], who deal with the diffusion case. It is split into two steps for the sake of clarity. In the first step, we carefully construct the topological spaces for the mean fields of states and controls so that we can identify compact sets in these spaces. Extra attention is required in the case of the mean field of controls, for which we use a randomization technique and a special topology called the stable topology in order to identify compact subsets. The second step is to define the function that maps the mean fields to the distribution of individual players' best responses and to show that such a mapping is continuous and stable with respect to the topology constructed in the first step. After these two steps, we apply Schauder's fixed-point theorem to conclude the proof of existence. Section 5 addresses the uniqueness of Nash equilibria. Finally, in Section 6, we formulate the model for a game with a finite number of players and show that a Nash equilibrium of the mean field game provides an approximate Nash equilibrium of the game with finite many players. Section 7, the appendix, contains a few useful results regarding BSDEs driven by multiple continuous-time Markov chains, which are used throughout the paper.

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 $diag(x)$ the square diagonal matrix whose diagonal elements are given by the entries of x, and we denote by $\|x\|$ the Euclidean norm 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 Doléans–Dade exponential of L. See Protter [29, chapter II.6] 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 Elliott et al. [16] and later developed in Cohen and Elliott [12, 13]. If $\mathbf{X} = (X_t)_{0 \le t \le T}$ 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 $E = \{e_1, \dots, e_m\}$. However, we frequently identify these states with their labels $1, 2, \dots, m$ for the sake of simplicity. We assume that the sample paths $t \to X_t$ are *càdlàg*, that is, 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 $\mathcal{F}_t := \sigma\{X_s, s \le t\}$ 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 p° on the set E. It is 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 p° and transition rates between any two different states equal to one. This means that, for $i, j \in \{1, \dots, m\}$, $i \ne j$, 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 Elliott et al. [16, appendix B], the process X has the representation

$$X_t = X_0 + \int_{(0,t]} Q^0 \cdot X_{s-} ds + \mathcal{M}_t, \tag{1}$$

where Q^0 is the square matrix with diagonal elements all equal to $-(m-1)$ and off-diagonal elements all equal to one, and $\mathcal{M} = (\mathcal{M}_t)_{t\ge 0}$ 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 1. The representation originally proposed in Elliott et al. [16] is

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

However, because 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 is suitable for the change-of-measure argument in what follows.

We refer to the probability measure $\mathbb{P}$ as the *reference measure* on the sample space. The first step of the weak formulation of the mean field game consists of 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

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

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:

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

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

Assumption 1.

i. *For all $(t, \alpha, p, v) \in [0,T] \times A \times \mathcal{S} \times \mathcal{P}(A)$, the matrix $Q(t, \alpha, p, v)$ 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 $i \ne j$, 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*

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

Here, $\mathcal{W}_1$ denotes the one-Wasserstein distance between probability measures on A.

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

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

Remark 2. Item (ii) in Assumption 1, which assumes that the transition rate between any two states admits a strictly positive lower bound, is analog to the nondegeneracy condition in the diffusion-based mean field game models.

It guarantees that the probability measure $\mathbb{Q}^{(\alpha,p,\nu)}$ defined in (7) is equivalent to the reference measure $\mathbb{P}$. This nondegeneracy condition is frequently assumed in the existing literature on finite-state mean field games. See, for example, Bayraktar and Cohen [2, assumption 3.1] and Cecchin and Pelino [11, section 2.1] in which the control is the transition rate of the controlled Markov chain and is assumed to have a strictly positive lower bound.

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 Kolokoltsov and Bensoussan [26] as well as in the extended version of the model that includes an attacker studied in Carmona and Wang [8]. When that happens, we need to use a different reference probability measure $\mathbb{P}$: we set the transition rate to one for all the jumps except for those that are forbidden, for which we set the transition rate to zero. 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 Cohen and Elliott [12].

Lemma 1. *The predictable quadratic variation of the martingale $\mathcal{M}$ under $\mathbb{P}$ is given by the formula:*

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

where ψ_t is given by

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

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

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

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

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

The seminorm $\|\cdot\|_{X_{t-}}$ can be rewritten in a more explicit way. For $i \in \{1,\ldots,m\}$, 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 $\|Z\|_{X_{t-}} = \sum_{i=1}^m \mathbb{1}(X_{t-} = i)\|Z\|_{e_i}$.

Because 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 the Girsanov theorem on changes of probability measure. See Protter [29, theorem III.41] or Sokol and Hansen [30, lemma 4.3].

Lemma 2. *Let $T > 0$ and $\mathcal{M} = (\mathcal{M}_t)_{t\geq 0}$, $\mathbf{L} = (L_t)_{t\geq 0}$ be two local martingales defined on $[0,T]$ such that $\Delta L_t := L_t - L_{t-} \geq -1$. 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 $[\mathcal{M}, \mathbf{L}]$ is integrable under $\mathbb{P}$, then $\mathcal{M} - \langle \mathcal{M}, \mathbf{L}\rangle$ is a local martingale under $\mathbb{Q}$, where the predictable quadratic covariation $\langle \mathcal{M}, \mathbf{L}\rangle$ 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 $\mathbb{A}$ to be the collection of $\mathbb{F}$-predictable processes $\boldsymbol{\alpha} = (\alpha_t)_{t\in[0,T]}$ 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 $\boldsymbol{\nu} = (\nu_t)_{t\in[0,T]}$ on A, we define the scalar martingale $\mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}$ under $\mathbb{P}$ by

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

Clearly, its jumps are given by

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

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

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

where the second-to-last equality is because $\sum_{k=1}^m (q(t,i,k,\alpha_t,p_t,\nu_t) - q_{i,k}^0) = 0$. Therefore, we have $\Delta L_t^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})} \geq -1$. By Protter [29, theorem III.45] and the remark that follows, in order to show that $\mathcal{E}(\mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})})$ is uniformly integrable, it suffices to show $\mathbb{E}[\exp(\langle \mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}, \mathbf{L}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}\rangle_T)] < \infty$. This is straightforward because we have

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

The integrand is bounded by some constant by Assumption 1.

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

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

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

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

By Lemma 2, we know that the process $\mathcal{M}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}$, defined as

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

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

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

This means that, under the measure $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}$, the stochastic intensity rate of $\mathbf{X}$ is given by $Q(t,\alpha_t,p_t,\nu_t)$. In addition, because $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}$ and $\mathbb{P}$ coincides on $\mathcal{F}_0$, the law of X_0 under $\mathbb{Q}^{(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu})}$ is the same as under the reference measure $\mathbb{P}$, which is p°. In particular, when $\boldsymbol{\alpha}$ is a Markov control, that is, 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 3. In the optimal control literature, admissible controls are often classified into the categories of open- and closed-loop controls. Open-loop controls are usually referred to as 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 setup, 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 $f:[0,T]\times E\times A\times\mathcal{S}\times\mathcal{P}(A)\to\mathbb{R}$ and $g:E\times\mathcal{S}\to\mathbb{R}$ 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. *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*

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

$$|g(e_i,p)-g(e_i,p')|\le C\|p-p'\|. \tag{11}$$

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

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

Each player aims at minimizing expected cost, that is, the player solves the optimization problem

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

The key idea of the theory of mean field games lies in the limit scenario of infinitely many players in the game, in which a single player's strategy $\boldsymbol{\alpha}$ does not alter the mean field $(\mathbf{p},\boldsymbol{\nu})$. Therefore, when a player searches for the best response to the other players, the player solves the player's own optimization problem considering $(\mathbf{p},\boldsymbol{\nu})$ 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 $(\mathbf{p},\boldsymbol{\nu})$. Intuitively, the Nash equilibrium should be for a game played by a typical individual against the fields of states and controls created by the other players. This justifies the following definition of a Nash equilibrium for the weak formulation of finite-state mean field games.

Definition 1. Let $\mathbf{p}^*:[0,T]\to\mathcal{S}$ and $\boldsymbol{\nu}^*:[0,T]\to\mathcal{P}(A)$ be two measurable functions, and $\boldsymbol{\alpha}^*\in\mathbb{A}$. We say that the tuple $(\boldsymbol{\alpha}^*,\mathbf{p}^*,\boldsymbol{\nu}^*)$ is a Nash equilibrium for the weak formulation of the mean field game if

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

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

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

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

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

3. Individual Players 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 a Continuous-Time Markov Chain

We first recall some of the results on BSDEs driven by continuous-time Markov chains obtained in Cohen and Elliott [12, 13]. 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 $\mathbb{R}^m$:

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

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 $t \to F(w, t, y, z)$ is predictable for all y, z.

Recalling the definition (4) of the stochastic seminorm $\|\cdot\|_{X_{t-}}$, we have the following existence and uniqueness result. See Cohen and Elliott [13, theorem 1.1].

Lemma 3. *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

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

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 proven by mimicking the argument used in the proof of Hu and Peng [23, theorem 2.1].

Lemma 4. *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 with the same constant. In addition, assume that the following conditions hold:*

i. $\lim_{n\to\infty} \mathbb{E}[|\xi^n - \xi^0|^2] = 0$.
ii. *For each $t \leq T$, $\lim_{n\to\infty} \mathbb{E}[(\int_t^T |F^n(w, s, Y_s^0, Z_s^0) - F^0(w, s, Y_s^0, Z_s^0)|ds)^2] = 0$.*
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

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

Finally, we state a crucial comparison result for linear BSDEs. See Cohen and Elliott [13, theorem 3.16].

Lemma 5. *Let $\boldsymbol{\gamma}$ be a bounded predictable process in $\mathbb{R}^m$, $\boldsymbol{\beta}$ a bounded predictable process in $\mathbb{R}$, $\boldsymbol{\phi}$ a nonnegative predictable process in $\mathbb{R}$ such that $\mathbb{E}[\int_0^T |\phi_t|^2 dt] < +\infty$, and ξ a nonnegative square-integrable $\mathcal{F}_T$-measurable random variable in $\mathbb{R}$, and let us assume that $(\mathbf{Y}, \mathbf{Z})$ solves the linear BSDE:*

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

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 need to consider BSDEs driven by multiple independent continuous-time Markov chains. It turns out that all the preceding results 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

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

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

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

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.

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

ii. *$\hat{a}_i$ 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 $i \in \{1, \ldots, m\}$, $z, z' \in \mathbb{R}^m$, $p, p' \in \mathcal{S}$:*

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

Remark 4. 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 holds is to assume strong convexity of the running cost function f and let the control be the transition rate. This is the set up in Gomes et al. [21], in which the Hamiltonian H_i can be regarded as the generalized Legendre transform of the running cost function, and one can show that $\hat{a}_i$ is Lipschitz with a uniform Lipschitz constant (see Gomes et al. [21, proposition 1]). More generally, Assumption 3 holds with a transition rate function that is linear on the control. We state a set of assumptions with a strongly convex cost function and linear transition rate function upholding Assumption 3:

Assumption 4.

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 $q_0 : [0, T] \times \{1, \cdots, m\}^2 \times \mathcal{S} \times \mathcal{P}(A) \to \mathbb{R}$ and $q_1 : [0, T] \times \{1, \cdots, m\}^2 \times \mathcal{S} \to \mathbb{R}^l$ 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 $i \in \{1, \ldots, m\}$, the mapping $f_0(\cdot, e_i, \cdot, \cdot)$ (respectively, $f_1(\cdot, e_i, \cdot, \cdot)$) is continuous on $[0, T] \times A \times \mathcal{P}$ ($[0, T] \times \mathcal{S} \times \mathcal{P}(A)$).*

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

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

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

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

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

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

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

From item (i) of Assumption 3 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 $\alpha \to H(t, x, z, \alpha, p, \nu)$, and the minimum equals $\hat{H}(t, x, z, p, \nu)$. In addition, from Assumptions 1–3 and the definition of the stochastic seminorm $\|\cdot\|_{X_{t-}}$, it is easy to deduce the regularity of the mappings $\hat{H}$ and $\hat{a}$.

In the rest of Section 3, we let Assumptions 1–3 hold.

Lemma 6. *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*

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

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

Proof. Inequality (27) is an easy consequence of Assumption 3 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

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

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

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

Exchanging the roles of z and z', we obtain

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

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

3.3. Player's Optimization Problem

In this section, we show that the optimization problem of the player facing a given mean field of states and controls can be characterized by a BSDE driven by the continuous-time Markov chain $\mathbf{X}$. Let us fix measurable flows $\mathbf{p} : [0,T] \to \mathcal{S}$ and $\boldsymbol{\nu} : [0,T] \to \mathcal{P}(A)$, an admissible strategy $\boldsymbol{\alpha} \in \mathbb{A}$, and let us consider the BSDE:

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

Lemma 7. *The BSDE (28) admits a unique solution $(\mathbf{Y}, \mathbf{Z})$ and $J(\boldsymbol{\alpha}, p, \nu) = \mathbb{E}^{\mathbb{P}}[Y_0]$.*

Proof. From the boundedness of the transition rate function q guaranteed by Assumption 1, it is easy to check that the driver function H of the BSDE (28) is Lipschitz in z with respect to the seminorm $\|\cdot\|_{X_{t-}}$. Therefore, by Lemma 3, it admits a unique solution $(\mathbf{Y}, \mathbf{Z})$. Moreover, we have

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

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

Now, we consider the following BSDE:

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

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

Proposition 1. *For any measurable function $\mathbf{p}$ from $[0,T]$ to $\mathcal{S}$ and any measurable function $\boldsymbol{\nu}$ from $[0,T]$ to $\mathcal{P}(A)$, the BSDE (29) admits a unique solution $(\mathbf{Y},\mathbf{Z})$. The value function of the optimal control problem (13) is given by $V(\mathbf{p},\boldsymbol{\nu}) = \mathbb{E}^{\mathbb{P}}[Y_0]$ and the process $\hat{\boldsymbol{\alpha}}^{(p,\nu)}$ defined by*

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

is an optimal control. In addition, if $\boldsymbol{\alpha}' \in \mathbb{A}$ is an optimal control, we have $\alpha'_t = \hat{\alpha}_t^{(\mathbf{p},\boldsymbol{\nu})}$, $dt \otimes d\mathbb{P}$ 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 6. Let $(\mathbf{Y},\mathbf{Z})$ be this unique solution and define the process $\hat{\boldsymbol{\alpha}}$ by $\hat{\alpha}_t := \hat{a}(t, X_{t-}, Z_t, p_t)$. Recall the definition of $\hat{a}$ in Equation (25). We have

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

Because $\hat{a}_i$ 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. Because both the processes $t \to X_{t-}$ and $\mathbf{Z}$ are predictable, we conclude that $\hat{\boldsymbol{\alpha}}$ is a predictable process and, therefore, an admissible control.

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

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

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

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

We can further decompose the driver of the preceding BSDE as

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

Define the processes $\boldsymbol{\psi}$ and $\boldsymbol{\gamma}$ 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 $\boldsymbol{\psi}$ and $\boldsymbol{\gamma}$ defined previously and $\boldsymbol{\beta} = 0$. Clearly, $\boldsymbol{\psi}$ and $\boldsymbol{\gamma}$ are both predictable. Because $\hat{\alpha}_t$ minimizes the Hamiltonian, $\boldsymbol{\psi}$ is nonnegative. The boundedness of $\boldsymbol{\gamma}$ 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 preceding inequality holds clearly. So we assume that $X_{t-} = e_i \neq e_j$. We have $\psi_t^+ \cdot (e_j - X_{t-}) = ((m-1)/m)e_j - \sum_{i \neq j}(1/m)e_i$. Therefore, when $X_{t-} = e_i \neq e_j$, we have

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

where the last equality is because $\sum_k (q(t, i, k, \alpha_t, p_t, \nu_t) - q_{i,k}^0) = 0$. 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 5, we conclude that $\Delta\mathbf{Y}$ is nonnegative, and in particular, $Y_0^{\alpha} \ge Y_0$. Because $\boldsymbol{\alpha}$ is an arbitrary admissible control, in light of Lemma 7, this means that $\mathbb{E}^{\mathbb{P}}[Y_0] \le \inf_{\alpha\in\mathbb{A}} J(\boldsymbol{\alpha},\mathbf{p},\boldsymbol{\nu}) = V(\mathbf{p},\boldsymbol{\nu})$. Finally, we notice that Y_0 is the expected total cost when the control is $\hat{\boldsymbol{\alpha}}$. We conclude that $\hat{\boldsymbol{\alpha}}$ is an optimal control and $\mathbb{E}^{\mathbb{P}}[Y_0] = V(p,\nu)$.

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:

$$Y_t' = \int_t^T H(s, X_{s-}, Z_s', \alpha_s', p_s, \nu_s)ds - \int_t^T (Z_s')^* \cdot d\mathcal{M}_s. \tag{32}$$

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

$$\begin{aligned}
Y_0 - Y_0' &= \int_0^T \left[H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t) - H(t, X_{t-}, Z_t', \alpha_t', p_t, \nu_t)\right]dt - \int_0^T (Z_t - Z_t')^* \cdot d\mathcal{M}_t \\
&= \int_0^T \left[X_{t-}^* \cdot \left(Q(t, \hat{\alpha}_t, p_t, \nu_t) - Q^0\right) \cdot Z_t - X_{t-}^* \cdot \left(Q(t, \alpha_t', p_t, \nu_t) - Q^0\right) \cdot Z_t'\right]dt \\
&\quad + \int_0^T \left[f(t, X_{t-}, \hat{\alpha}_t, p_t, \nu_t) - f(t, X_{t-}, \hat{\alpha}_t, p_t, \nu_t)\right]dt - \int_0^T (Z_t - Z_t')^* \cdot d\mathcal{M}_t \\
&= \int_0^T \left[f(t, X_{t-}, \hat{\alpha}_t, p_t, \nu_t) - f(t, X_{t-}, \alpha_t', p_t, \nu_t) + X_{t-}^* \cdot \left(Q(t, \hat{\alpha}_t, p_t, \nu_t) - Q(t, \alpha_t', p_t, \nu_t)\right) \cdot Z_t\right]dt \\
&\quad - \int_0^T (Z_t - Z_t')^* \cdot \left[d\mathcal{M}_t - \left(Q^*(t, \alpha_t', p_t, \nu_t) - Q^0\right) \cdot X_{t-}dt\right] \\
&= \int_0^T \left[H(t, X_{t-}, Z_t, \hat{\alpha}_t, p_t, \nu_t) - H(t, X_{t-}, Z_t, \alpha_t', p_t, \nu_t)\right]dt - \int_0^T (Z_t - Z_t')^* \cdot d\mathcal{M}_t^{(\alpha',\mathbf{p},\nu)}.
\end{aligned}$$

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

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

where the last inequality is because $\hat{\alpha}_t$ 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 because the minimizer of the Hamiltonian is unique by Assumption 3. Because $\mathbb{P}$ is equivalent to $\mathbb{Q}^{(\alpha',\mathbf{p},\nu)}$, we have $\hat{\alpha}_t = \alpha_t'$, $dt \otimes d\mathbb{P}$ a.e. □

4. Existence of Nash Equilibria

We state the main result of this section:

Theorem 1. Under Assumptions 1–3, there exists a Nash equilibrium $(\boldsymbol{\alpha}^*, \mathbf{p}^*, \boldsymbol{\nu}^*)$ for the weak formulation of the finite-state mean field game in the sense of Definition 1.

The rest of this section is devoted to the proof of Theorem 1. As in the case of diffusion-based mean field games, we rely on a fixed-point argument to show existence of Nash equilibria. We start from a measurable function $\mathbf{p} : [0,T] \to \mathcal{S}$ and a measurable function $\boldsymbol{\nu} : [0,T] \to \mathcal{P}(A)$, where we recall that $\mathcal{S}$ is the m-dimensional simplex, which we identify with the space of probability measures on E, and $\mathcal{P}(A)$ 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 $\hat{\boldsymbol{\alpha}}^{(\mathbf{p},\nu)}$ 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 identify a Nash equilibrium if we find a fixed point for the mapping $(\mathbf{p},\boldsymbol{\nu}) \to \hat{\mathbb{Q}}^{(\mathbf{p},\nu)}_{\#(X_t,\hat{\alpha}_t^{(\mathbf{p},\nu)})}$.

In practice, however, the implementation of the aforementioned fixed-point argument 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 resort to different descriptions of the mean field for the state and the control. For the mean field of the state, because we have assumed from the very beginning that $\mathbf{X}$ is a càdlàg process, we 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 Jacod and Shiryaev [25] 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 Carmona and Lacker [7]. 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-one 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 $\boldsymbol{\nu}$ 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 Jacod and Mémin [24], 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.

In the rest of Section 4, we let Assumptions 1–3 hold.

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 $d_E(x,y):=\mathbb{1}(x\neq y)$. Clearly, (E,d_E) is a Polish space. Then, the Skorokhod space

$$D:=\big\{x:[0,T]\to E, x \text{ is cadlag and left continuous on } T\big\} \tag{33}$$

is endowed with the J1 metric:

$$d_D(x,y):=\inf_{\lambda\in\Lambda}\max\left\{\sup_{t\le T}|\lambda(t)-t|,\sup_{t\le T}|y(\lambda(t))-x(t)|\right\}, \tag{34}$$

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

$$\mathcal{P}_0:=\left\{\mathbb{Q}:\frac{d\mathbb{Q}}{d\mathbb{P}}=L,\ \text{ with }\ \mathbb{E}^{\mathbb{P}}\big[L^2\big]\le C_0\right\}. \tag{35}$$

Here, C_0 is a constant that we specify later (see the proof of Proposition 4). We have the following result:

Proposition 2. *The set $\mathcal{P}_0$ 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 that 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 two consecutive jumps are greater or equal than δ. We now show that $D_{\delta,K}$ is compact in D. Because D is a 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, and $\delta\le t_n^1<t_n^2<\cdots<t_n^{k_n}\le T-\delta$ 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 $x_n^0,x_n^1,\dots,t_n^{k_n}$ are the values taken by x_n in each interval defined by the jumps. Then, we can represent x_n using the vector y_n of dimension $2(K+1)$:

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

In this representation, the first coordinate of y_n is the number of jumps. Coordinate two to $K+1$ are the times elapsed between jumps defined previously, and if there are fewer than K jumps, we complete the vector by zero. Coordinates $K+2$ to $2(K+1)$ are the values taken by the path x_n and completed with zero. 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 $i \le k_n$ and $\sum_{i=1}^{k_n} \Delta t_n^i \le T - \delta$, whereas the rest of the coordinates of y_n belong 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, because 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 $n \ge N_0$, we have $k_n = k$ and $x_n^i = x^i$ for all $i \le k$. In addition, we have that Δt_n^i converges to Δt^i for all $i \le k$, where $\Delta t^i \ge \delta$ for all $i \le k$ and $\sum_{i=1}^k \Delta t^i \le T - \delta$. We consider the path represented by the vector y:

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

Clearly, x belongs to the set $D_{\delta.K}$, and it is straightforward to verify that x is the limit of the sequence $(x_n)_{n\ge 0}$ in the 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 $\mathbb{P}(D_{\delta.K}) \ge 1 - \epsilon$. 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 the first jump as well as the time between consecutive jumps thereafter, which we denote by $\Delta t_1, \Delta t_2, \ldots$, are independent and identically distributed exponential random variables of parameter $(m-1)$ under the measure $\mathbb{P}$. We have

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

For each $k = 1, \ldots, K$, we have

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

It follows that

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

We can first pick K large enough for $(\exp(-(m-1)T) \sum_{k=0}^K (m-1)^k T^k / k!)$ to be greater than $(1 - \epsilon/2)$. Then, we pick δ small enough so that the rest of the terms are 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 $\mathbb{P}(D \setminus D_{\delta,K}) \le \epsilon^2 / C_0$. For all $\mathbb{Q} \in \mathcal{P}_0$, we have $d\mathbb{Q}/d\mathbb{P} = L$ and $\mathbb{E}^{\mathbb{P}}[L^2] \le C_0$, and by the Cauchy–Schwartz inequality, we obtain

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

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 a path space to the convergence in $\mathcal{S}$, that is, measures on state space. We define the function π by

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

and prove the following result:

Lemma 8. *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 $t \notin \mathcal{D}(\mu)$,*

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

Proof. Define $\mathcal{D}(\mu) := \{0 \le t \le T;\ \mu(X_t - X_{t-} \neq 0) > 0\}$. By Jacod and Shiryaev [25, lemma 3.12], the set $\mathcal{D}(\mu)$ is, at most, countable with probability one. In addition, we have $T \notin \mathcal{D}(\mu)$ because all the paths in D are left-continuous on T. In light of Jacod and Shiryaev [25, proposition 3.14], we have that $\mu^n_{\#X_t}$ converges to $\mu_{\#X_t}$ weakly for all $t \notin \mathcal{D}(\mu)$. To conclude, we use the fact that $\mu^n_{\#X_t}$ for all $t \in [0,T]$ and n are counting measures on the discrete set E. □

We now turn to the mean field of controls. 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-one distance (see Villani [31, theorem 6.9]). $(\mathcal{P}(A), \mathcal{W}_1)$ is a Polish space. Because 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 $\sigma-$algebra denoted by $\mathcal{B}(\mathcal{P}(A))$. We endow $[0,T]$ with its Borel $\sigma-$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 the mean field of control. We introduce the stable topology on this space:

Definition 2. 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 $\eta \to \int g(t,m)\eta(dt,dm)$ are continuous for all bounded and measurable mappings g defined on $[0,T] \times \mathcal{P}(A)$ such that $m \to g(t,m)$ 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 3. *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 Jacod and Mémin [24, proposition 2.10] 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 Jacod and Mémin [24, theorem 2.8] and obtain the compactness of $\mathcal{R}$.

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

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

Lemma 9. *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 $g : \mathcal{P}(A) \to \mathbb{R}$ is a bounded Lipschitz function (with respect to the Wasserstein-one distance on $\mathcal{P}(A)$). Then, the stable topology introduced in* Definition 2 *is the coarsest topology that makes the mappings $\eta \to \int_{[0,T]\times\mathcal{P}(A)} f(t,\nu)\eta(dt,d\nu)$ 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 $g : \mathcal{P}(A) \to \mathcal{R}$ is a bounded and uniformly continuous function. Then, clearly, we have $\mathcal{H} \subset \mathcal{H}_0$. By Jacod and Mémin [24, proposition 2.4], the stable topology is the coarsest topology under which the mappings $\eta \to \int_{[0,T]\times\mathcal{P}(A)} f(t,\nu)\eta(dt,d\nu)$ 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 $\int f(t,\nu)\eta^n(dt,d\nu) \to \int f(t,\nu)\eta^0(dt,d\nu)$ for all $f \in \mathcal{H}$, then we have $\int f(t,\nu)\eta^n(dt,d\nu) \to \int f(t,\nu)\eta^0(dt,d\nu)$ 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 Georganopoulos [19] (see also Miculescu [28]) shows that g can be approximated uniformly by a bounded Lipschitz continuous function. That is, for all $\epsilon > 0$, we can find $g_\epsilon \in \mathcal{H}$ such that $\sup_{\nu\in\mathcal{P}(A)} |g_\epsilon(\nu) - g(\nu)| \le \epsilon/3$. By our assumption, we have $\int 1_B(t) g_\epsilon(\nu)\eta^n(dt, d\nu) \to \int 1_B(t) g_\epsilon(\nu)\eta^0(dt, d\nu)$. Therefore, there exists N_0 such that $|\int 1_B(t) g_\epsilon(\nu)\eta^n(dt, d\nu) - \int 1_B(t) g_\epsilon(\nu)\eta^0(dt, d\nu)| \le \epsilon/3$ for all $n \ge N_0$. Combining these facts we have, for $n \ge N_0$,

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

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

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 10. *$\mathcal{R}_0$ is a convex and compact subset of $\mathcal{R}$.*

Proof. We apply Jacod and Mémin [25, theorem 2.8]. 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. □

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

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

Because we changed the way we represent the mean field of controls, we need to modify accordingly the definition of the transition rate matrix as well as the cost functionals in order to make them compatible with the randomization procedure. For any function $F : \mathcal{P}(A) \to \mathbb{R}$ possibly containing other arguments, we denote $\underline{F} : \mathcal{P}(\mathcal{P}(A)) \to \mathbb{R}$ 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. Because the terminal cost g does not depend on the mean field of controls, we do not need to consider its randomized version.

Recall from Assumption 3 that the minimizer $\hat{a}_i$ of the reduced Hamiltonian is only a function of t, z, and p. Consequently, for $\underline{H}$, $\underline{H}_i$, $\underline{\hat{H}}$, and $\underline{\hat{H}}^i$, which are the randomized versions of H, H_i, $\hat{H}$, and $\hat{H}_i$, respectively, we still have

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

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

Lemma 11. *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*

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

and

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

Here, $\bar{\mathcal{W}}_1(m,m')$ is the Wasserstein-one distance on the space of probability measure on $\mathcal{P}(A)$ defined by

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

Proof. We have

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

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

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

and combined with the preceding estimation, we obtain the desired inequality for $\underline{\hat{H}}$. The Lipschitz property for $\underline{q}$ can be proved in the same way. □

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:

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

Denote by $\hat{\boldsymbol{\alpha}}^{(\mu,\eta)}$ the predictable process $t \to \hat{a}(t, X_{t-}, Z_t^{(\mu,\eta)}, \pi(t, \mu))$, which is the optimal control of the player facing 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

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

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

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

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

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

$$
\Phi^\eta : \mathcal{P} \times \mathcal{R}_0 \ni (\mu, \eta) \to \mathcal{L}(dt) \times \delta_{\hat{\mathbb{Q}}^{(\mu,\eta)}_{\#\hat{\alpha}_t^{(\mu,\eta)}}}(d\nu) \in \mathcal{R}_0, \tag{45}
$$

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

Remark 5. Before delving into its properties, 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 $\mathcal{L}(dt) \times \delta_{\hat{\mathbb{Q}}^{(\mu,\eta)}_{\#\hat{\alpha}_t^{(\mu,\eta)}}}(d\nu))$ do not depend on which solution to the BSDE (41) we use to construct $\hat{\boldsymbol{\alpha}}^{(\mu,\eta)}$, $\mathbf{L}^{(\mu,\eta)}$, and $\mathcal{E}(L^{(\mu,\eta)})$. To this end, let us consider $(\mathbf{Y}, \mathbf{Z})$ and $(\mathbf{Y}', \mathbf{Z}')$ to be two solutions to BSDE (41), $\hat{\boldsymbol{\alpha}}$ and $\hat{\boldsymbol{\alpha}}'$ 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 the uniqueness of the solution to (41), we have $\mathbb{E}[\int_0^T \|Z_t' - Z_t\|_{X_{t-}}^2 dt] = 0$. Using the Lipschitz continuity of $\hat{a}$ and q, it is straightforward to show $\mathbb{E}[\int_0^T \|\alpha_t' - \alpha_t\|^2 dt] = 0$ and eventually $\mathbb{Q} = \mathbb{Q}'$.

Proposition 4. *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 under 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

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

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

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

because $dL_t^{(\mu,\eta)} = X_{t-}^* \cdot (\underline{Q}(t, \hat{\alpha}_t^{(\mu,\eta)}, \pi(t,\mu), \eta_t) - Q^0) \cdot \psi_t^+ \cdot d\mathcal{M}_t$ 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

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

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

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

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 preceding SDE between zero and $T \wedge T_n$ and taking the expectation under $\mathbb{P}$, we obtain

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

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

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

Here, 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 $\mathbb{E}^{\mathbb{P}}[W_T^2] \le C_0$. □

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 $(\mu^{(n)}, \eta^{(n)})_{n\ge1}$ 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 $\mathbf{Y}^{(\mu^{(n)},\eta^{(n)})}$, $\mathbf{Z}^{(\mu^{(n)},\eta^{(n)})}$, $\hat{\boldsymbol{\alpha}}^{(\mu^{(n)},\eta^{(n)})}$, $\mathbf{L}^{(\mu^{(n)},\eta^{(n)})}$, and $\hat{\mathbb{Q}}^{(\mu^{(n)},\eta^{(n)})}$, respectively, by $\mathbf{Y}^{(n)}$, $\mathbf{Z}^{(n)}$, $\hat{\boldsymbol{\alpha}}^{(n)}$, $\mathbf{L}^{(n)}$, and $\mathbb{Q}^{(n)}$ for $n \ge 0$. 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 12. *Without any loss of generality, we may assume that there exists a constant C such that $\|Z_t^{(0)}\|_{X_{t-}} \le C$ for all $(\omega, t) \in \Omega \times [0, T]$.*

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

$$0 = \frac{dV_i(t)}{dt} + \underline{\hat{H}}_i\left(t, V(t), p_t^{(0)}, \eta_t^{(0)}\right) + \sum_{j \neq i}[V_j(t) - V_i(t)],$$

$$V_i(T) = g\left(e_i, p_T^{(0)}\right), \quad i = 1, \dots, m. \tag{47}$$

Set $\zeta: [0,T] \times \mathbb{R}^m \ni (t,v) \to [\zeta_1(t,v), \dots, \zeta_m(t,v)] \in \mathbb{R}^m$, where $\zeta_i(t,v) := \underline{\hat{H}}_i(t, v, p_t^{(0)}, \eta_t^{(0)}) + \sum_{j \neq i}[v_j - v_i]$. By Lemma 11, we see that $t \to \zeta(t,v)$ is measurable for all $v \in \mathbb{R}^m$, and $v \to \zeta(t,v)$ is Lipschitz in v uniformly in t. By Filippov [17, theorems 1 and 2], 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 $\mathbf{Y}$, we obtain

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

where, in the last equality, we use the fact that $dX_s = Q^0 \cdot X_{s-}ds + d\mathcal{M}_s$ and $V_t = Z_t$. Therefore, $(\mathbf{Y}, \mathbf{Z})$ and $(\mathbf{Y}^{(0)}, \mathbf{Z}^{(0)})$ solve the same BSDE. As we discuss in Remark 5, we may assume that $\mathbf{Z}^{(0)} = \mathbf{Z}$. Therefore, $Z_t^{(0)} = V(t)$. It follows from the continuity of $t \to V(t)$ that $\|Z_t^{(0)}\|_{X_{t-}}$ is bounded for all $\omega \in \Omega$ and $t \in [0,T]$ by a uniform constant. □

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

Proposition 5. *We have*

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

Proof. By Lemma 4, it suffices to check that

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

converges to zero for all $t \le T$ and that $I_n(t)$ is bounded by C uniformly in t and n. We also need to check that $J_n := \mathbb{E}[|g(X_T, p_T^{(n)}) - g(X_T, p_T^{(0)})|^2]$ converges to zero. By the Lipschitz property of the cost functional g and Lemma 8, we have

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

as $n \to +\infty$. To check the uniform boundedness of $I_n(t)$, we recall from Lemma 6 that

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

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 because $p_t^{(n)}$ is in the simplex $\mathcal{S}$. On the other hand, we have

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

Because 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

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

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

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

By Lemma 8, we have $(1 + \|Z_s^{(0)}\|_{X_{s-}})^2 \|p_s^{(n)} - p_s^{(0)}\|^2 \to 0$, $dt \otimes d\mathbb{P}$ 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

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

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

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

converges to zero. For a fix $w \in \Omega$ and $t \leq T$, we have

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

where we define $\kappa(s, \nu) := 1_{t \leq s \leq T} H(s, X_{s-}, Z_s^{(0)}, p_s^{(0)}, \nu)$. Clearly, κ is continuous in ν for all s. On the other hand, by inequality (26) in Lemma 6, for all $t \leq s \leq T$ and $\nu \in \mathcal{P}(A)$, we have

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

Therefore, by Lemma 12 and the boundedness of $\mathcal{P}(A)$, we conclude that the mapping $(s, \nu) \to \kappa(s, \nu)$ is bounded. It follows from the definition of the stable topology and $\eta^{(n)} \to \eta^{(0)}$ that

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

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

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

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

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

Lemma 13. *Denote by $\mathbf{W}^{(n)}$ the Doléans–Dade exponential of $\mathbf{L}^{(n)}$. Then, the Itô differential of $\log(\mathbf{W}^{(n)})$ satisfies*

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

where $\underline{O}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)})$ 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. Because $\mathbf{W}^{(n)}$ 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

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

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, X, we obtain

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

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

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

Using matrix notation and recalling the definition of $\underline{O}$ in the statement of Lemma 13, we may write

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

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 $\log(W_t)$. □

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

Proposition 6. *$\mathbb{Q}^{(n)}$ 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_{TV} between $\mathbb{Q}$ and $\mathbb{Q}'$ is

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

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

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

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

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

Using Lemma 13, we have

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

By Assumption 1, the process $t \to \int_0^t X_{s-}^* \cdot (O(s, \hat{\alpha}_s^{(0)}, p_s^{(0)}, \nu_s^{(n)}) - O(s, \hat{\alpha}_s^{(n)}, p_s^{(n)}, \nu_s^{(n)})) \cdot d\mathcal{M}_s$ is a true martingale. Therefore, it has 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 $\mathbb{E}^0[\int_0^T X_{t-}^* \cdot (\underline{O}(t, \hat{\alpha}_t^{(n)}, p_t^{(n)}, \eta_t^{(n)}) - \underline{O}(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \eta_t^{(0)})) \cdot Q^0 \cdot X_{t-} dt]$ can be dealt with in the exact same way. Using the Lipschitz property of $\hat{a}$ and $\underline{Q}$ in Lemmas 6 and 11, we obtain

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

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

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

This converges to zero by Proposition 5. For the second expectation, we notice from Lemma 12 that $\|Z_t^{(0)}\|_{X_{t-}}$ is bounded by a constant for all $(\omega, t) \in \Omega \times [0, T]$. Therefore, we have

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

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

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

This converges to zero because $\eta^{(n)}$ converges to $\eta^{(0)}$ in the stable topology and the mapping $\nu \to Q(t, \hat{\alpha}_t^{(0)}, p_t^{(0)}, \nu)$ is continuous for all t. Notice also that the integrand is bounded by a constant because q is bounded according to Assumption 1. Then, by the dominated convergence theorem, the third expectation converges to zero as well. This completes the proof. □

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

Lemma 14. *Let $(\nu_t^{(n)})_{t\le T}$ be a sequence of measurable functions from $[0,T]$ to $\mathcal{P}(A)$ such that $\int_0^T \mathcal{W}_1(\nu_t^{(n)}, \nu_t^{(0)}) \to 0$. 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 $n \ge 0$, and let $f : [0,T] \times \mathcal{P}(A) \to \mathbb{R}$ 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

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

By Lemma 9, we conclude that $\lambda^{(n)}$ converges to $\lambda^{(0)}$ for the stable topology. □

Proposition 7. $\mathcal{L}(\cdot) \times \delta_{\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}}(\cdot)$ *converges to* $\mathcal{L}(\cdot) \times \delta_{\mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}}(\cdot)$ *in* $\mathcal{R}_0$ *in the sense of the stable topology.*

Proof. By Lemma 14, we only need to show that $\int_0^T \mathcal{W}_1(\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}}, \mathbb{Q}^{(0)}_{\#\hat{\alpha}_t^{(0)}}) dt$ converges to zero. Notice that

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

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

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

which converges to zero according to the proof of Proposition 6. By Villani [33, theorem 6.16], because A is bounded and $\mathbb{Q}^{(n)}_{\#\hat{\alpha}_t^{(n)}} \in \mathcal{P}(A)$, there exists a constant C such that

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

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

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

Now, for the other term, we have

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

The Lipschitz property of $\hat{a}$ (see Lemma 6) and Proposition 5 imply that $\mathbb{E}[T \int_0^T \|\hat{\alpha}_t^{(n)} - \hat{\alpha}_t^{(0)}\|^2 dt] \to 0$. □

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

Proof of Theorem 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 3, Γ is a Polish space. By Proposition 2 and Lemma 10, $\Gamma_0 := \bar{\mathcal{P}}_0 \times \mathcal{R}_0$ is a compact and convex subset of Γ, and it is stable for the mapping Φ defined in (46). In addition, we see from Propositions 6 and 7 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 $\alpha_t^* := \hat{a}(t, X_{t-}, Z_t^*, \pi(t, \mu^*))$, 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 $\nu_t^* := \mathbb{P}^*_{\#\alpha_t^*}$. Because (μ^*, η^*) is the fixed point of the mapping Φ, we have $\eta_t^* = \delta_{\nu_t^*}$ and $\mu^* = \mathbb{P}^*$. It follows that $p_t^* = \pi(t, \mathbb{P}^*) = [\mathbb{P}^*(X_t = e_i)]_{1\le i\le m}$. By Proposition 1, we see that $\boldsymbol{\alpha}^*$ 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 $\boldsymbol{\nu}^*$. This implies that $(\boldsymbol{\alpha}^*, \mathbf{p}^*, \boldsymbol{\nu}^*)$ is a Nash equilibrium. □

5. Approximate Nash Equilibrium for Games with Finitely 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 càdlág mappings from $[0, T]$ to $E = \{e_1, \dots, e_m\}$, which are continuous on T; $t \to X_t$ is the canonical process; and $\mathbb{F} := (\mathcal{F}_t)_{t\le T}$ 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 p° 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 $\psi_t^n := diag(Q^0 \cdot X_{t-}^n) - Q^0 \cdot diag(X_{t-}^n) - diag(X_{t-}^n) \cdot Q^0$.

Throughout this section, we let Assumptions 1–3 hold. In addition, we adopt the following assumption:

Assumption 5. *The transition rate function q depends neither on the mean field of states nor on the mean field of controls.*

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 $\boldsymbol{\alpha}^n \in \mathbb{A}^N$. We define the martingale $L^{(\alpha^1,\dots,\alpha^N)}$ by

$$L_t^{(\alpha^1,\dots,\alpha^N)} := \int_0^t \sum_{n=1}^N (X_{s-}^n)^* \cdot (Q(s, \alpha_s^n) - Q^0) \cdot (\psi_s^n)^+ \cdot d\mathcal{M}_s^n, \tag{50}$$

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

$$\frac{d\mathbb{Q}^{(\alpha^1,\dots,\alpha^N)}}{d\mathbb{P}^N} = \mathcal{E}_T^{(\alpha^1,\dots,\alpha^N)}, \tag{51}$$

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:

$$p_t^N := \frac{1}{N}\left[\sum_{n=1}^N \mathbb{1}(X_t^n = e_1), \sum_{n=1}^N \mathbb{1}(X_t^n = e_2), \dots, \sum_{n=1}^N \mathbb{1}(X_t^n = e_m)\right] \in \mathcal{S}, \tag{52}$$

as well as the empirical distribution of the controls:

$$\nu(\alpha_t^1, \dots, \alpha_t^N) := \frac{1}{N}\sum_{n=1}^N \delta_{\alpha_t^n} \in \mathcal{P}(A), \tag{53}$$

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

$$J^{n,N}(\boldsymbol{\alpha}^1, \dots, \boldsymbol{\alpha}^N) := \mathbb{E}^{\mathbb{Q}^{(\alpha^1,\dots,\alpha^N)}}\left[\int_0^T f(t, X_t^n, \alpha_t^n, p_t^N, \nu(\alpha_t^1, \dots, \alpha_t^N))dt + g(X_T^n, p_T^N)\right]. \tag{54}$$

We now consider a Nash equilibrium $(\boldsymbol{\alpha}^*, \mathbf{p}^*, \boldsymbol{\nu}^*)$ of the mean field game in the sense of Definition 1. Recall that $\boldsymbol{\alpha}^*$ 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 $\hat{\boldsymbol{\alpha}}^n$ of player n by

$$\hat{\boldsymbol{\alpha}}^n(w^1, \dots, w^N) := \boldsymbol{\alpha}^*(w^n). \tag{55}$$

Clearly, $\hat{\boldsymbol{\alpha}}^n$ 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 $\hat{\boldsymbol{\alpha}}^{(N)}$ defined by

$$\hat{\boldsymbol{\alpha}}^{(N)} := \left(\hat{\boldsymbol{\alpha}}^1, \dots, \hat{\boldsymbol{\alpha}}^N\right) \tag{56}$$

is a distributed strategy profile, which means that every player's strategy is only based on the observation of the player's own path.

Definition 3 (Approximative Nash Equilibrium). Given a family of strategy profiles $(\boldsymbol{\alpha}^{(N)})_{N\geq 1}$ indexed by the number of players N, where, for each N, $\boldsymbol{\alpha}^{(N)} := (\boldsymbol{\alpha}^{1,N}, \dots, \boldsymbol{\alpha}^{N,N})$, we say that the family $(\boldsymbol{\alpha}^{(N)})_{N\geq 1}$ is an approximative Nash equilibrium for games with finitely many players if there exists a positive sequence $(\epsilon_N)_{N\geq 1}$ such that $\lim_{N\to+\infty} \epsilon_N = 0$, and for any $N \geq 1$, any individual player $n \leq N$, and any admissible strategy $\beta \in \mathbb{A}^N$ chosen by player n, we have

$$J^{n,N}\left(\boldsymbol{\alpha}^{1,N}, \dots, \boldsymbol{\alpha}^{N,N}\right) \leq J^{n,N}\left(\hat{\boldsymbol{\alpha}}^{1,N}, \dots, \hat{\boldsymbol{\alpha}}^{n-1,N}, \boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{n+1,N}, \dots, \hat{\boldsymbol{\alpha}}^{N,N}\right) + \epsilon_N. \tag{57}$$

In the rest of Section 5, we show that the sequence of strategy profiles $(\hat{\boldsymbol{\alpha}}^{(N)})_{N\geq 0}$ derived from the equilibrium strategy of mean field games defined in (55) and (56) is an approximative Nash equilibrium for games with finitely many players. To this end, we first give a result on the propagation of chaos, which compares player 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 $(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-n,N})$ for $(\hat{\boldsymbol{\alpha}}^1, \dots, \hat{\boldsymbol{\alpha}}^{n-1}, \boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{n+1}, \dots, \hat{\boldsymbol{\alpha}}^N)$, $\hat{\mathbb{Q}}^{(N)}$ for $\mathbb{Q}^{\hat{\boldsymbol{\alpha}}^{(N)}}$, $\hat{\mathbb{E}}^{(N)}$for $\mathbb{E}^{\hat{\mathbb{Q}}^{(N)}}$, and finally, $\hat{\mathcal{E}}^{(N)}$ for $\mathcal{E}^{\hat{\boldsymbol{\alpha}}^{(N)}}$. We start from the following lemmas:

Lemma 15. *There exists a sequence $(\delta_N)_{N\geq 0}$ such that $\delta_N \longrightarrow 0$ as $N \longrightarrow +\infty$ and such that, for all $N \geq 1$, $n \leq N$, and $t \leq T$, we have*

$$\max\{\hat{\mathbb{E}}^{(N)}[\mathcal{W}_1^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)],\ \ \hat{\mathbb{E}}^N[\|p_t^N - p_t^*\|^2]\} \leq \delta_N. \tag{58}$$

Proof. From $\hat{\mathbb{Q}}^{(N)} = \mathbb{Q}^{\hat{\boldsymbol{\alpha}}^{(N)}}$ and the fact that $(\boldsymbol{\alpha}^*, \mathbf{p}^*, \boldsymbol{\nu}^*)$ 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

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

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 $\mathbb{R}^l$. We have

$$\begin{aligned}
\hat{\mathbb{E}}^{(N)}[\mathcal{W}_1^2(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)] &\leq C\hat{\mathbb{E}}^{(N)}[\mathcal{W}_1(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu_t^*)] \\
&\leq C\hat{\mathbb{E}}^{(N)}\left[\mathcal{W}_1\left(\nu(\beta_t, \hat{\alpha}_t^{-1,N}), \nu\left(\hat{\alpha}_t^{(N)}\right)\right) + \mathcal{W}_1\left(\nu\left(\hat{\alpha}_t^{(N)}\right), \nu_t^*\right)\right] \\
&\leq C\left(\hat{\mathbb{E}}^{(N)}\left[\frac{1}{N}\|\beta_t - \hat{\alpha}_t^{1,N}\|\right] + \hat{\mathbb{E}}^{(N)}\left[\mathcal{W}_1\left(\nu\left(\hat{\alpha}_t^{(N)}\right), \nu_t^*\right)\right]\right) \\
&\leq C\left(\frac{1}{N} + \hat{\mathbb{E}}^{(N)}\left[\mathcal{W}_1\left(\nu\left(\hat{\alpha}_t^{(N)}\right), \nu_t^*\right)\right]\right),
\end{aligned}$$

where C is a constant only depending on $\sup_{a\in A}\|a\|$, which changes its value from line to line. Now applying Fournier and Guillin [18, theorem 1], we have

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

Combining this with the estimates previously shown, we obtain the desired result. □

Lemma 16. *There exists a constant C that only depends on the bound of the transition rate q such that, for all $N > 0$ and $\boldsymbol{\beta} \in \mathbb{A}$, we have*

$$\hat{\mathbb{E}}^{(N)}\left[\left(\frac{\mathcal{E}_T^{(\beta,\hat{\alpha}^{-1,N})}}{\hat{\mathcal{E}}_T^{(N)}}\right)^2\right] \le C. \tag{59}$$

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

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

Recall that

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

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

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

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 Ξ_t^β be the m by m matrix with zero diagonal elements and the entry on the ith row and the jth column $\frac{q(t,i,j,\beta_t) - q(t,i,j,\hat{\alpha}_t^1)}{q(t,i,j,\hat{\alpha}_t^1)}$. Then, it is clear that $\Delta W_t = e_i^* \cdot \Xi_t^\beta \cdot (e_j - e_i)$. It follows that

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

Injecting the preceding equation into the Itô decomposition of W_t, we obtain

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

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 $i = 1, \ldots, m$, it is plain to check the following equality:

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

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. In particular, we make use of Assumption 1 that the transition rate q is bounded uniformly with regard to the controls. □

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

Proposition 8. *There exists a sequence $(\epsilon_N)_{N\geq 0}$ such that $\epsilon_N \to 0$ as $N \to +\infty$ and such that, for all $N \geq 0$, $n \leq N$, and $\boldsymbol{\beta} \in \mathbb{A}$,*

$$\left| J^{n,N}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-n,N}) - \mathbb{E}^{\mathbb{Q}^{(\beta,\hat{\alpha}^{-n,N})}}\left[\int_0^T f(t, X_t^n, \beta_t, p_t^*, \nu_t^*)dt + g(X_T^n, p_T^*)\right]\right| \leq \epsilon_N. \tag{60}$$

Proof. Because of symmetry, we only need to show the claim for $n = 1$. Let $N > 0$ and $\boldsymbol{\beta} \in \mathbb{A}$. Using, successively, the Cauchy–Schwartz inequality, Assumption 2, and Lemmas 15 and 16, we have

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

where δ_N is as in Lemma 15, and C is a constant only depending on T, the Lipschitz constant of f and g, and the constant appearing in Lemma 16. This gives us the desired inequality. □

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

Theorem 2. *There exists a sequence ϵ_N converging to zero such that, for all $N > 0$, $\boldsymbol{\beta} \in \mathbb{A}$, and $n \leq N$, we have*

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

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

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

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

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

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

$$Y_t = g\big(X_T^1, p_T^*\big) + \int_t^T H\big(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*\big)ds - \int_t^T \big(Z_s^1\big)^* \cdot d\mathcal{M}_s^1. \tag{61}$$

By the optimality of the equilibrium, we know that, for all $t \in [0, T]$, $\hat{\alpha}_t^1$ minimizes the mapping $\alpha \to H(t, X_{t-}^1, Z_t^1, \alpha, p_t^*, \nu_t^*)$. Clearly, the solution of BSDE (61) is also the unique solution to the following BSDE:

$$\begin{aligned} Y_t = g\big(X_T^1, p_T^*\big) &+ \int_t^T \left[H\big(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*\big) + \sum_{n=2}^N \big(X_t^n\big)^* \cdot \big(Q\big(s, \hat{\alpha}_s^n\big) - Q^0\big) \cdot Z_s^n\right] ds \\ &- \int_t^T \sum_{n=1}^N \int_t^T \big(Z_s^n\big)^* \cdot d\mathcal{M}_s^n, \end{aligned} \tag{62}$$

with $Z_t^n = 0$ for $n = 2, \ldots, N$. Indeed, the existence and uniqueness of BSDE (62) can be checked easily by applying Proposition 9. On the other hand, by following exactly the same argument as in the proof of Lemma 7, we can show that $K^{1,N}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-1,N}) = \mathbb{E}^{\mathbb{P}^N}[Y_0^{(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-1,N})}]$, where $Y_0^{(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-1,N})}$ is the solution (at time $t = 0$) of

$$\begin{aligned} Y_t = g\big(X_T^1, p_T^*\big) &+ \int_t^T \left[H\big(s, X_{s-}^1, Z_s^1, \beta_s, p_s^*, \nu_s^*\big) + \sum_{n=2}^N \big(X_t^n\big)^* \cdot \big(Q\big(s, \hat{\alpha}_s^n\big) - Q^0\big) \cdot Z_s^n\right] ds \\ &- \int_t^T \sum_{n=1}^N \int_t^T \big(Z_s^n\big)^* \cdot d\mathcal{M}_s^n. \end{aligned} \tag{63}$$

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 $H(s, X_{s-}^1, Z_s^1, \hat{\alpha}_s^1, p_s^*, \nu_s^*) \le H(s, X_{s-}^1, Z_s^1, \beta_s, p_s^*, \nu_s^*)$. Applying the comparison principle as stated in Proposition 10 to the BSDEs (62) and (63), we conclude that $K^{1,N}(\hat{\boldsymbol{\alpha}}^{(N)}) \le K^{1,N}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-1,N})$ for all $\boldsymbol{\beta} \in \mathbb{A}$. Now, thanks to symmetry, we have $K^{n,N}(\hat{\boldsymbol{\alpha}}^{(N)}) \le K^{n,N}(\boldsymbol{\beta}, \hat{\boldsymbol{\alpha}}^{-n,N})$ for all $\boldsymbol{\beta} \in \mathbb{A}$ and $n = 1, \ldots, N$. The desired results immediately follow by applying Proposition 8. ◻

Appendix. SDEs 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 $\mathbf{X}^1, \ldots, \mathbf{X}^N$. For each $n = 1, \ldots, N$, 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 in which all the diagonal elements equal $-(m_n - 1)$ and all the off-diagonal elements equal one. We denote by $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,T]}$ the natural filtration generated by $(\mathbf{X}^1, \ldots, \mathbf{X}^N)$. 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, because of the independence of the Markov chains, for all $n_1 \neq n_2$ and $t \le T$, $\mathbb{P}$-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 $\tilde{X}_t := X_t^1 \otimes X_t^2 \otimes \cdots \otimes X_t^N$, 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 $\mathbb{R}^{m_1 \times \cdots \times m_N}$. We have the following result on the decomposition of $\tilde{\mathbf{X}}$.

Lemma A.1. *$\tilde{\mathbf{X}}$ is a continuous-time Markov chain with transition rate matrix $\tilde{Q}^0$ given by*

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

In addition, it has the canonical decomposition:

$$d\tilde{X}_t = \tilde{Q}^0 \cdot \tilde{X}_{t-} dt + d\tilde{\mathcal{M}}_t, \tag{A.2}$$

where $\tilde{\mathcal{M}}$ is an $\mathbb{F}$-martingale that satisfies

$$d\tilde{\mathcal{M}}_t = \sum_{n=1}^N \big(X_{t-}^1 \otimes \cdots \otimes X_{t-}^{n-1} \otimes I_{m_n} \otimes X_{t-}^{n+1} \otimes \cdots \otimes X_{t-}^N\big) \cdot d\mathcal{M}_t^n. \tag{A.3}$$

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 these 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 do not detail here. □

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

$$Y_t = \xi + \int_t^T F(w, s, Y_s, Z_s^1, \ldots, Z_s^n) ds - \sum_{n=1}^N \int_t^T (Z_s^n)^* \cdot d\mathcal{M}_s^n. \tag{A.4}$$

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

Proposition A.1. *Assume that there exists a constant $C > 0$ such that $dt \times \mathbb{P}$-a.s.; we have*

$$|F(w, t, y, z_1, \ldots, z_N) - F(w, t, \tilde{y}, \tilde{z}_1, \ldots, \tilde{z}_N)| \leq C\left(|y - \tilde{y}| + \sum_{n=1}^N \|z_n - \tilde{z}_n\|_{X_{t-}^n}\right). \tag{A.5}$$

Then, the BSDE (A.4) admits a solution $(\mathbf{Y}, \mathbf{Z}^1, \ldots, \mathbf{Z}^N)$ satisfying

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

Moreover, the solution is unique in the sense that, if $(\mathbf{Y}^{(1)}, \mathbf{Z}^{(1),1}, \ldots, \mathbf{Z}^{(1),N})$ and $(\mathbf{Y}^{(2)}, \mathbf{Z}^{(2),1}, \ldots, \mathbf{Z}^{(2),N})$ 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 $\tilde{Z} \in \mathbb{R}^{m_1 \times m_2}$:

$$\|\tilde{Z}\|_{\tilde{X}_{t-}}^2 = \left\|\left(I_{m_1} \otimes (X_{t-}^2)^*\right) \cdot \tilde{Z}\right\|_{X_{t-}^1}^2 + \left\|\left((X_{t-}^1)^* \otimes I_{m_2}\right) \cdot \tilde{Z}\right\|_{X_{t-}^2}^2. \tag{A.6}$$

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

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

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

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

where we have used the fact that, for any two vectors X^1, X^2, we have $\mathrm{diag}(X^1 \otimes X^2) = \mathrm{diag}(X^1) \otimes \mathrm{diag}(X^2)$. This immediately leads to the equality (A.6). 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

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

By equality (A.6) and the assumption on the regularity of F, we have

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

Applying Lemma 3, 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 $Z^1_t := (I_{m_1} \otimes (X^2_{s-})^*) \cdot \tilde{Z}_s$ and $Z^2_t := ((X^1_{s-})^* \otimes I_{m_2}) \cdot \tilde{Z}_s$. From the definition of the driver $\tilde{F}$ and $\tilde{\mathcal{M}}$ in equation (A.3), we see that

$$\begin{aligned}
Y_t &= \xi + \int_t^T F\Big(w, s, Y_s, \Big(I_{m_1} \otimes \big(X^2_{s-}\big)^*\Big) \cdot \tilde{Z}_s, \Big(\big(X^1_{s-}\big)^* \otimes I_{m_2}\Big) \cdot \tilde{Z}_s\Big)ds \\
&\quad + \int_t^T \tilde{Z}^*_s \cdot \big[\big(I_{m_1} \otimes X^2_{s-}\big) \cdot dM^1_s + \big(X^1_{s-} \otimes I_{m_2}\big) \cdot dM^2_s\big] \\
&= \xi + \int_t^T F\big(w, s, Y_s, Z^1_s, Z^2_s\big)ds + \int_t^T \big(Z^1_s\big)^* \cdot dM^1_s + \int_t^T \big(Z^2_s\big)^* \cdot dM^2_s
\end{aligned}.$$

This shows that (Y, Z^1, Z^2) is a solution to BSDE (A.4). □

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

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

$$Y_t = \xi + \int_t^T \left(\phi_s + \beta_s Y_s + \sum_{n=1}^N \big(\gamma^n_s\big)^* \cdot Z^n_s\right) ds - \sum_{n=1}^N \int_t^T \big(Z^n_s\big)^* \cdot d\mathcal{M}^n_s. \tag{A.7}$$

Assume that, for all $n = 1, \dots, N$, $t \in (0, T]$, and j such that $(e^n_j)^* \cdot Q^{0,n} \cdot X^n_{t-} > 0$, we have $1 + (\gamma^n_t)^* \cdot (\psi^n_t)^+ \cdot (e^n_j - X^n_{t-}) \ge 0$, where $(\psi^n_t)^+$ is the Moore–Penrose inverse of the matrix ψ^n_t. Then, Y is nonnegative.

Proof. As before, we treat the case for $N = 2$, for which the argument can be trivially generalized to any $N > 2$. Because γ^n and β are bounded processes and $\sum_{i=1}^{m_n}[\gamma^n_t]_i = 0$ for all $t \le T$ and $n \le N$, we easily verify that the Lipschitz condition (A.5) stated in Proposition A.1 is satisfied, and therefore, BSDE (A.7) admits a unique solution. Now, consider the following BSDE driven by $\mathcal{M}$:

$$Y_t = \xi + \int_t^T (\phi_s + \beta_s Y_s + \boldsymbol{\gamma}^*_s \cdot \boldsymbol{Z}_s)ds - \sum_{n=1}^2 \int_t^T \boldsymbol{Z}^*_s \cdot dM_s, \tag{A.8}$$

where $\boldsymbol{\gamma}_t := (\gamma_t^1 \otimes X_{t-}^2) + (X_{t-}^1 \otimes \gamma_t^2)$. It is easy to verify the BSDE (A.8) admits a unique solution (Y, Z), and following the same argument as in the proof of Proposition A.1, we verify that $(Y_t, Z_t^1, Z_t^2) := (Y_t, (I_{m_1} \otimes (X_{s-}^2)^*) \cdot \boldsymbol{Z}_s, ((X_{s-}^1)^* \otimes I_{m_2}) \cdot \boldsymbol{Z}_s)$ 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 end, we need to apply the comparison principal for the case of a single Markov chain as is stated in Lemma 5. Note that X^1 and X^2 do not jump simultaneously, and $\boldsymbol{X}_t = X_t^1 \otimes X_t^2$. For the jump of $\boldsymbol{X}$ resulting from the jump of X^1, we need to show that, for $k = 1, \ldots, m_1$,

$$1 + \boldsymbol{\gamma}_t^* \cdot \boldsymbol{\psi}_t^+ \cdot (e_k^1 \otimes X_{t-}^2 - X_{t-}^1 \otimes X_{t-}^2) \geq 0. \tag{A.9}$$

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

$$1 + (\gamma_t^1)^* \cdot (\psi_t^1)^+ \cdot (e_k^1 - e_i^1) \geq 0. \tag{A.10}$$

It can be easily verified that

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

so that we have

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

It follows that

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

where, in the last equality, we use 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

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

Combining this with the inequality (A.10), we obtain the inequality (A.9). Proceeding in a similar way, we can also show that, for $k = 1, \ldots, m_2$,

$$1 + \boldsymbol{\gamma}_t^* \cdot \boldsymbol{\psi}_t^+ \cdot (X_{t-}^1 \otimes e_k^2 - X_{t-}^1 \otimes X_{t-}^2) \geq 0.$$

Applying Lemma 5 to the BSDE (A.8), we obtain the desired result. □

References

[1] Basna R, Hilbert A, Kolokoltsov VN (2014) An epsilon-Nash equilibrium for non-linear Markov games of mean-field-type on finite spaces. *Comm. Stochastic Anal.* 8(4):449–468.
[2] Bayraktar E, Cohen A (2018) Analysis of a finite state many player game using its master equation. *SIAM J. Control Optim.* 56(5):3538–3568.
[3] Benazzoli C, Campi L, Di Persio L (2017) Mean-field games with controlled jumps. arXiv:1703.01919.
[4] Cardaliaguet P, Lehalle CA (2017) Mean field game of controls and an application to trade crowding. *Math. Financial Econom.* 12(3):335–363.
[5] Carmona R (2016) *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*, vol. 1 (SIAM, Philadelphia).
[6] Carmona R, Delarue F (2013) Probabilistic analysis of mean-field games. *SIAM J. Control Optim.* 51(4):2705–2734.

[7] Carmona R, Lacker D (2015) A probabilistic weak formulation of mean field games and applications. *Ann. Appl. Probab.* 25(3):1189–1231.
[8] Carmona R, Wang P (2016) Finite state mean field games with major and minor players. arXiv:1610.05408.
[9] Carmona R, Wang P (2017) An alternative approach to mean field game with major and minor players, and applications to herders impacts. *Appl. Math. Optim.* 76(1):5–27.
[10] Cecchin A, Fischer M (2018) Probabilistic approach to finite state mean field games. *Appl. Math. Optim.* 81:253–300.
[11] Cecchin A, Pelino G (2019) Convergence, fluctuations and large deviations for finite state mean field games via the master equation. *Stochastic Processes Appl.* 129(11):4510–4555.
[12] Cohen SN, Elliott RJ (2008) Solutions of backward stochastic differential equations on Markov chains. *Comm. Stochastic Anal.* 2(2):251–262.
[13] Cohen SN, Elliott RJ (2010) Comparisons for backward stochastic differential equations on Markov chains and related no-arbitrage conditions. *Ann. Appl. Probab.* 20(1):267–311.
[14] Doncel J, Gast N, Gaujal B (2016) Mean-field games with explicit interactions. *J. Dynam. Games* 6(3):1–19.
[15] Elie R, Mastrolia T, Possamaï D (2019) A tale of a principal and many, many agents. *Math. Oper. Res.* 44(2):440–467.
[16] Elliott RJ, Aggoun L, Moore JB (1995) *Hidden Markov Models: Estimation and Control*. Number 29 in Applications of Mathematics (Springer, New York).
[17] Filippov AF, Arscott FM (2013) *Differential Equations with Discontinuous Righthand Sides: Control Systems*, vol. 18 (Kluwer Academic Publishers).
[18] Fournier N, Guillin A (2015) On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* 162(3–4):707–738.
[19] Georganopoulos G (1967) Sur l'approximation des fonctions continues par des fonctions lipschitziennes. *Comptes Rendus Acad. Sci. Paris* 264(7):319–321.
[20] Gomes D, Voskanyan V (2016) Extended deterministic mean-field games. *SIAM J. Control Optim.* 54(2):1030–1055.
[21] Gomes DA, Mohr J, Souza RR (2013) Continuous time finite state mean field games. *Appl. Math. Optim.* 68(1):99–143.
[22] Gomes DA, Velho RM, Wolfram MT (2014) Socio-economic applications of finite state mean field games. *Philos. Trans. Roy. Soc. A*. 372(2028).
[23] Hu Y, Peng S (1997) A stability theorem of backward stochastic differential equations and its application. *Comptes Rendus Acad. Sci.Ser. I Math*. 324(9):1059–1064.
[24] Jacod J, Mémin J (1981) Sur un type de convergence intermédiaire entre la convergence en loi et la convergence en probabilité. *Séminaire de Probabilités XV*, vol. 80 (Springer), 529–546.
[25] Jacod J, Shiryaev AN (1987) Skorokhod topology and convergence of processes. *Limit Theorems for Stochastic Processes* (Grundlehren der Mathematischen Wissenschaften 288 Berlin Heidelberg), 288–347.
[26] Kolokoltsov V, Bensoussan A (2016) Mean-field-game model for botnet defense in cyber-security. *Appl. Math. Optim.* 74(3):669–692.
[27] Lasry JM, Lions PL (2007) Mean field games. *Japanese J. Math.* 2(1):229–260.
[28] Miculescu R (2000) Approximation of continuous functions by Lipschitz functions. *Real Anal. Exchange* 26(1):449–452.
[29] Protter PE (2005) Stochastic differential equations.*Stochastic Integration and Differential Equations* (Springer, New York).
[30] Sokol A, Hansen NR (2015) Exponential martingales and changes of measure for counting processes. *Stoch. Anal. Appl.* 33(5):823–843.
[31] Villani C (2008) *Optimal Transport: Old and New*, vol. 338 (Springer Science and Business Media, Berlin).