ESAIM: COCV 27 (2021) 5
https://doi.org/10.1051/cocv/2021003

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

CONTROLLED MARKOV CHAINS WITH NON-EXPONENTIAL DISCOUNTING AND DISTRIBUTION-DEPENDENT COSTS[*]

Hongwei Mei[1,**] and George Yin[2]

Abstract. This paper deals with a controlled Markov chain in continuous time with a non-exponential discounting and distribution-dependent cost functional. A definition of closed-loop equilibrium is given and its existence and uniqueness are established. Due to the time-inconsistency brought by the non-exponential discounting and distribution dependence, it is proved that the equilibrium is locally optimal in some appropriate sense. Moreover, it is shown that our problem is equivalent to a mean-field game for infinite-many symmetric players with a non-exponential discounting cost.

Mathematics Subject Classification. 93E20, 60J27, 90C40.

Received June 8, 2020. Accepted January 2, 2021.

1. Introduction

Owing to a wide range of applications in existing and emerging applications in control engineering, biology, ecology, communication systems, social networks, and finance and economics, stochastic control has been used extensively in the last a few decades. It has long been recognized that time consistency is an important issue in control theory [27]. Recently, time-inconsistent control problems have attracted increasing attention. Compared to a time-consistent problem, it is impossible to find a global optimal control or strategy for a time-inconsistent problem. For example, if the cost functional or reward is non-exponential discounting [35] or it is of mean-variance form, the problem is time-inconsistent. For specific examples of the time-inconsistency brought by the non-exponential discounting, one may refer to Example 2.1 and Example 2.2 in [35]. One of the ideas to deal with time inconsistency is to find a strategy (termed an equilibrium), which is only optimal locally. After the breakthrough using equilibria [4–6, 35], the study in time-inconsistent problems has taken root and flourished; see the survey [33] and references therein. For time-inconsistent linear-quadratic problems, we refer to [11, 12]; for time-inconsistent stochastic Volterra integral systems, see [28, 29]. Among the aforementioned works, the strategies may be divided into two categories, open-loop [5] and closed-loop [35]. In the open-loop setup, the main effort is devoted to deriving a local maximum principle (or variational principle) by using local optimality. In contrast, in the closed-loop setup, through the investigation on an N-player game, one can derive an equilibrium Hamilton-Jacobi (HJ for short) equation to determine an equilibrium strategy which verifies a

[*] The research of the second author was supported in part by the Air Force Office of Scientific Research under grant FA9550-18-1-0268.

Keywords and phrases: time inconsistency, distribution dependence, controlled Markov chain.

[1] Department of Statistics, Rice University, Houston, TX 77005, USA.
[2] Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA.

[**] Corresponding author: hongwei.mei@rice.edu

local optimality [35]. For similar closed-loop problems, time-inconsistent control problems with recursive cost functions were considered in [32] and time-inconsistent control problems with switching states were treated in [21, 31].

In this paper, we consider a controlled Markov chain in continuous-time, whose cost functional depends on a non-exponential discounting factor and the distribution of the Markov chain. Being time-inconsistent, to get a closed-loop equilibrium that verifies the appropriate local optimality, one could adopt the idea of deriving an equilibrium HJ equation using an N-player game in [35]. Nevertheless, due to the novel feature of distribution dependence in our problem, such a method is not applicable because each player cannot find his or her optimal strategy even in a short time. A modification is to consider the space of probability measures and a non-exponential discounting control problem corresponding to the distribution process (infinite dimensional). For a similar N-player game, an equilibrium HJ equation on the space of probability measures can be derived where the existence and uniqueness of the viscosity solutions have to be examined. The equilibrium we get through such a method is essentially a local optimal strategy extending the results to time-consistent case. If the problem is exponential discounting, the problem reduces to an optimal control problem in the space of probability measures. Such results for time-consistent cases have been investigated in [25, 26], where the authors proved the existence and uniqueness of the viscosity solution to an HJ equation. As observed in [35], if the cost functional is non-exponential discounting, it is required that the equilibrium HJ equation (on the space of probability measures) admits a unique classical solution instead of a viscosity solution to identify an equilibrium that verifies the local optimality. This is an even more difficult problem to deal with. In this paper, we present a different approach.

A viable alternative for dealing with the distribution dependence is based on the recently developed mean-field game theory; see [13–20] and also related works [23, 24]. In a mean-field game problem, one considers a backward HJ equation coupled with a forward transport equation on the space of probability measures. Using the fixed-point theory, one can derive a mean-field equilibrium, which is essentially a Nash equilibrium point. Applying similar idea to the control problem with non-exponential discounting and distribution-dependent cost functional, the first step is to solve a classical HJ equation from the classical optimal control theory yielding an auxiliary (fixed) process $\rho(\cdot)$ in the cost functional. A feedback control can be determined if the HJ equation is "regular enough". Then the second step is to verify that the auxiliary process $\rho(\cdot)$ coincides with the distribution law of the system using the feedback control. If the two-step verification is fulfilled, the feedback control is called a mean-field equilibrium, which is essentially a Nash equilibrium verifying certain appropriate local optimality condition. However, it needs to be mentioned that the equilibrium is not an optimal strategy in general even for the exponential-discounting cases.

In this paper, to consider continuous-time and finite-state controlled Markov chains with time-inconsistent and distribution-dependent cost functionals, we present a new definition of equilibrium with distribution dependence, which verifies a local optimality in an appropriate sense (combining the non-exponential discounting and distribution dependence together). Then we prove the existence and uniqueness of the equilibrium under some appropriate conditions. Moreover, we show that the equilibrium is equivalent to that for a mean-field game of infinite-many equivalent players with non-exponential discounting costs. Because of the non-exponential discounting feature, one has to modify the Nash optimality appropriately. Previous works on dealing with the two time inconsistencies together include [22, 30, 36]. To the best of our knowledge, most of the papers are concerned with a special class of linear stochastic differential equations with quadratic-type costs. There are few papers concerned with the theory on continuous-time controlled Markov chains with non-exponential and distribution-dependent costs. This paper aims to fill in this gap.

In view of the developments, one would question why we should consider controlled Markov chains. The answers to the question stem from the following aspects. To begin, controlled Markov chains are the foundation of controlled Markovian systems. Such a class is one of the simplest in terms of the formulation. Nevertheless, they have a broad range of applicability. Numerous systems can be formulated as controlled Markov chains and/or Markov decision processes. Not only is considering such systems necessary but has broader impacts. In addition, under suitable scaling, controlled Markov chains approximate that of controlled diffusions. In fact, in the literature, controlled Markov chains are used to build numerical schemes for stochastic control problems. Moreover, from the discussion in the previous paragraph, to take care of the regularity issues in time-inconsistent problems is important. As will be seen in this paper, treating controlled Markov chains with simple structures enables us to deal with the regularity issue effectively without complex conditions. This together with aforementioned approximation may lead to future consideration of numerical approximation of time-inconsistent controlled diffusions, which is of practical concerns.

The rest of the paper is arranged as follows. Section 2 introduces some notation together with certain preliminary results for controlled Markov chains. Section 3 formulates the main problem and proves the existence and uniqueness of the equilibrium. Section 4 establishes that the problem under consideration is equivalent to an infinite-player mean-field game with time-inconsistent costs. Finally some concluding remarks are made in Section 5.

2. Preliminaries

2.1. Notation

Let $M=\{1,\dots,m\}$ and $\mathbb{T}=[0,T]$. Denote by $\mathcal{M}$ the set of all functions defined on M equipped with the sup-norm $\|\cdot\|_{\mathcal{M}}$. Set

$$L^1([0,T],\mathcal{M}):=\left\{\theta:\mathbb{T}\times M\to\mathbb{R}\Big|\ \int_0^T\|\theta_t\|_{\mathcal{M}}\mathrm{d}t<\infty\right\},$$

$$C([0,T],\mathcal{M}):=\left\{\theta:\mathbb{T}\times M\to\mathbb{R}\big|\ \theta_t(i)\text{ is continuous w.r.t. }t\right\},$$

and

$$D([0,T],\mathcal{M}):=\left\{\theta:\mathbb{T}\times M\to\mathbb{R}\big|\ \theta_t(i)\text{ is right-continuous with left-limit w.r.t. }t\right\}.$$

Let $\mathcal{P}$ be the collection of all probability measures on M equipped with metric $\mathrm{d}(\cdot,\cdot)$ defined by

$$\mathrm{d}(\rho,\gamma):=\sum_{i=1}^m|\rho(i)-\gamma(i)|.$$

Denote by $C([0,T],\mathcal{P})$ the set of all continuous $\mathcal{P}$-valued curves on $\mathbb{T}$. Let U be the space of actions equipped with the metric w and v_0 be a fixed element in U. Let $\mathcal{U}$ be the set of maps from M to U equipped with the following metric

$$\mathrm{d}_{\mathcal{U}}(u,u')=\sup_{i\in M}w(u(i),u'(i)).$$

2.2. Controlled Markov chain

We consider a finite-state controlled Markov chain in continuous time with generator $Q_t^v=[q_t^v(i,j)]_{M\times M}$ satisfying

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}t}=\mu_tQ_t^v,\quad t\in\mathbb{T},v\in U,\tag{2.1}$$

where the the state space is M and the action space is U.

Define the admissible action set for state i as

$$U_t(i) := \{v \in U : \ q_t^v(i,\cdot) \text{ is a generator}\},$$

where $q_t^v(i,\cdot)$ is a generator in that $q_t^v(i,j) \geq 0$ if $j \neq i$ and $\sum_{j=1}^m q_t^v(i,j) = 0$. Throughout the paper, we assume that $U_t(i)$ is measurable and nonempty under the topology of U induced by the metric w.

In this paper, we focus on closed-loop Markov strategies. Let the set of all admissible closed-loop Markov strategies on time-interval $[t_0, T]$ be

$$L^1([t_0,T],\mathcal{U}) := \Big\{\pi \Big| \pi : (s,i) \in [t_0,T] \times M \to \pi_s(i) \in U_t(i) \text{ with } \int_{t_0}^T w(\pi_s(i), v_0)\mathrm{d}s < \infty\Big\}.$$

Define a subset $D([t_0,T],\mathcal{U}) \subset L^1([t_0,T],\mathcal{U})$ by

$$\begin{aligned} &D([t_0,T],\mathcal{U}) \\ &\qquad := \{\pi \in L^1([t_0,T],\mathcal{U}) : \pi_t(i) \text{ is right-continuous with left-limit w.r.t. } t \text{ for each } i \in M\}. \end{aligned}$$

We define a map $\phi_t : L^1([0,T],\mathcal{U}) \to L^1([t,T],\mathcal{U})$ by

$$(\phi_t[\pi])_s(i) := \pi_s(i), \quad \text{for } i \in M, \ t \leq s \leq T.$$

Given any $\pi \in L^1([0,T],\mathcal{U})$, we write $\mu_t^{t_0,\rho,\pi}$ as the solution of (2.1) at time t with initial data ρ at time t_0 under the strategy π. If the initial data $\rho = \delta_x$, *i.e.*, the Dirac measure concentrated on the point x, we write $\mu_t^{t_0,x,\pi} = \mu_t^{t_0,\delta_x,\pi}$. The initial time t_0 will be omitted if $t_0 = 0$. Now we pose the following assumptions to guarantee the regularity of the dynamic (2.1).

Assumption (A)

(A1) The admissible action set $U_.$ is right continuous with left limits. That is,

$$U_t(i) = \lim_{\varepsilon \to 0^+} U_{t+\varepsilon}(i) \text{ and } U_{t-}(i) := \lim_{\varepsilon \to 0^-} U_{t+\varepsilon}(i) \text{ exists.}$$

(A2) $q_.^v(i,j)$ is right continuous with left limits in the sense that for any $v_n \in U_{t+\varepsilon_n}(i)$ with limit $v \in U_t(i)$ as $\varepsilon_n \to 0^+$,

$$\lim_{n\to\infty} q_{t+\varepsilon_n}^{v_n}(i,j) = q_t^v(i,j).$$

For any $v_n \in U_{t+\varepsilon_n}(i)$ with limit $v \in U_{t-}(i)$ as $\varepsilon_n \to 0^-$,

$$q_{t-}^v(i,j) := \lim_{n\to\infty} q_{t+\varepsilon_n}^{v_n}(i,j) \text{ exists.}$$

(A3) There exist constants K_1 and $\kappa_1 > 0$ such that for any $t \in [0,T]$ and $i, j \in M$,

$$\left\{ \begin{aligned} &|q_t^v(i,j) - q_t^{v'}(i,j)| \leq \kappa_1 w(v,v'), \quad \text{for } v, v' \in U_t(i); \\ &\sup_{j\in M} \sup_{v \in U_t(j)} |q_t^v(i,j)| \leq K_1. \end{aligned} \right. \tag{2.2}$$

Remark 2.1. It follows from (A1) and (A2), $q_t^v(i,j)$ is right continuous with left limits w.r.t. t; in view of (A3), $q_t^v(i,j)$ is Lipschitz with respect to v. Thus the existence and uniqueness of the solution of (2.1) for any any $\pi \in L^1([0,T],\mathcal{U})$ are guaranteed. Note that the assumptions are much stronger than necessary for the existence and uniqueness of the solution to (2.1). The reason for the assumptions is to conclude the following corollary to be used for the control problem.

Corollary 2.2. *Under Assumption* (A)*, for any $\pi \in D([0,T],\mathcal{M})$, $q_\cdot^{\pi_\cdot(i)}(i,j)$ is right continuous with left limits for any fixed $i,j \in M$.*

To proceed, we present some regularity results about (2.1) whose proof can be found in ([34], p. 19, Thm. 2.5).

Proposition 2.3. *Under Assumption* (A)*, for any $\pi \in L^1([0,T],\mathcal{U})$, let $\mathbf{P}^\pi_{t,s}$ be the solution of*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{P}^\pi_{t,s} = \mathbf{P}^\pi_{t,s} Q_s^{\pi_s}, \quad \textit{with} \quad \mathbf{P}^\pi_{t,t} = I; \quad \textit{for } s \geq t. \tag{2.3}$$

The unique solution of (2.1) *is $\mu_t^{\rho,\pi} = \rho \mathbf{P}^\pi_{0,t}$, i.e., for each $j \in M$,*

$$\mu_t^{\rho,\pi}(j) = \sum_{j=1}^{m} p^\pi_{0,t}(i,j)\rho(i). \tag{2.4}$$

If $\pi \in L^1([0,T],\mathcal{U})$ for any $h \in \mathcal{M}$, then

$$\lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}\Big(\sum_{j=1}^{m} h(j)\mu_{t+\varepsilon}^{t,i,\pi}(j) - h(i)\Big) = \sum_{j=1}^{m} h(j) q_t^{\pi(i)}(i,j) \ \textit{holds for a.e.} \ \ t \in \mathbb{T}. \tag{2.5}$$

Moreover, if $\pi \in D([0,T],\mathcal{U})$, (2.5) *holds for all $t \in \mathbb{T}$.*

The following lemma indicates that the solution of dynamic (2.1) is Lipschitz dependent on the strategies as well.

Lemma 2.4. *Under Assumption* (A)*, given any two admissible strategies $\pi, \pi' \in L^1([0,T],\mathcal{U})$ and $\rho, \gamma \in \mathcal{P}$, we have*

$$\mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\gamma,\pi'}) \leq \mathrm{d}(\rho,\gamma) + \kappa_1 \int_0^t \mathrm{d}_{\mathcal{U}}(\pi_s, \pi'_s)\mathrm{d}s. \tag{2.6}$$

Proof. Noting that $\sum_{j=1}^m p_t^\pi(i,j) = 1$, simple calculation yields

$$\begin{aligned} \mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\gamma,\pi}) &= \mathrm{d}(\rho\mathbf{P}^\pi_t, \gamma\mathbf{P}^\pi_t) \\ &\leq \sum_{i=1}^{M}\sum_{j=1}^{M} |\rho(i) - \gamma(i)| p_t^\pi(i,j) \\ &\leq \mathrm{d}(\rho,\gamma). \end{aligned} \tag{2.7}$$

By (2.2), for any $\varepsilon > 0$, there exists a $\delta_\varepsilon > 0$ small enough (independent of t) such that when $0 \le t \le \delta_\varepsilon$,

$$\begin{aligned} \mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\rho,\pi'}) &= \mathrm{d}(\rho \mathbf{P}_t^{\pi}, \rho \mathbf{P}_t^{\pi'}) \\ &\le \sum_{j=1}^{M} \Big| \sum_{i=1}^{M} \rho(i)[p_t^{\pi}(i,j) - p_t^{\pi'}(i,j)] \Big| \\ &\le (1+\varepsilon)\kappa_1 \mathrm{d}_{\mathcal{U}}(\pi_s, \pi_s')\mathrm{d}s. \end{aligned} \tag{2.8}$$

By (2.7),

$$\begin{aligned} \mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\gamma,\pi'}) &\le \mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\rho,\pi'}) + \mathrm{d}(\mu_t^{\rho,\pi'}, \mu_t^{\gamma,\pi'}) \\ &\le (1+\varepsilon)\kappa_1 \int_0^t \mathrm{d}_{\mathcal{U}}(\pi_s, \pi_s')\mathrm{d}s + \mathrm{d}(\rho,\gamma). \end{aligned} \tag{2.9}$$

For any $t \in [\delta_\varepsilon, T]$, using (2.8) and (2.9), we have

$$\mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\gamma,\pi'}) \le \mathrm{d}(\mu_{t-\delta_\varepsilon}^{\rho,\pi}, \mu_{t-\delta_\varepsilon}^{\gamma,\pi'}) + (1+\varepsilon)\kappa_1 \int_{t-\delta_\varepsilon}^t \mathrm{d}_{\mathcal{U}}(\pi_s, \pi_s')\mathrm{d}s$$

By simple recursions and the arbitrariness of $\varepsilon > 0$, one can easily see that

$$\mathrm{d}(\mu_t^{\rho,\pi}, \mu_t^{\gamma,\pi'}) \le \mathrm{d}(\rho,\gamma) + \kappa_1 \int_0^t \mathrm{d}_{\mathcal{U}}(\pi_s, \pi_s')\mathrm{d}s.$$

The proof is complete. □

3. Time-inconsistent control problem

In this section, we introduce our time-inconsistent control problems and give the definition of an equilibrium. Moreover, we prove the existence and uniqueness under some general conditions. This section is divided into several subsections.

3.1. Definition of an equilibrium

Here, we introduce the definition of an equilibrium for our time-inconsistent problem. We begin with the non-exponential discounting and distribution-dependent cost functional.

Let the running cost and terminal cost rates be maps defined as

$$\begin{cases} f : \mathbb{T} \times \mathbb{T} \times M \times U \times \mathcal{P} \to \mathbb{R} \text{ by } (\tau, t, i, v, \rho) \to f_{\tau,t}(i, v; \rho), \\ g : \mathbb{T} \times M \times \mathcal{P} \to \mathbb{R} \text{ by } (\tau, i, \rho) \to g_\tau(i; \rho). \end{cases}$$

In this paper, we are concerned with the following distribution-dependent cost functional with non-exponential discounting factor τ,

$$\mathbf{J}_{\tau,t}(\rho,\pi) := \int_t^T \sum_{j=1}^{m} f_{\tau,s}(j, \pi_s(j); \mu_s^{t,\rho,\pi}) \mu_s^{t,\rho,\pi}(j)\mathrm{d}s + \sum_{j=1}^{m} g_\tau(j; \mu_T^{t,\rho,\pi}) \mu_T^{t,\rho,\pi}(j) \tag{3.1}$$

and the corresponding (equilibrium) value functional

$$\mathbf{V}_t(\rho,\pi) := \mathbf{J}_{t,t}(\rho,\pi). \tag{3.2}$$

$\mathbf{J}$ and $\mathbf{V}$ are called *distribution-dependent* functionals because the running cost f and the terminal cost g depend on the distribution term μ. If f and g are distribution independent, the problem reduces to the classical control problem for Markov chains with a non-exponential discounting only.

As alluded to in the introduction, it is impossible to find a global optimal strategy because of the time-inconsistency. Thus we look for the *equilibrium with distribution dependence*, which verifies some local optimality. Because our cost functional is distribution dependent, the local optimality is similar to that in [32] with a slight difference (*i.e.*, we have to fix a distribution $\nu_{\mathbb{T}}$). The definition is given as follows.

Definition 3.1. Given a $\nu_{\mathbb{T}} \in C([0,T],\mathcal{P})$, write

$$J_{\tau,t}(i,\phi_t[\pi];\nu_{\mathbb{T}}) := \int_t^T \sum_{j=1}^m f_{\tau,s}(j,\pi_s(j);\nu_s)\mu_s^{t,i,\pi}(j)\mathrm{d}s + \sum_{j=1}^m g_\tau(j;\nu_T)\mu_T^{t,i,\pi}(j)$$

and

$$V_t(i,\phi_t[\pi];\nu_{\mathbb{T}}) = J_{t,t}(i,\phi_t[\pi];\nu_{\mathbb{T}}).$$

A pair $(\rho,\pi) \in \mathcal{P}\times D([0,T],\mathcal{U})$ is called an *equilibrium with distribution dependence* if the following local-optimality holds,

$$\limsup_{\varepsilon\to 0^+} \frac{V_t(i,\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi];\mu_{\mathbb{T}}^{\rho,\pi}) - V_t(i,\phi_t[\pi];\mu_{\mathbb{T}}^{\rho,\pi})}{\varepsilon} \geq 0, \qquad \text{for any } (t,i,\pi^\varepsilon) \in \mathbb{T}\times M\times D([t,t+\varepsilon),\mathcal{U}), \tag{3.3}$$

where the perturbed strategy $\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi] \in D([t,T],\mathcal{U})$ is defined as

$$\Big(\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi]\Big)_s(i) := \begin{cases} \pi_s(i), & t+\varepsilon\leq s\leq T,\\ \pi_s^\varepsilon(i), & t\leq s< t+\varepsilon. \end{cases}$$

One can see that our local optimality is not directly related to the cost functional $\mathbf{J}$ or $\mathbf{V}$. In the definition, we define two new cost functionals J and V rather than $\mathbf{J}$ and $\mathbf{V}$. In fact, we have the following relationship between them.

$$\mathbf{J}_{\tau,t}(i,\pi) = J_{\tau,t}(i,\phi_t[\pi];\mu_{\mathbb{T}}^{i,\pi}) \text{ and } \mathbf{V}_t(i,\pi) = V_t(i,\phi_t[\pi];\mu_{\mathbb{T}}^{i,\pi}).$$

The J and V are used to indicate that the $\nu_{\mathbb{T}}$ should be given *a priori* for local optimality.

Now let us briefly introduce the idea of the local optimality (3.3). If we substitute the distribution term in the cost functional by a given $\nu_{\mathbb{T}}$, the problem becomes a non-exponential discounting but distribution independent control problem. Thus we can adopt the idea from [35] to get a strategy verifying (3.3) for such pure non-exponential discounting problem (time-inconsistent too). Since our problem is distribution-dependent, we have to verify the solution of (2.1) under the strategy achieved in the previous step coincides with the given $\nu_{\mathbb{T}}$. Therefore, we proceed with the verification in two steps.

(a) Find the unique solution $\mu_{\mathbb{T}}^{\rho,\pi}$ of the (2.1) using (ρ,π). This is to solve (2.1) using strategy π.
(b) Using $\nu_{\mathbb{T}} = \mu_{\mathbb{T}}^{\rho,\pi}$ as *a priori* given curve, find the distribution-independent equilibrium strategy, which verifies the local-optimality (3.3), which is to solve an equilibrium HJ equation.

If such a two-step recursion is fulfilled, one can see that (ρ, π) is an equilibrium with distribution dependence. With such an observation, it is intuitive to adopt the fixed-point theory to prove the existence and uniqueness of the equilibrium with distribution dependence.

To proceed, we present the following example to avoid some possible confusions on the cost functional $\mathbf{J}$.

Example 3.2. Suppose that the terminal cost in (3.1) is given by

$$g_\tau(i,\rho) = (i - \sum_{j=1}^m j\rho(j))^2.$$

Then $\sum_{i=1}^m g_\tau(i,\mu_T)\mu_T(i)$ is the variance of the distribution μ_T. In this case, if we let

$$\tilde{g}_\tau(i,\rho) = i^2 - (\sum_{j=1}^m j\rho(j))^2.$$

$\sum_{i=1}^m \tilde{g}_\tau(i,\mu_T)\mu_T(i)$ is the variance of the distribution μ_T as well. Thus g and $\tilde{g}$ give the same terminal functional in $\mathbf{V}$. While in the process of deriving the fixed-point, the terminal conditions in the Hamilton-Jacobi equation will be different (*i.e.*, J will be different from the choice of g or $\tilde{g}$ even we have the same $\mathbf{J}$). As a consequence, the equilibrium might be different as well. Therefore it is natural to question which of the equilibria is the correct one to use. In fact, in Section 4, we introduce a mean-field game with infinite-many equivalent players with non-exponential discounting cost. We can clearly identify the correct forms of f and g from the problem itself. Thus, generally speaking, the forms of f or g depend on the model used.

3.2. Distribution-independent equilibrium with A given $\nu_{\mathbb{T}}$

In this subsection, we derive the process to find a (classical) equilibrium if $\nu_{\mathbb{T}}$ in the cost functional is given and fixed. This is Step (b) from previous section. Adopting the idea from [35], one can derive an equilibrium HJ equation to deal with the time-inconsistency due to the non-exponential discounting. While in our paper, we will omit the detailed calculation for the N-player game and present the time-inconsistent HJ equation directly. Further details regarding the derivation of the equilibrium HJ equation can be found in [35]. Instead, we will prove that the strategy obtained from the equilibrium HJ equation verifies the required local-optimality. We need the following assumptions. Given $u \in \mathcal{U}$, define an operator $\mathcal{Q}_t^u : \mathcal{M} \to \mathcal{M}$ by

$$\mathcal{Q}_t^u[h](i) := \sum_{j=1}^m q_t^{u(i)}(i,j)h(j).$$

Assumption (B)

(B1) $g_\tau(i)$ is right continuous with left limits with respect to τ for each $i \in M$. There exists $\bar{f}_{\tau,t} : M \times \mathcal{P} \to \mathbb{R}$ and $\Psi_{\tau,t} : M \times U \to \mathbb{R}$ which are right-continuous with left-limits with respect to t and τ such that

$$f_{\tau,t}(i,v;\rho) = \bar{f}_{\tau,t}(i;\rho) + \Psi_{\tau,t}(i,v),$$

and

$$\sup_{i\in M} \sup_{0\le\tau,t\le T} \sup_{v\in U_t(i)} \Psi_{\tau,t}(i,v) < K_1.$$

(B2) There exist constants $K_2, K_3, K_4 \geq 0$ such that

$$\begin{cases} 0 \leq \bar{f}_{\tau,t}(i,\rho),\ g_\tau(i,\rho) \leq K_2, \\ |\bar{f}_{\tau,t}(i,\rho) - \bar{f}_{\tau,t}(i,\rho')| + |g_\tau(i,\rho) - g_\tau(i,\rho')| \leq K_3 \mathrm{d}(\rho,\rho') \\ |\Psi_{\tau,t}(i,v) - \Psi_{\tau,t}(i,v')| \leq K_4 w(v,v'). \end{cases}$$

(B3) There exists a map $\psi_t : \mathcal{M} \to \mathcal{U}$ such that for any $h \in \mathcal{M}$ such that for all $i \in M$,

$$\Psi_{t,t}(i, \psi_t[h](i)) + \mathcal{Q}_t^{\psi_t[h](i)}[h](i) = \min_{v \in U_t(i)} \Big[\Psi_{t,t}(i,v) + \mathcal{Q}_t^v[h](i) \Big]. \tag{3.4}$$

Moreover, For any $h, h' \in \mathcal{M}$,

$$\begin{cases} \mathrm{d}_{\mathcal{U}}(\psi_t[h], \psi_t[h']) \leq \kappa_2 \|h - h'\|_{\mathcal{M}}, \\ \lim\limits_{\varepsilon \to 0^+} w(\psi_{t+\varepsilon}[h](i), \psi_t[h](i)) = 0, \\ \lim\limits_{\varepsilon \to 0^-} w(\psi_{t+\varepsilon}[h](i), \psi_{t-}[h](i)) = 0, \end{cases} \tag{3.5}$$

where ψ_{t-} is similarly defined as (3.4) using $t-$ since $\Psi_{t,t}$, $\mathcal{Q}_t$ and $U_t(i)$ are all right continuous with left limits in appropriate sense.

Remark 3.3. (1) The definition of ψ_t induces a map $\psi : L^1([0,T], \mathcal{M}) \to L^1([0,T], \mathcal{U})$ by point-wise defining

$$(\psi[\theta])_t(i) = \psi_t[\theta_t](i).$$

Moreover, if $\theta \in D([0,T], \mathcal{M})$, by Assumption (B3), $\psi[\theta] \in D([0,T], \mathcal{U})$.

(2) In (B1), we assume $f_{\tau,t}$ and $g_{\tau,t}$ are right-continuous with left-limits with respect to τ. Such an assumption will be used to solve the equilibrium HJ equation. We will explain why we need such assumption later.

Example 3.4. Here we present an example where Assumption (B) is verified. Let

$$q_t^v(i,j) = \alpha_t(i,j) + \beta_t(j)v,$$

and the action set $U = [-1, 1]$. Suppose that $\Psi_t(i,v) = \frac{1}{2}v^2$, $\alpha_t(i,j) \geq 0$ if $i \neq j$ and

$$\sum_{j=1}^m \alpha_t(i,j) = \sum_{j=1}^m \beta_t(j) = 0.$$

One can easily see that $U_t(i) \neq \emptyset$ since $0 \in U_t(i)$. If $v, v' \in U_t(i)$, *i.e.*,

$$\alpha_t(i,j) + \beta_t(j)v,\ \alpha_t(i,j) + \beta_t(j)v' \geq 0 \text{ for } i \neq j,$$

then for any $\lambda \in [0,1]$

$$\alpha_t(i,j) + \beta_t(j)[\lambda v + (1 - \lambda v')] \geq 0 \text{ for } i \neq j.$$

This proves that

$$\lambda v + (1-\lambda)v' \in U_t(i), \text{ for any } \lambda \in [0,1],$$

i.e., $U_t(i)$ is a convex subset. It can also be seen that $U_t(i)$ is closed. Thus $U_t(i)$ is a closed subinterval of $[-1,1]$, and as a result,

$$\psi_t(i) = \operatorname*{argmin}_{v\in U_t(i)} \Big[\frac{v^2}{2} + v\sum_{j=1}^m h(j)\beta_t(j)\Big]$$

is well-defined uniquely by the strong convexity. Moreover, if $\beta_t(i)$ is continuous with respect to t, (3.5) holds directly.

Now we are ready to introduce the equilibrium HJ equation for our problem. Given $\nu_{\mathbb{T}} \in C([0,T],\mathcal{P})$, consider the following equilibrium HJ equation,

$$\begin{cases} \partial_t\Theta_{\tau,t}(i) + f_{\tau,t}(i,\pi_t(i);\nu_t) + \mathcal{Q}_t^{\pi_t}[\Theta_{\tau,t}](i) = 0 \\ \pi_t(i) = \psi_t[\Theta_{t,t}](i),\ \Theta_{\tau,T}(i) = g_\tau(i;\nu_T) \end{cases} \text{for any } i\in M. \tag{3.6}$$

To investigate the solution of (3.6), let us define $\mathbf{A}^\pi_{t_0,t_1}: \mathcal{M}\to\mathcal{M}$ by

$$(\mathbf{A}^\pi_{t_0,t_1}h)(i) = \sum_{j=1}^m h(j)p^\pi_{t_0,t_1}(i,j)$$

where $p^\pi_{t_0,t_1}(j)$ is defined in (2.3). Then it is not difficult to see that (3.6) is equivalent to

$$\begin{cases} \Theta_{\tau,t} = \displaystyle\int_t^T \mathbf{A}^\pi_{t,s}f_{\tau,s}(\cdot,\pi_s(\cdot);\nu_s)\mathrm{d}s + \mathbf{A}^\pi_{t,T}g_\tau(\cdot;\nu_T) \\ \pi_t(i) = \psi_t[\Theta_{t,t}](i),\ \Theta_{\tau,T}(i) = g_\tau(i). \end{cases} \tag{3.7}$$

In view of the first equality of the HJ equation, formally we would expect that the solution $\Theta\in\mathscr{M}[0,T]^2$, where

$$\mathscr{M}([0,T]^2 := \{\Theta : (\tau,t,i)\in\mathbb{T}^2\times M \to \Theta_{\tau,t}(i)\in\mathbb{R}\big|\ \Theta_{\tau,\cdot}\in C([0,T],\mathcal{M}) \text{ for each } \tau\in\mathbb{T}\}.$$

While for such Θ, it is not required $\Theta_{t,t}(i)$ being continuous in t, or even measurable. Let us look at the following example.

Example 3.5. Let $\mathbb{T}_1$ be a non-measurable subset of $[0,T]$. Let

$$\Theta_{\tau,t} = \mathbf{I}(\tau\in\mathbb{T}_1).$$

Obviously $\Theta_{\tau,t}$ is continuous with respect t for each fixed τ. While $\Theta_{t,t} = \mathbf{I}(t\in\mathbb{T}_1)$ is not a measurable function.

Therefore, to guarantee the regularity of $\Theta_{t,t}$, we have assumed $f_{\tau,t}$ to be right-continuous with left-limits with respect to τ in Assumption (B1). With the observations and the assumptions presented, we can present the existence and uniqueness of the solution to (3.6) in the following theorem.

Theorem 3.6. *Under Assumptions* (A), (B), *given any* $\nu_{\mathbb{T}}\in C([0,T],\mathcal{P})$, *there exists a solution pair* $(\Theta,\pi)\in\mathscr{M}[0,T]^2\times D([0,T],\mathcal{U})$ *for* (3.6) *with* $\Theta_{t,t}(i)$ *being right-continuous with left limits for each* $i\in M$. *As a consequence,*

$$\pi_t(i) = \psi_t(i,\Theta_{t,t})\in D([0,T],\mathcal{U}).$$

Proof. The proof is based on the fixed-point theory. Since $\nu_{\mathbb{T}}$ is given *a priori*, we omit $\nu_{\mathbb{T}}$ in the proof. Moreover, K is a generic constant whose values may be different in different appearances. Given two $\theta, \theta' \in D([0,T], \mathcal{M})$, let $(\Theta, \pi), (\Theta', \pi')$ be the solutions of

$$\begin{cases} \partial_t \Theta_{\tau,t}(i) + f_{\tau,t}(i, \pi_t(i)) + \mathcal{Q}_t^{\pi_t}[\Theta_{\tau,t}](i) = 0 \\ \pi = \psi[\theta], \ \Theta_{\tau,T}(i) = g_\tau(i) \end{cases} \text{for any } i \in M,$$

and

$$\begin{cases} \partial_t \Theta'_{\tau,t}(i) + f_{\tau,t}(i, \pi'_t(i)) + \mathcal{Q}_t^{\pi'_t}[\Theta'_{\tau,t}](i) = 0 \\ \pi' = \psi[\theta'], \ \Theta'_{\tau,T}(i) = g_\tau(i) \end{cases} \text{for any } i \in M.$$

By (3.5), it follows that

$$\int_t^T \mathrm{d}_{\mathcal{U}}(\pi_s(i), \pi'_s(i)) \mathrm{d}s \ \le \kappa_3 \int_t^T \|\theta_s - \theta'_s\|_{\mathcal{M}} \mathrm{d}s$$

and we have the contraction inequality

$$\begin{aligned} &\|\Theta_{\tau,t} - \Theta'_{\tau,t}\|_{\mathcal{M}} \\ &\le \int_t^T \|\mathbf{A}^\pi_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot)) - \mathbf{A}^{\pi'}_{t,s} f_{\tau,s}(\cdot, \pi'_s(\cdot))\|_{\mathcal{M}} \mathrm{d}s + \|\mathbf{A}^\pi_{t,T} g_\tau - \mathbf{A}^{\pi'}_{t,T} g_\tau\|_{\mathcal{M}} \\ &\le \int_t^T \|\mathbf{A}^\pi_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot)) - \mathbf{A}^\pi_{t,s} f_{\tau,s}(\cdot, \pi'_s(\cdot))\|_{\mathcal{M}} \mathrm{d}s + \int_t^T \|\mathbf{A}^\pi_{t,s} f_{\tau,s}(\cdot, \pi'_s(\cdot)) - \mathbf{A}^{\pi'}_{t,s} f_{\tau,s}(\cdot, \pi'_s(\cdot))\|_{\mathcal{M}} \mathrm{d}s \\ &\quad + \|\mathbf{A}^\pi_{t,T} g_\tau - \mathbf{A}^{\pi'}_{t,T} g_\tau\|_{\mathcal{M}} \\ &\le K \int_t^T \mathrm{d}_{\mathcal{U}}(\pi_s, \pi'_s) \mathrm{d}s \le \kappa_3 K \int_t^T \|\theta_s - \theta'_s\|_{\mathcal{M}} \mathrm{d}s. \end{aligned} \tag{3.8}$$

In the second inequality, we used (2.6) and Assumption (B).

Now we show that if $\theta \in D([0,T], \mathcal{M})$, $\Theta_{t,t}$ is also right continuous with left limits. Then we note that

$$\begin{aligned} &\|\Theta_{t+\varepsilon,t+\varepsilon} - \Theta_{t,t}\|_{\mathcal{M}} \\ &\le \|\Theta_{t+\varepsilon,t+\varepsilon} - \Theta_{t,t+\varepsilon}\|_{\mathcal{M}} + \|\Theta_{t,t+\varepsilon} - \Theta_{t,t}\|_{\mathcal{M}} \\ &\le \int_{t+\varepsilon}^T \|\mathbf{A}^\pi_{t+\varepsilon,s} f_{t+\varepsilon,s}(\cdot, \pi_s(\cdot)) - \mathbf{A}^\pi_{t+\varepsilon,s} f_{t,s}(\cdot, \pi_s(\cdot))\|_{\mathcal{M}} \mathrm{d}s + \|\mathbf{A}^\pi_{t+\varepsilon,T} g_{t+\varepsilon} - \mathbf{A}^\pi_{t+\varepsilon,T} g_t\|_{\mathcal{M}} \\ &\quad + \|\Theta_{t,t+\varepsilon} - \Theta_{t,t}\|_{\mathcal{M}}. \end{aligned}$$

Using Assumption (B1) and letting $\varepsilon \to 0^+$, since $\mathbf{A}$ is a linear operator, it follows that

$$\lim_{\varepsilon \to 0^+} \|\Theta_{t+\varepsilon,t+\varepsilon} - \Theta_{t,t}\|_{\mathcal{M}} = 0.$$

This proves that $\Theta_{t,t}$ is right continuous with respect to t. Similarly, $\Theta_{t,t}$ has a left limit with respect to t.

Now we are in a position to prove the existence and uniqueness of the solution by adopting the fixed point theory. Given $\theta^{(1)} \in D([0,T], \mathcal{M})$, let $\Theta^{(1)}$ be the solution of (3.6). Let $\theta^{(2)}_t(i) = \Theta^{(1)}_{t,t}(i)$ for each (t,i). By our claim, we know that $\theta^{(2)} \in D([0,T], \mathcal{M})$. Then let $\Theta^{(2)}$ be the solution of (3.6) using $\theta^{(2)}$. Repeating such process, one gets a sequence of functions $\{(\theta^{(n)},\Theta^{(n)}\}$. By (3.8), we have

$$\begin{aligned}&\sup_{t\le s\le T}\sup_{0\le\tau\le T}\|\Theta^{(n)}_{\tau,s}-\Theta^{(n+1)}_{\tau,s}\|_{\mathcal{M}}\\&\quad\le\kappa_3K(T-t)\int_t^T\|\theta^{(n)}_s-\theta^{(n+1)}_s\|_{\mathcal{M}}\mathrm{d}s\\&\quad\le\kappa_3K(T-t)\sup_{t\le s\le T}\sup_{0\le\tau\le T}\|\Theta^{(n-1)}_{\tau,s}-\Theta^{(n)}_{\tau,s}\|_{\mathcal{M}}.\end{aligned}$$

As a result, $T-t<\delta$ for small $\delta>0$, there exists a limit pair (θ,Θ) with $\Theta\in\mathscr{M}[T-\delta,T]^2$ and

$$\int_{T-\delta}^T|\theta_s(i)|\mathrm{d}s<\infty.$$

For the time interval $[0,T]$, one can divide $[0,T]$ into $[T,T-\delta]$, $[T-\delta,T-2\delta],\dots$. Recursively from T to 0, one can see that there exists a unique solution pair $(\theta,\Theta)\in D([0,T],\mathcal{M})\times\mathscr{M}[0,T]^2$. Moreover we have $\pi=\psi[\theta]\in D([0,T],\mathcal{U})$. □

Moreover, we have the following uniform estimate for Θ, independent of the given $\nu_{\mathbb{T}}$.

Proposition 3.7. *For any $\nu_{\mathbb{T}}\in C([0,T],\mathcal{P})$, there exists a uniform constant κ_2 (independent of $\nu_{\mathbb{T}}$) such that*

$$0\le\Theta_{\tau,t}(i)\le\kappa_2. \tag{3.9}$$

Proof. By Assumption (B2) and the representation of Θ in (3.7), (3.9) holds with some uniform constant κ_2 independent of the choice of $\nu_{\mathbb{T}}$. □

Now let us show why the strategy from (3.6) verifies the local-optimality (3.3). This illustrates why we deal with the equilibrium HJ equation (3.6) for our control problem.

Theorem 3.8. *Under Assumptions* (A), (B), *given a* $(\nu_{\mathbb{T}},\Theta,\pi)$, *where* (Θ,π) *solves* (3.6) *with the given* $\nu_{\mathbb{T}}$, *it follows that*

$$\liminf_{\varepsilon\to0^+}\frac{V_t(i,\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})-V_t(i,\phi_t[\pi];\nu_{\mathbb{T}})}{\varepsilon}\ge0$$
$$\text{for any } (t,i,\pi^\varepsilon)\in\mathbb{T}\times M\times D([t,t+\varepsilon),\mathcal{U})$$

with $\pi^\varepsilon_t=u$ *for any* $\varepsilon>0$.

Proof. Note that

$$\begin{aligned}&J_{t,t}(i,\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\\&\quad=\int_t^{t+\varepsilon}\sum_{j=1}^m f_{t,s}(j,\pi^\varepsilon_s(i);\nu_s)\mu_s^{t,i,\pi^\varepsilon}(j)\mathrm{d}s+\sum_{j=1}^m J_{t,t+\varepsilon}(j,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\mu_{t+\varepsilon}^{t,i,\pi^\varepsilon}(j).\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&J_{t,t}(i,\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})-J_{t,t}(i,\phi_t[\pi];\nu_{\mathbb{T}})\\
&\quad=\int_t^{t+\varepsilon}\sum_{j=1}^m\Big(f_{t,s}(j,\pi_s^\varepsilon(i);\nu_s)\mu_s^{t,i,\pi_\varepsilon}(j)-f_{t,s}(j,\pi_s(j);\nu_s)\mu_s^{t,i,\pi}(j)\Big)\mathrm{d}s\\
&\qquad+\sum_{j=1}^m J_{t,t+\varepsilon}(j,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\Big(\mu_{t+\varepsilon}^{t,i,\pi^\varepsilon}(j)-\mu_{t+\varepsilon}^{t,i,\pi}(j)\Big).
\end{aligned}$$

Since π^ε and π are right-continuous, it follows that

$$\begin{aligned}
&\limsup_{\varepsilon\to0^+}\frac{1}{\varepsilon}\Big[\int_t^{t+\varepsilon}\sum_{j=1}^m\Big(f_{t,s}(j,\pi_s^\varepsilon(j);\nu_s)\mu_s^{t,i,\pi^\varepsilon}(j)-f_{t,s}(j,\pi(s);\nu_s)\mu_s^{t,i,\pi}(j)\Big)\mathrm{d}s\Big]\\
&\quad=f_{t,t}(i,u(i);\nu_t)-f_{t,t}(i,\pi_t(i);\nu_t).
\end{aligned}\tag{3.10}$$

Note that $\Theta_{\tau,t}(i)=J_{\tau,t}(i,\phi_t[\pi];\nu_{\mathbb{T}})$, we have

$$\begin{aligned}
&\sum_{j=1}^m J_{t,t+\varepsilon}(j,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\mu_{t+\varepsilon}^{t,i,\pi^\varepsilon}(j)-J_{t,t+\varepsilon}(i,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\\
&=\sum_{j=1}^m\Theta_{t,t+\varepsilon}(j)[\delta_i\mathbf{P}_{t,t+\varepsilon}^{\pi^\varepsilon}](j)-\Theta_{t,t}(i)-\Big(\Theta_{t,t+\varepsilon}(i)-\Theta_{t,t}(i)\Big)\\
&=\sum_{j=1}^m\Theta_{t,t}(j)[\delta_i\mathbf{P}_{t,t+\varepsilon}^{\pi^\varepsilon}](j)-\Theta_{t,t}(i)+\sum_{j=1}^m(\Theta_{t,t}(j)-\Theta_{t,t}(j))\Big([\delta_i\mathbf{P}_{t,t+\varepsilon}^{\pi^\varepsilon}](j)-\delta_i(j)\Big).
\end{aligned}$$

Since π^ε,π are right-continuous and by Assumption (A),

$$\left\{\begin{aligned}
&\lim_{\varepsilon\to0^+}\frac{1}{\varepsilon}\Big(\sum_{j=1}^m J_{t,t+\varepsilon}(j,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\mu_{t+\varepsilon}^{t,i,\pi^\varepsilon}(j)-\Theta_{t,t}(i)\Big)=\mathcal{Q}_t^u[\Theta_{t,t}](i)\\
&\lim_{\varepsilon\to0^+}\frac{1}{\varepsilon}\Big(\sum_{j=1}^m J_{t,t+\varepsilon}(j,\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})\mu_{t+\varepsilon}^{t,i,\pi}(j)-\Theta_{t,t}(i)\Big)=\mathcal{Q}_t^{\pi_t}[\Theta_{t,t}](i).
\end{aligned}\right.\tag{3.11}$$

As a result, by (3.10), (3.11), and the definition of $\pi_t(i)=\psi_t[\Theta_{t,t}](i)$,

$$\begin{aligned}
&\liminf_{\varepsilon\to0^+}\frac{J_{t,t}(i,\pi^\varepsilon\oplus\phi_{t+\varepsilon}[\pi];\nu_{\mathbb{T}})-J_{t,t}(i,\phi_t[\pi];\nu_{\mathbb{T}})}{\varepsilon}\\
&\quad=f_{t,t}(i,u(i);\nu_t)-f_{t,t}(i,\pi_t(i);\nu_t)+\mathcal{Q}_t^u[\Theta_{t,t}](i)-\mathcal{Q}_t^{\pi_t}[\Theta_{t,t}](i)\\
&\quad\geq0.
\end{aligned}$$

The proof is complete. $\square$

3.3. Equilibrium with distribution dependence

In this section, we use the following two-step recursion to prove the existence and uniqueness of the equilibrium with distribution dependence.

Step 1: Given a $\nu_{\mathbb{T}}^{(1)}\in C([0,T],\mathcal{P})$ with initial $\nu_0^{(1)}=\rho$, let $(\Theta^{(1)},\pi^{(1)})$ be the solution pair of (3.6).

Step 2: Using the strategy $\pi^{(1)}$, let $\nu^{(2)}_{\mathbb{T}}$ be the solution of the dynamic equation (2.1) with initial $\nu^{(2)}_0 = \rho$.

By recursively repeating Step 1 and Step 2, we get a sequence of triples $(\nu^{(n)}_{\mathbb{T}}, \Theta^{(n)}, \pi^{(n)})$. Now we aim to prove that such a sequence is convergent in an appropriate sense. We need the following lemma.

Lemma 3.9. *Under Assumptions* (A), (B)*, given* $\nu_{\mathbb{T}}, \tilde{\nu}_{\mathbb{T}} \in C([0,T], \mathcal{P})$*, let* (Θ, π) *and* $(\tilde{\Theta}, \tilde{\pi})$ *be the solutions of* (3.6)*, respectively. Then there exists a constant* $\kappa_3 > 0$ *such that*

$$\sup_{0\le t\le T} \sup_{0\le \tau\le T} \|\Theta_{\tau,t} - \tilde{\Theta}_{\tau,t}\|_{\mathcal{M}} \le \kappa_3 \sup_{0\le t\le T} \mathrm{d}(\nu_t, \tilde{\nu}_t). \tag{3.12}$$

Proof. Note that

$$\begin{aligned}
&\|\Theta_{\tau,t} - \tilde{\Theta}_{\tau,t}\|_{\mathcal{M}} \\
&\quad \le \int_t^T \|\mathbf{A}^{\pi}_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot); \nu_s) - \mathbf{A}^{\tilde{\pi}}_{t,s} f_{\tau,s}(\cdot, \tilde{\pi}_s(\cdot); \tilde{\nu}_s)\|_{\mathcal{M}} \mathrm{d}s + \|\mathbf{A}^{\pi}_{t,T} g_\tau(\cdot; \nu_T) - \mathbf{A}^{\tilde{\pi}}_{t,T} g_\tau(\cdot; \tilde{\nu}_T)\|_{\mathcal{M}} \\
&\quad \le \int_t^T \|\mathbf{A}^{\pi}_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot); \nu_s) - \mathbf{A}^{\tilde{\pi}}_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot); \tilde{\nu}_s)\|_{\mathcal{M}} \mathrm{d}s \\
&\qquad + \int_t^T \|\mathbf{A}^{\tilde{\pi}}_{t,s} f_{\tau,s}(\cdot, \pi_s(\cdot); \nu_s) - \mathbf{A}^{\tilde{\pi}}_{t,s} f_{\tau,s}(\cdot, \tilde{\pi}_s(\cdot); \tilde{\nu}_s)\|_{\mathcal{M}} \mathrm{d}s \\
&\qquad + \|\mathbf{A}^{\pi}_{t,T} g_\tau(\cdot; \nu_T) - \mathbf{A}^{\tilde{\pi}}_{t,T} g_\tau(\cdot; \nu_T)\|_{\mathcal{M}} + \|\mathbf{A}^{\tilde{\pi}}_{t,T} g_\tau(\cdot; \nu_T) - \mathbf{A}^{\tilde{\pi}}_{t,T} g_\tau(\cdot; \tilde{\nu}_T)\|_{\mathcal{M}} \\
&\quad \le K(T+1) \sup_{t\le s\le T} \mathrm{d}(\nu_s, \tilde{\nu}_s) + K \int_t^T \|\Theta_{s,s} - \tilde{\Theta}_{s,s}\|_{\mathcal{M}} \mathrm{d}s.
\end{aligned}$$

Using Grownwall's inequality, one can see that there exists a $\kappa_3 > 0$

$$\sup_{0\le \tau\le T} \sup_{0\le t\le T} \|\Theta_{\tau,t} - \tilde{\Theta}_{\tau,t}\|_{\mathcal{M}} \le \kappa_3 \sup_{0\le t\le T} \mathrm{d}(\nu_t, \tilde{\nu}_t).$$

□

Now we are ready to present the main theorem of this paper.

Theorem 3.10. *Under Assumption* (A), (B)*, there exists an equilibrium with distribution dependence* $(\rho, \pi^*) \in \mathcal{P} \times D([0,T], \mathcal{U})$*. If* $\kappa_1\kappa_2\kappa_3 T < 1$*, the equilibrium is unique.*

Proof. By (3.12), we have

$$\sup_{0\le s\le T} \sup_{0\le \tau\le T} \|\Theta^{(3)}_{\tau,s} - \Theta^{(2)}_{\tau,s}\|_{\mathcal{M}} \le \kappa_3 \sup_{0\le s\le T} \mathrm{d}(\nu^{(3)}_s, \nu^{(2)}_s).$$

By (2.6) and Assumption (B), we have

$$\begin{aligned}
\sup_{0\le s\le T} \mathrm{d}(\nu^{(3)}_s, \nu^{(2)}_s)\mathrm{d}s &\le \kappa_1 \int_0^T \mathrm{d}_{\mathcal{U}}(\pi^{(2)}_s, \pi^{(1)}_s)\mathrm{d}s \\
&\le \kappa_1\kappa_2 \int_0^T \|\Theta^{(2)}_{t,t} - \Theta^{(1)}_{t,t}\|_{\mathcal{M}} \mathrm{d}s \\
&\le \kappa_1\kappa_2 T \sup_{0\le s\le T} \sup_{0\le \tau\le T} \|\Theta^{(2)}_{\tau,s} - \Theta^{(1)}_{\tau,s}\|_{\mathcal{M}}.
\end{aligned}$$

Thus we have

$$\sup_{0\le s\le T}\sup_{0\le\tau\le T}\|\Theta^{(3)}_{\tau,s}-\Theta^{(2)}_{\tau,s}\|_{\mathcal{M}}\le\kappa_1\kappa_2\kappa_3T\sup_{0\le s\le T}\sup_{0\le\tau\le T}\|\Theta^{(2)}_{\tau,s}-\Theta^{(1)}_{\tau,s}\|_{\mathcal{M}},$$

and similarly

$$\sup_{0\le s\le T}\sup_{0\le\tau\le T}\mathrm{d}(\nu^{(3)}_s,\nu^{(2)}_s)\le\kappa_1\kappa_2\kappa_3T\sup_{0\le s\le T}\sup_{0\le\tau\le T}\mathrm{d}(\nu^{(2)}_s,\nu^{(1)}_s).$$

Since $q^{\pi}_{i,j}$ is uniformly bounded by K_1, (2.4) implies that $\{\nu^{(i)}_{\mathbb{T}}\}$ is a class of equicontinuous curves valued in a compact space $\mathcal{P}$. Using the well-known Schauder's fixed point theorem, there exists a $\nu^*_{\mathbb{T}}$ which is fixed point of the two-step recursion before. Correspondingly, we can find π^* by Theorem 3.6 and (ρ,π^*) is an equilibrium with distribution dependence.

If $\kappa_1\kappa_2\kappa_3T<1$, we can easily see that $\Theta^{(n)}_{\tau,t}$ is a Cauchy sequence with limit $\Theta^*_{\tau,t}$ and $\nu^{(n)}_{\mathbb{T}}$ is a Cauchy sequence in $C([0,T],\mathcal{P})$ with limit $\nu_{\mathbb{T}}$. At the same time, we can get a pair (ρ,π^*) where

$$\theta^*_t(i):=\Theta^*_{t,t}(i)\text{ and }\pi^*=\psi[\theta^*]\in D([0,T],\mathcal{U}).$$

By Theorem 3.8, (ρ,π^*) is the equilibrium with distribution dependence. □

We have established the equilibrium for controlled Markov chains with non-exponential discounting and distribution-dependent costs. If the functions f and g in $\mathbf{J}$ given in (3.1) are independent of the distribution, such problem reduces to the classical time-inconsistent control problem (due to non-exponential discounting only) for Markov chains. Then by Theorem 3.6, one can find the equilibrium π^*, which is independent of the choice ρ. Moreover, the assumption $\kappa_1\kappa_2\kappa_3T<1$ if naturally true because $\kappa_3=0$. On the other hand, we will see that if $\mathbf{J}$ is exponential discounting but distribution-dependent, our problem coincides with a mean-field game for infinite-many symmetric players.

4. Mean-field game

In this section, we introduce an infinite-player game of which the mean-field equilibrium is equivalent to that defined in Definition 3.1.

On a complete probability measure space $\{\Omega,\mathbf{P},\mathcal{F}\}$, consider a controlled Markov chain with a finite state space M and transition probability

$$\mathbf{P}(X_{t+\Delta t}=j|X_t=i;v)=\begin{cases} q^v_t(i,j)\Delta t+o(\Delta t),\\ 1+q^v_t(i,i)\Delta t+o(\Delta t)\end{cases}$$

where v is taken in the action space U. Denote $\mathbb{N}=\{1,\dots,N\}$ and $\mathbb{T}=[0,T]$.

Let $\{X^n_{\mathbb{T}}:n\in\mathbb{N}\}$ be solution the dynamic equations of N players with the same transition with different initial data. Let the empirical measure process $\rho^{\mathbb{N},-k}_{\mathbb{T}}$ be

$$\rho^{\mathbb{N},-k}_t(i;\omega)=\frac{1}{N-1}\sum_{n\ne k}\mathbf{I}[X^n_t(\omega)=i].$$

Let the set of admissible strategies $L^1([0,T],\mathcal{U})$ be similarly defined as before and $D^{\mathbb{N}}([0,T],\mathcal{U})$ be the set of all admissible strategies for N players.

A classical mean-field game problem is to find a Nash equilibrium $\pi^{\mathbb{N}} = (\pi^1, \ldots, \pi^N) \in \mathcal{U}^{\mathbb{N}}[0,T]$ such that

$$\hat{J}_t^k(i, \phi_t[\pi^k]; \rho_{\mathbb{T}}^{\mathbb{N},-k}) \le \hat{J}_t^k(i, \phi_t[\pi^\varepsilon]; \rho_{\mathbb{T}}^{\mathbb{N},-k}) \text{ for any } \pi^\varepsilon \in L^1([0,T],\mathcal{U}) \tag{4.1}$$

and $\hat{J}_t^k$ is the cost functional for player k defined as

$$\hat{J}_t^k(i, \phi_t[\pi^k]; \rho_{\mathbb{T}}^{\mathbb{N},-k}) = \mathbb{E}\Big[\int_t^T \hat{f}_s^k(X_s^k, \pi_s^k(X_s^k); \rho_s^{\mathbb{N},-k})\mathrm{d}s + \hat{g}^k(X_T^k; \rho_T^{\mathbb{N},-k})\big|X_t^k = i; \pi^{\mathbb{N}}\Big] \tag{4.2}$$

for some appropriate $\hat{f}^k$ and $\hat{g}^k$.

Different from (4.2), we suppose that the cost function for each player is in the same form corresponding with their dynamic equations respectively with an additional non-exponential discounting factor τ in the form of

$$J_{\tau,t}^k(i, \phi_t[\pi^k]; \rho_{\mathbb{T}}^{\mathbb{N},-k}) = \mathbb{E}\Big[\int_t^T f_{\tau,s}(X_s^k, \pi_s^k(X_s^k); \rho_s^{\mathbb{N},-k})\mathrm{d}s + g_\tau(X_T^k; \rho_T^{\mathbb{N},-k})\big|X_t^k = i; \pi^{\mathbb{N}}\Big].$$

for some appropriate f and g. Here every player is assumed to make decisions according to the same non-exponential cost functional. Each player is solving a non-exponential discounting (hence time-inconsistent) control problem and hence it is impossible to find an equilibrium strategy which verifies the optimality (4.1). Thus for such a game, we aim to find an N-player strategy in the following sense.

Definition 4.1. An N-player strategy $\pi^{\mathbb{N}} = (\pi^1, \ldots, \pi^N) \in D^{\mathbb{N}}([0,T],\mathcal{U})$ is called an *equilibrium* if

$$\begin{aligned} &\liminf_{\varepsilon\to 0^+} \frac{J_{t,t}^k(i, \pi^\varepsilon \oplus \phi_{t+\varepsilon}[\pi^k]; \rho_{\mathbb{T}}^{\mathbb{N},-k}) - J_{t,t}^k(i, \phi_t[\pi^k]; \rho_{\mathbb{T}}^{\mathbb{N},-k})}{\varepsilon} \ge 0 \\ &\qquad\qquad \text{for any } k \in \mathbb{N} \text{ and } \pi^\varepsilon \in D([t, t+\varepsilon),\mathcal{U}). \end{aligned} \tag{4.3}$$

The N-player equilibrium can be understood in the following way. Consider a player ℓ, if it is assumed that the strategies of the other players are known from the equilibrium, the strategy of player ℓ finds in the time-inconsistent control problem, coincides with player ℓ's strategy determined in the equilibrium. The above equilibrium essentially indicates that with all the strategies of the other players fixed, ℓ-player's strategy is locally optimal (due to the non-exponential discounting).

Let $N \to \infty$, *i.e.*, the number of the players tends to infinity, and suppose that $\rho_0^{\mathbb{N},-k} \to \gamma$ for any k. It can be seen that all the players are equivalent in such game. Thus we conclude that every player should obey the same strategy π. By the law of large numbers, $\rho_{\mathbb{T}}^{\mathbb{N},-k}$ converges to the curve $\mu_{\mathbb{T}}^{\gamma,\pi}$ determined by

$$\frac{\mathrm{d}\mu_t}{\mathrm{d}t} = \mu_t Q_t^{\pi_t} \text{ with } \mu_0 = \gamma.$$

In this case the equilibrium is fully determined by (γ, π) since every player is equivalent in such mean-field game. It is not difficult to see that (γ, π) coincides with that in Definition 3.1. In this sense, we can call the equilibrium defined in Definition 3.1 a mean-field equilibrium too.

Moreover, if the cost functional is independent of the non-exponential factor τ, the problem reduces to the time-consistent distribution-independent case. Through the Bellman principle, one can see that the local optimality in (4.3) yields (4.1) in this case. Thus our results generalizes mean-field game problems for controlled Markov chains with non-exponential discounting costs.

5. Concluding remarks

We have developed an approach for treating non-exponential discounting and distribution-dependent control problems for continuous-time controlled Markov chains. Due to the time-inconsistency, the problems have certain unique but interesting features. We proved the existence and uniqueness of an appropriately defined equilibrium and verified the local optimality. Moreover, we found that the equilibrium being essentially equivalent to that of an infinite-player game problem with non-exponential discounting cost.

For future work, the following questions will be of interest. Can we consider similar problems if the state space is countable? Can time delay be incorporated in the setup? These deserve further thoughts and careful investigations.

References

[1] T. Björk, M. Khapko and A. Murgoci, On time-inconsistent stochastic control in continuous time. *Finance Stoch.* **21** (2017) 331–360.
[2] F. Cherbonnier, Optimal insurance for time-inconsistent agents. Preprint (2016).
[3] M. Dodd, Obesity and time-inconsistent preferences. *Obes. Res. Clin. Practice* **2** (2008) 83–89.
[4] I. Ekeland and T.A. Pirvu, Investment and consumption without commitment. *Math. Finan. Econ.* **2** (2008) 57–86.
[5] I. Ekeland and A. Lazrak, The golden rule when preferences are time inconsistent. *Math. Finan. Econ.* **4** (2010) 29–55.
[6] I. Ekeland, O. Mbodji and T.A. Pirvu, Time-consistent portfolio management. *SIAM J. Financ. Math.* **3** (2012) 1–32.
[7] T.S. Findley and F.N. Caliendo, Short horizons, time inconsistency and optimal social security. *Int. Tax & Public Finance* **16** (2009) 487–513.
[8] T.S. Findley and J.A. Feigenbaum, Quasi-hyperbolic discounting and the existence of time-inconsistent retirement. *Theor. Econ. Lett.* **3** (2013) 119–123.
[9] S.L. Green, Time inconsistency, self-control, and remembrance. *Faith Econ.* **42** (2003) 51–60.
[10] S.R. Grenadier and N. Wang, Investment under uncertainty and time-inconsistent preferences. *J. Financ. Econ.* **84** (2007) 2–39.
[11] Y. Hu, H. Jin and X.Y. Zhou, Time-inconsistent stochastic linear–quadratic control. *SIAM J. Control Optim.* **50** (2012) 1548–1572.
[12] Y. Hu, H. Jin and X.Y. Zhou, Time-inconsistent stochastic linear–quadratic control: characterization and uniqueness of equilibrium. *SIAM J. Control Optim.* **55** (2017) 1261–1279.
[13] M. Huang, R.P. Malhamé and P.E. Caines, Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6** (2006) 221–252.
[14] M. Huang, P.E. Caines and R.P. Malhamé, Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized-nash equilibria. *IEEE Trans. Autom. Control* **52** (2007) 1560–1571.
[15] M. Huang, P.E. Caines and R.P. Malhamé, The Nash Certainty Equivalence Principle and McKean-Vlasov Systems: an Invariance Principle and Entry Adaptation. *46th IEEE Conference on Decision and Control* (2007) 121–123.
[16] M. Huang, P.E. Caines and R.P. Malhamé, An invariance principle in large population stochastic dynamic games. *J. Syst. Sci. Complex.* **20** (2007) 162–172.
[17] J.-M. Lasry and P.-L. Lions, Jeux champ moyen. I. Le cas stationnaire. *C.R. Math. Acad. Sci. Paris* **343** (2006) 619–625.
[18] J.-M. Lasry and P.-L. Lions, Jeux champ moyen. II. Horizon fini et contrle optimal. *C.R. Math. Acad. Sci. Paris* **343** (2006) 679–684.
[19] J.-M. Lasry and P.-L. Lions, Large investor trading impacts on volatility. *Ann. Inst. Henri Poincaré Anal. Non Linaire* **24** (2007) 311–323.
[20] J.-M. Lasry and P.-L. Lions, Mean field games. *Jpn. J. Math.* **2** (2007) 229–260.
[21] H. Mei and J. Yong, Equilibrium strategies for time-inconsistent stochastic switching systems. *ESAIM: COCV* **25** (2019) 64.
[22] Y. Ni, J. Zhang and M. Krstic, Time-inconsistent mean-field stochastic LQ problem: open-loop time-consistent control. *IEEE Trans. Autom. Control* **63** (2018) 2771–2786.
[23] S.L. Nguyen, D.T. Nguyen and G. Yin, A stochastic maximum principle for switching diffusions using conditional mean-fields with applications to control problems. *ESAIM: COCV* **26** (2020) 69.
[24] S. Nguyen, G. Yin and T. Hoang, On laws of large numbers for systems with mean-field interactions and Markovian switching. *Stochastic Process. Appl.* **130** (2020) 262–296.
[25] H. Pham and X. Wei, Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM: COCV* **24** (2018) 437–461.
[26] H. Pham and X. Wei, Dynamic programming for optimal control of Stochastic McKean–Vlasov dynamics. *SIAM J. Control Optim.* **55** (2017) 1069–1101.
[27] R.A. Pollak, Consistent planning. *Rev. Econ. Stud.* **35** (1968) 185–199.
[28] H. Wang, J. Sun and J. Yong, Recursive Utility Processes, Dynamic Risk Measures and Quadratic Backward Stochastic Volterra Integral Equations. *Appl. Math. Optim.* (2019) 1–46.

[29] H. Wang and J. Yong, Time-inconsistent stochastic optimal control problems and back-ward stochastic volterra integral equations. Preprint arXiv:1911.04995 (2019).

[30] T. Wang, Characterizations of equilibrium controls in time inconsistent mean-field stochastic linear quadratic problems. I. Preprint arXiv:1802.01080 (2018).

[31] J. Wei, Time-inconsistent optimal control problems with regime switching. *Math. Control Rel. Fields* **7** (2017) 585–622.

[32] Q. Wei, J. Yong and Z. Yu, Time-inconsistent recrusive stochastic optimal control problems. *SIAM J. Control Optim.* **55** (2017) 4156–4201.

[33] W. Yan and J. Yong, Time-inconsistent optimal control problems and related issues. Modeling, Stochastic Control, Optimization, and Applications, edited by G. Yin., Q. Zhang. In Vol. 164 of *The IMA Volumes in Mathematics and its Applications*. Springer (2019) 533–569.

[34] G. Yin and Q. Zhang, Continuous-time Markov chains and applications: a singular perturbation approach. Vol. 37. Springer (2012).

[35] J. Yong, Time-inconsistent optimal control problems and the equilibrium HJB equation. *Math. Control Relat. Fields* **2** (2012) 271–329.

[36] J. Yong, Linear-quadratic optimal control problems for mean-field stochastic differential equations – time-consistent solutions. *Trans. AMS* **369** (2017) 5467–5523.