Deep Learning Methods for Mean Field Control Problems with Delay

Jean-Pierre Fouque [*]   Zhaoyu Zhang [†]

April 23, 2019

Abstract

We consider a general class of mean field control problems described by stochastic delayed differential equations of McKean-Vlasov type. Two numerical algorithms are provided based on deep learning techniques, one is to directly parameterizing the optimal control using neural networks, the other is based on numerically solving the McKean-Vlasov forward anticipated backward stochastic differential equation (MV-FABSDE) system. In addition, we establish the necessary and sufficient stochastic maximum principle of this class of mean field control problems with delay based on the differential calculus on function of measures, and the existence and uniqueness results are proved for the associated MV-FABSDE system under suitable conditions.

Keywords: ************************
Mathematical Subject Classification (2000): ************************

1   Introduction

Stochastic games were introduced to study the optimal behaviors of agents interacting with each other. It is used to study the topic of systemic risk in the context of finance. For example, in [6], the authors proposed a linear quadratic inter-bank borrowing and lending model, and solved explicitly for the Nash equilibrium with a finite number of players. Later, this model was extended in [5] by considering delay in the control in the state dynamic to account for the debt repayment. The authors analyzed the problem via a probabilistic approach which relies on stochastic maximum principle, as well as via an analytic approach which is build on top of an infinite dimensional dynamic programming principle.

---

[*]Department of Statistics & Applied Probability, University of California, Santa Barbara, CA 93106-3110, e-mail: fouque@pstat.ucsb.edu. Work supported by NSF grant DMS-1814091.

[†]Department of Statistics & Applied Probability, University of California, Santa Barbara, CA 93106-3110, e-mail: zhaoyu_zhang@ucsb.edu

Both mean field control and mean field games are used to characterize the asymptotic behavior of a stochastic game as the number of players grows to infinity under the assumption that all the agents behave similarly, but with different notation of equilibrium. The mean field games consist of solving a standard control problem, where the measure is fixed, and solving a fixed point problem such that the measure matches the distribution of the dynamic of a representative agent. Whereas, the mean field control problem is a nonstandard control problem in the sense that the law of state is present in the McKean-Vlasov dynamic, and optimization is performed while imposing the constrain of distribution of the state. More detail can be found in [3] and [2].

In this paper, we considered a general class of mean field control problem with delay effect in the McKean-Vlasov dynamic. We derived the adjoint process associated with the McKean-Vlasov stochastic delayed differential equation, which was a type of anticipated backward stochastic differential equation of McKean-Vlasov type due to the fact that the conditional expectation of the future of adjoint process was involved. This type of backward stochastic differential equations (BSDE) was introduced in [13], and for the general theory of BSDE, we refer [16]. We also establish a necessary and sufficient stochastic maximum principle based on differential calculus on functions of measures. In the meantime, we also prove the existence and uniqueness of the system of McKean-Vlasov forward anticipated backward stochastic differential equations (MV-FABSDE) under some suitable conditions using the method of continuation, which can be found in [12], [1] and [5]. For a comprehensive study of FBSDE theory, we refer [11].

When there was no delay effect in the dynamic, [10] proved the relation between the solution to the FBSDE and quasi-linear partial differential equation (PDE) via "Four Step Scheme". [7] and [14] explored the use of deep learning for solving high dimensional PDEs. However, the class of fully coupled MV-FABSDE considered in our paper has no explicit solution. Here, we presented one algorithm to tackle the above problem by means of deep learning techniques. Due to the non-Markovian nature of the state dynamic, we applied the long short-term memory (LSTM) network, which was able to capture the arbitrary long-term dependencies in the data sequence. It also partially solved the vanishing gradient problem in vanilla recurrent neural networks (RNNs). This was shown in [9]. The idea of our algorithm is to approximate the solution to the adjoint process and the conditional expectations of the adjoint process. The optimal control was readily obtained after the MV-FABSDE being solved. We may also emphasis that the way that we present here for numerically compute conditional expectation may have a wide range of applications, and it is simple to implement. We also present another algorithm solving the mean field control problem by directly parameterizing the optimal control. Similar idea can be found in the the policy gradient method in the regime of reinforcement learning [15]. Numerically, two algorithms we proposed in this paper yield the same result. Besides, out approaches embrace the case with no delay.

The paper is organized as follows. We start with a N player game with delay, and let num-

ber of players goes to infinity to introduce a mean field control problem in Section 2. Next, in Section 3, we mathematically formulate the feedforward neural network and LSTM network, and propose two algorithms to numerically solve the mean field control problem using deep learning techniques. One is based on directly parameterizing the control, and the other depends on numerically solving the MV-FABSDE system. The adjoint process associated with the delayed dynamic is derived, as well as the stochastic maximum principle is proved in Section 4. Finally, the uniqueness and existence solution for this class of MV-FABSDE are proved under suitable assumptions via continuation method in Section 5.

2   Formulation of The Problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $T > 0, \tau > 0$ be given constants. We consider a N-player game with delay in both state and control. The dynamic $(X_t^i)_{0 \le t \le T}$ for player $i \in \{1, \ldots, N\}$ is given by a stochastic delayed differential equation (SDDE),

$$
\begin{aligned}
dX_t^i &= b^i(t, \mathbf{X}_t, \mathbf{X}_{t-\tau}, \alpha_t^i, \alpha_{t-\tau}^i)dt + \sigma^i(t, \mathbf{X}_t, \mathbf{X}_{t-\tau}, \alpha_t^i, \alpha_{t-\tau}^i)dW_t^i, \quad t \in (0, T] \\
X_0^i &= x_0^i, \\
X_t^i &= \alpha_t^i = 0; \quad t \in [-\tau, 0),
\end{aligned}
\tag{2.1}
$$

where $\mathbf{X}_t = (X_t^1, \cdots, X_t^N)$, and where $((W_t^i)_{t \in [0,T]})_{i=1,\cdots,N}$ are N independent Brownian motions defined on the space $(\Omega, \mathcal{F}, \mathbb{P})$, $(\mathcal{F}_t)_{0 \le t \le T}$ being the natural filtration of Brownian motions.

$$
(b, \sigma) : [0, T] \times \Omega \times \mathbb{R}^N \times \mathbb{R}^N \times A \times A \to \mathbb{R} \times \mathbb{R},
$$

are continuously differentiable $\mathcal{F}_t$-measurable functions. We denote A a closed convex subset of $\mathbb{R}$, the set of actions that player i can take, and denote $\mathbb{A}$ the set of admissible control processes. For each $i \in \{1, \ldots, N\}$, A-valued measurable processes $(\alpha_t^i)_{0 \le t \le T}$ satisfy an integrability condition such that $\mathbb{E}\left[\int_{-\tau}^T |\alpha_t^i|^2 dt\right] < +\infty$.

Given an initial condition $\mathbf{x}_0 = (x_0^1, \cdots, x_0^N) \in \mathbb{R}^N$, each player would like to minimize his objective functional:

$$
J^i(\boldsymbol{\alpha}) = \mathbb{E}\left[\int_0^T f^i(t, \mathbf{X}_t, \mathbf{X}_{t-\tau}, \alpha_t^i)dt + g^i(\mathbf{X}_T)\right].
\tag{2.2}
$$

for some Borel measurable functions $f^i : [0, T] \times \Omega \times \mathbb{R}^N \times \mathbb{R}^N \times A \to \mathbb{R}$, and $g^i : \Omega \times \mathbb{R}^N \to \mathbb{R}$.

In order to study the mean-field limit of $(\mathbf{X}_t)_{t \in [0,T]}$, we assume that the system (2.1) satisfy a symmetric property, that is to say, for each player i, the other players are indistinguishable.

3

Therefore, drift b^i and volatility σ^i in (2.1) take the form of

$$(b^i, \sigma^i)(t, \mathbf{X}_t, \mathbf{X}_{t-\tau}, \alpha^i_t, \alpha^i_{t-\tau}) = (b^i, \sigma^i)(t, X^i_t, \mu^N_t, X^i_{t-\tau}, \mu^N_{t-\tau}, \alpha^i_t, \alpha^i_{t-\tau})$$

and the running cost f^i and terminal cost g^i are of the form

$$f^i(t, \mathbf{X}_t, \mathbf{X}_{t-\tau}, \alpha^i_t) = f^i(t, X^i_t, \mu^N_t, X^i_{t-\tau}, \mu^N_{t-\tau}, \alpha^i_t) \text{ and } g^i(\mathbf{X}_T) = g^i(X^i_T, \mu^N_T),$$

where we use the notation μ^N_t for the empirical distribution of $\mathbf{X} = (X^1, \cdots, X^N)$, which is defined as

$$\mu^N_t = \frac{1}{N} \sum_{j=1}^{N} \delta_{X^j_t}.$$

Next, we let the number of players N goes to $+\infty$ before we perform the optimization. According to symmetry property and the theory of propagation of chaos, the joint distribution of the N dimensional process $(\mathbf{X}_t)_{0 \leq t \leq T} = (X^1_t, \ldots, X^N_t)_{0 \leq t \leq T}$ converges to a product distribution, and the distribution of each single marginal process converges to the distribution of $(X_t)_{0 \leq t \leq T}$ of the following Mckean-Vlasov stochastic delayed differential equation (MV-SDDE). For more detail on the argument without delay, we refer [3] and [4].

$$
\begin{aligned}
dX_t &= b(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau})dt + \sigma(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau})dW_t; \quad t \in (0, T], \\
X_0 &= x_0, \\
X_t &= \alpha_t = 0; \quad t \in [-\tau, 0).
\end{aligned}
$$
(2.3)

We then optimize after taking the limit. The objective for each play of (2.2) now becomes

$$J(\alpha) = \mathbb{E}\left[\int_0^T f(X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t)dt + g(X_T, \mu_T)\right],$$
(2.4)

where we denote $\mu_t := \mathcal{L}(X_t)$ the law of X_t.

3 Solving Mean-Field Control Problem Using Deep Learning Techniques

Due to the non-Markovian structure, the above mean-field optimal control problem (2.3)-(2.4) is difficult to solve either analytically or numerically. Here we propose two algorithms together with four approaches to tackle the above problem based on deep learning techniques. We would like to use two types of neural networks, one is called the feedforward neural network, and the other one

is called Long Short-Term Memory (LSTM) network.

For a feedforward neural network, we first define the set of layers $\mathbb{M}_{d,h}^{\rho}$, for $x \in \mathbb{R}^d$, as

$$\mathbb{M}_{d,h}^{\rho} := \{M : \mathbb{R}^d \to \mathbb{R}^h | M(x) = \rho(Ax+b), A \in \mathbb{R}^{h \times d}, b \in \mathbb{R}^h\}. \tag{3.1}$$

d is called input dimension, h is known as the number of hidden neurons, $A \in \mathbb{R}^{h \times d}$ is the weight matrix, $b \in \mathbb{R}^h$ is the bias vector, and ρ is called the activation function. The following activation functions will be used in this paper, for some $x \in \mathbb{R}$,

$$\rho_{ReLU}(x) := x^+ = \max(0, x); \quad \rho_s(x) := \frac{1}{1 + e^{-x}}; \quad \rho_{\tanh}(x) := \tanh(x); \quad \rho_{Id}(x) := x.$$

Then feedforward neural network is defined as as a composition of layers, so that the set of feedforward neural network with l hidden layers we used in this paper is defined as

$$\begin{aligned} \mathbb{NN}_{d_1,d_2}^l = \{\tilde{M} : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2} | \tilde{M} = M_{l+1} \circ M_l \circ \cdots \circ M_1 \circ M_0, \\ M_0 \in \mathbb{M}_{d_1,h_1}^{\rho_{ReLU}}, M_{l+1} \in \mathbb{M}_{h_{l+1},d_2}^{\rho_{Id}}, M_i \in \mathbb{M}_{h_i,h_{i+1}}^{\rho_{ReLU}}, h_. \in \mathbb{Z}^+, i = 1, \ldots, l\}. \end{aligned} \tag{3.2}$$

The LSTM network is one of recurrent neural network(RNN) architectures, which are powerful for capturing long-range dependence of the data. It is proposed in [9], and it is designed to solve the shrinking gradient effects which basic RNN often suffers from. The LSTM network is a chain of cells. Each LSTM cell composes of a cell state, which contains information, and three gates, which regulate the flow of information. Mathematically, the rule inside tth cell follows,

$$\begin{aligned} \Gamma_{f_t} &= \rho_s(A_f x_t + U_f a_{t-1} + b_f), \\ \Gamma_{i_t} &= \rho_s(A_i x_t + U_i a_{t-1} + b_i), \\ \Gamma_{o_t} &= \rho_s(A_o x_t + U_o a_{t-1} + b_o), \\ c_t &= \Gamma_{f_t} \odot c_{t-1} + \Gamma_{i_t} \odot \rho_{\tanh}(A_c x_t + U_c a_{t-1} + b_c), \\ a_t &= \Gamma_{o_t} \odot \rho_{\tanh}(c_t), \end{aligned} \tag{3.3}$$

where the operator $\odot$ denotes the Hadamard product. $(\Gamma_{f_t}, \Gamma_{i_t}, \Gamma_{o_t}) \in \mathbb{R}^h \times \mathbb{R}^h \times \mathbb{R}^h$ represents forget gate, input gate and output gate respectively, h refers the number of hidden neurons. $x_t \in \mathbb{R}^d$ is the input vector with d features. $a_t \in \mathbb{R}^h$ is known as the output vector with initial value $a_0 = 0$, and $c_t \in \mathbb{R}^h$ is known as the cell state with initial value $c_0 = 0$. $A_. \in \mathbb{R}^{h \times d}$ are the weight matrices connecting input and hidden layers, $U_. \in \mathbb{R}^{h \times h}$ are the weight matrix connecting hidden and output layers, and $b_. \in \mathbb{R}^h$ represents bias vector. The weight matrices and bias vectors are shared through all time steps, and are going to be learned during training process by back-propagation through

time (BPTT) method, which can be implemented in Tensorflow platform. Here we define the set of LSTM network up to time t as

$$\mathbb{LSTM}_{d,h,t} = \left\{ M : (\mathbb{R}^d)^t \times \mathbb{R}^h \times \mathbb{R}^h \to \mathbb{R}^h \times \mathbb{R}^h \mid M(x_0, \ldots, x_t, a_0, c_0) = (a_t, c_t), \right.$$

$$\left. c_t = \Gamma_{f_t} \odot c_{t-1} + \Gamma_{i_t} \odot \rho_{\tanh}(A_c x_t + U_c a_{t-1} + b_c), a_t = \Gamma_{o_t} \odot \rho_{\tanh}(c_t), a_0 = c_0 = 0 \right\}, \quad (3.4)$$

where $\Gamma_{f_.}, \Gamma_{i_.}, \Gamma_{o_.}$ are defined in (3.3).

In particular, we specify the model in a linear-quadratic form, which is inspired by [5] and [8]. The objective function is defined as

$$J(\alpha) = \mathbb{E}\left[\int_0^T \left(\frac{1}{2}\alpha_t^2 + \frac{1}{2}(X_t - m_t)^2 \right) dt + \frac{c}{2}(X_T - m_T)^2 \right], \quad (3.5)$$

subject to

$$\begin{aligned} dX_t &= (\alpha_t - \alpha_{t-\tau})dt + \sigma dW_t, \quad t \in [0, T] \\ X_0 &= x_0 \\ X_t = \alpha_t &= 0, \quad t \in [-\tau, 0) \end{aligned} \quad (3.6)$$

where $\sigma, c > 0$ are given constants, and $m_t := \int_{\mathbb{R}} x d\mu_t(x)$ denotes the mean of X at time t. In the following subsections, we solve the above problem numerically using two algorithms together with four approaches. The first two approaches are to directly approximate the control by either a LSTM network or a feedforward neural network, and minimize the objective (3.5) using stochastic gradient descent algorithm. The third and fourth approaches are to introduce the adjoint process associated with (3.6), and approximate the adjoint process, conditional expectation of adjoint process using neural networks.

3.1 Approximating the optimal control using neural networks

We first set $\Delta t = T/N = \tau/D$ for some positive integer N. The time discretization becomes

$$-\tau = t_{-D} \leq t_{-D+1} \leq t_0 = 0 = t_0 \leq t_1 \leq \cdots \leq t_N \leq T,$$

for $t_i - t_{i-1} = \Delta t$, for $i \in \{-D, -D+1, \cdots, 0, \cdots, N-1, N\}$. The discretizing the SDDE (3.6) according to Euler–Maruyama scheme now reads

$$X_{t_{i+1}} = X_{t_i} + (\alpha_{t_i} - \alpha_{t_{i-D}})\Delta t + \sigma\sqrt{\Delta t}\Delta W_{t_i}, \text{ for } i \in \{0, \cdots, N-1\} \quad (3.7)$$

6

where $(\Delta W_{t_i})_{0 \leq i \leq N}$ are independent, normal distributed sequence of random variables with mean 0 and variance 1.

First, from the definition of open loop control, and due to non-Markovian feature of (3.6), the open-loop optimal control is a function of the path of the Brownian motions up to time t, i.e., $\alpha(t, (W_s)_{0 \leq s \leq t})$. We are able to describe this dependency by a LSTM network by parametrizing the control as a function of current time and the discretized path of Brownian motion, i.e.,

$$
\begin{aligned}
(a_{t_i}, c_{t_i}) = &\varphi^1(t_i, (W_s)_{t_0 \leq s \leq t_i} | \Phi^1) \\
&\text{for } \varphi^1 \in \mathbb{LSTM}_{2, h_1, t} \text{ and } \Phi^1 = (A_f, A_i, A_o, A_c, U_f, U_i, U_o, U_c, b_f, b_i, b_o, b_c),
\end{aligned}
$$
$$
\alpha(t_i, (W_s)_{t_0 \leq s \leq t_i}) \approx \psi^1(a_{t_i} | \Psi^1) \text{ for } \psi^1 \in \mathbb{M}_{h_1, 1}^{Id} \text{ and } \Psi^1 = (A, b),
$$
$$(3.8)$$

for some $h_1 \in \mathbb{Z}^+$. We remark that the last dense layer is used to match the desired output dimension.

The second approach is again directly approximate the control but with a feedforward neural network. Due to the special structure of our model, where the mean of dynamic in (3.6) is constant, the mean field control problem coincides with the mean field game problem. In [8], authors solved the associated mean field game problem using infinite dimensional PDE approach, and found that the optimal control is a function of current state and the past of control. Therefore, the feedforward neural network with l layers, which we use to approximate the optimal control, is defined as, for $\Psi^2 = (W, b)$

$$
\begin{aligned}
\alpha_{t_i}(X_{t_i}, (\alpha_s)_{t_i - D \leq s < t_i}) \approx &\psi^2(X_{t_i}, (\alpha_s)_{t_i - D \leq s \leq t_i} | \Psi^2) \\
&\text{for } \psi^2 \in \mathbb{NN}_{D+1, 1}^l, \ \Psi^2 = (A_0, b_0, \ldots, A_{l+1}, b_{l+1}).
\end{aligned}
$$
$$(3.9)$$

From Monte Carlo algorithm, and trapezoidal rule, the objective function (3.5) now becomes

$$
\begin{aligned}
J = \frac{1}{M} \sum_{j=1}^{M} \Bigg[&\left(\frac{1}{2}(\alpha_{t_0}^{(j)})^2 + \frac{1}{2}(X_{t_0}^{(j)} - \bar{X}_{t_0})^2 + \sum_{i=1}^{N-1} \left((\alpha_{t_i}^{(j)})^2 + (X_{t_i}^{(j)} - \bar{X}_{t_i})^2 \right) \right. \\
&\left. + \frac{1}{2}(\alpha_{t_N}^{(j)})^2 + \frac{1}{2}(X_{t_N}^{(j)} - \bar{X}_{t_N})^2 \right) \frac{\Delta t}{2} + \frac{c}{2}(X_{t_N}^{(j)} - \bar{X}_{t_N})^2 \Bigg], \quad (3.10)
\end{aligned}
$$

where M denotes the number of realizations and $\bar{X} := \frac{1}{M} \sum_{j=1}^{M} X^{(j)}$. After plugging in the neural network either given by (3.8) or (3.9), the optimization problem becomes to find the best set of parameters either (Φ^1, Ψ^1) or Ψ^2 such that the objective $J(\Phi^1, \Psi^1)$ or $J(\Psi^2)$ is minimized with respect to those parameters.

The algorithm works as follows,

7

Algorithm 1: Algorithms for solving mean field control problem with delay by directly approximating the optimal control using neural networks

---

Initialization of parameters $\Theta_0 = (\Phi^1, \Psi^1)$ for approach 1 (3.8) or $\Theta_0 = (\Psi^2)$ for approach 2 (3.9);

for *each epoch $e = 1, 2, \ldots$* **do**

     • Generate $\Delta W \in \mathbb{R}^{M \times N}$ for $\Delta W_{t_i}^{(j)} := W_{ji} \sim N(0, 1), j \in \{1, \ldots, M\}$ and $i \in \{1, \ldots, N\}$;

     • $\alpha_{t_i}^{(j)} = 0$ for $i = \{-D, \ldots, -1\}$;

     • $X_0^{(j)} = \bar{X}_0 = x_0, \alpha_0^{(j)} \approx \varphi_0^{(j)}(\Theta_e)$ for some network φ given by (3.8) or by (3.9) at t_0 with proper inputs;

     • $J = \frac{1}{M} \sum_{j=1}^{M} \frac{1}{2} (\varphi_0^{(j)})^2 \frac{\Delta t}{2}$;

     for *($i = 0, \ldots, N-1$)* **do**

         • $X_{t_{i+1}}^{(j)} = X_{t_i}^{(j)} + (\alpha_{t_i}^{(j)} - \alpha_{t_{i-D}}^{(j)}) \Delta t + \sigma \sqrt{\Delta t} \Delta W_{t_i}^{(j)}$;

         • $\bar{X}_{t_{i+1}} = \frac{1}{M} \sum_{j=1}^{M} X_{t_{i+1}}^{(j)}$;

         • $\alpha_{t_{i+1}}^{(j)} \approx \varphi_{t_{i+1}}^{(j)}(\Theta_e)$ is given by either (3.8) or (3.9) at t_{i+1} ;

         if *($j = N-1$)* **then**

             • $J + = \frac{1}{M} \sum_{j=1}^{M} \left(\frac{1}{2} (\varphi_{t_{i+1}}^{(j)})^2 + \frac{1}{2} (X_{t_{i+1}}^{(j)} - \bar{X}_{t_{i+1}})^2 \right) \frac{\Delta t}{2}$;

         else

             • $J + = \frac{1}{M} \sum_{j=1}^{M} \left((\varphi_{t_{i+1}}^{(j)})^2 + (X_{t_{i+1}}^{(j)} - \bar{X}_{t_{i+1}})^2 \right) \frac{\Delta t}{2}$;

         end

     end

     • $J + = \frac{1}{M} \sum_{j=1}^{M} \frac{c}{2} (X_{t_N}^{(j)} - \bar{X}_{t_N})^2$;

     • Compute the gradient $\nabla J(\Theta)$ by backpropagation through time;

     • Update $\Theta_{e+1} = \Theta_e + \eta \frac{1}{M} \sum_{j=1}^{M} \nabla J(\Theta_e)$, according to stochastic gradient descent algorithm, for some learning rate $\eta > 0$ small ;

     • Stop if $J(\Theta)$ converges, or $|\nabla J(\Theta)| < \delta$ for some threshold δ, and return Θ;

end

In the following graphics, we choose $x_0 = 0, c = 1, \sigma = 1, T = 10, \tau = 4, \Delta t = 0.1, M = 2000$. For a specific representative path, the underlying Brownian motion paths are the same for different approaches. Figure 3.1 compares one representative optimal trajectory of the dynamic and the control, and they coincide. Figure 3.2s plot the sample average of optimal trajectory of the dynamic and the control, which are trajectories of 0. This is the same as the theoretical mean.
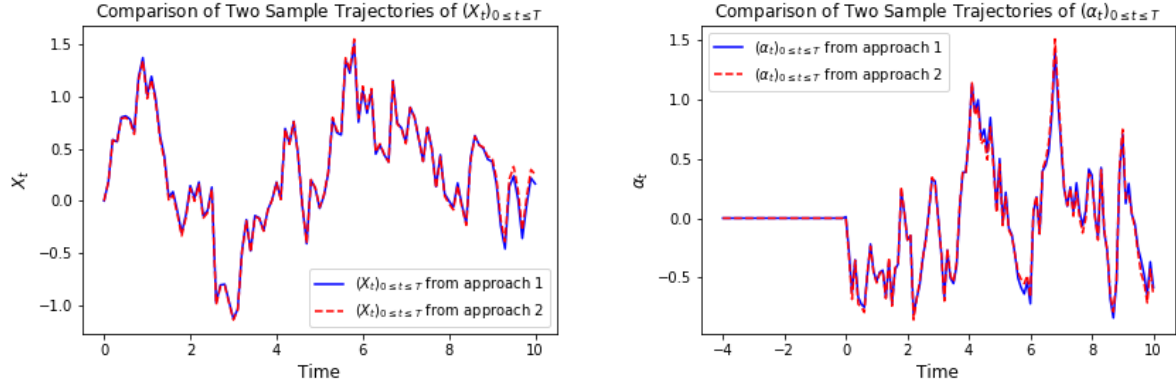
Figure 3.1: On the left, we compare one representative optimal trajectory of $(X_t)_{0 \leq t \leq T}$. The plot on the right show the comparison of one representative optimal trajectory of $(\alpha_t)_{0 \leq t \leq T}$ between approach 1 and approach 2.
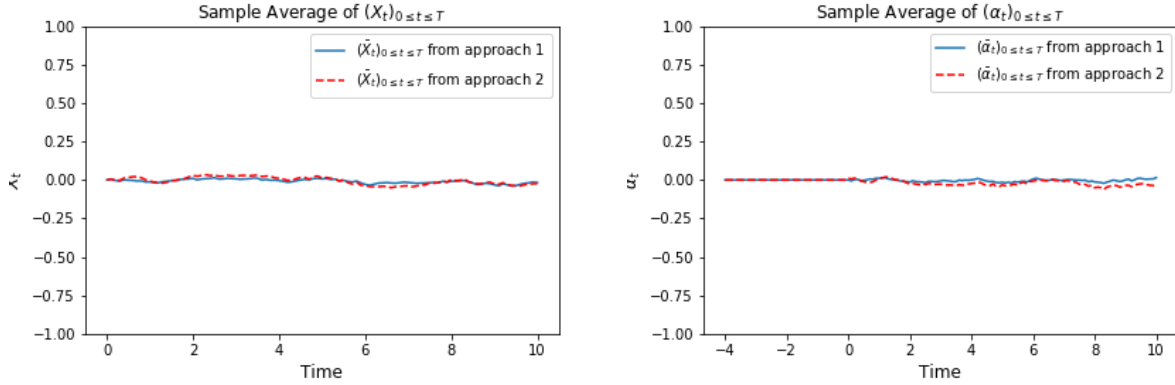

Figure 3.2: On the left, we compare the sample mean of optimal trajectories of $(X_t)_{0 \leq t \leq T}$. The plot on the right show the comparison of sample mean of trajectories of optimal control $(\alpha_t)_{0 \leq t \leq T}$ between approach 1 and approach 2.

3.2 Approximating the adjoint process using neural networks

The third and fourth approaches are based on numerically solving the MV-FABSDE system using LSTM network and feedforward neural networks. From Section 4, we are able to write the backward stochastic differential equation associated to (3.6) as,

$$dY_t = -\epsilon(X_t - m_t)dt + Z_t dW_t, \ t \in [0, T] \tag{3.11}$$

with terminal condition $Y_T = c(X_T - m_T)$, and $Y_s = 0$ for $s \in (T, T + \tau]$. The optimal control $(\alpha_t)_{0 \leq t \leq T}$ can be obtained in terms of the adjoint process Y_t from the maximum principle, and it is given by

$$\alpha_t = -Y_t + \mathbb{E}[Y_{t+\tau}|\mathcal{F}_t].$$

9

From the Euler-Maruyama method, the discretized version of (3.6) and (3.11) now reads,

$$X_{t_{i+1}} = X_{t_i} + (\alpha_{t_i} - \alpha_{t_{i-D}})\Delta t + \sigma\sqrt{\Delta t}\Delta W_{t_i}, \text{where } \Delta W_{t_i} \sim N(0,1). \quad (3.12)$$

$$\tilde{Y}_{t_{i+1}} = Y_{t_i} - \epsilon(X_{t_i} - \bar{X}_{t_i})\Delta t + Z_{t_i}\Delta W_{t_i}, \quad (3.13)$$

where we use the sample average $\bar{X}. = \frac{1}{M}\sum_{j=1}^{M} X.^{(j)}$ to approximate the expectation of $X.$. In order to solve the the above MV - FABSDE system, we need to approximate $(Y_{t_i}, \mathbb{E}[Y_{t_{i+D}}|\mathcal{F}_{t_i}], Z_{t_i})_{0 \le t_i \le t_N}$.

The third approach consists of approximating $(Y_{t_i}, \mathbb{E}[Y_{t_{i+D}}|\mathcal{F}_{t_i}], Z_{t_i})_{0 \le t_i \le t_N}$ using three LSTM networks as functions of current time and the discretized path of Brownian motions respectively, i.e.,

$$(a_{t_i}^Y, c_{t_i}^Y) = \varphi^Y(t_i, (W_s)_{t_0 \le s \le t_i}|\Phi^Y)$$
$$\text{for } \varphi^Y \in \mathbb{LSTM}_{2,h_Y,t_i} \text{ and } \Phi^Y = (A_f^Y, A_i^Y, A_o^Y, A_c^Y, U_f^Y, U_i^Y, U_o^Y, U_c^Y, b_f^Y, b_i^Y, b_o^Y, b_c^Y),$$
$$Y_{t_i} \approx \psi^Y(a_{t_i}^Y|\Psi^Y) \text{ for } \psi^Y \in \mathbb{M}_{h_Y,1}^{Id} \text{ and } \Psi^Y = (A^Y, b^Y),$$

$$(a_{t_i}^{EY}, c_{t_i}^{EY}) = \varphi^{EY}(t_i, (W_s)_{t_0 \le s \le t_i}|\Phi^{EY})$$
$$\text{for } \varphi^{EY} \in \mathbb{LSTM}_{2,h_{EY},t_i}$$
$$\text{and } \Phi^{EY} = (A_f^{EY}, A_i^{EY}, A_o^{EY}, A_c^{EY}, U_f^{EY}, U_i^{EY}, U_o^{EY}, U_c^{EY}, b_f^{EY}, b_i^{EY}, b_o^{EY}, b_c^{EY}),$$
$$E[Y_{t_{i+D}}|\mathcal{F}_{t_i}] \approx \psi^{EY}(a_{t_i}^{EY}|\Psi^{EY}) \text{ for } \psi^{EY} \in \mathbb{M}_{h_{EY},1}^{Id} \text{ and } \Psi^{EY} = (A^{EY}, b^{EY}),$$

$$(a_{t_i}^Z, c_{t_i}^Z) = \varphi^Z(t_i, (W_s)_{t_0 \le s \le t_i}|\Phi^Z)$$
$$\text{for } \varphi^Z \in \mathbb{LSTM}_{2,h_Z,t_i} \text{ and } \Phi^Z = (A_f^Z, A_i^Z, A_o^Z, A_c^Z, U_f^Z, U_i^Z, U_o^Z, U_c^Z, b_f^Z, b_i^Z, b_o^Z, b_c^Z),$$
$$Z_{t_i} \approx \psi^Z(a_{t_i}^Z|\Psi^Z) \text{ for } \psi^Z \in \mathbb{M}_{h_Z,1}^{Id} \text{ and } \Psi^Z = (A^Z, b^Z),$$

$$(3.14)$$

for some $h_Y, h_{EY}, h_Z \in \mathbb{Z}^+$. Again, the last dense layers are used to match the desired output dimension.

Since approach 3 consist three neural networks with large number of parameters, which is hard to train in general, we would like to make the following simplification in approach 4 for approximating $(Y_t, \mathbb{E}[Y_{t+\tau}|\mathcal{F}_t], Z_t)_{0 \le t \le T}$ via combination of one LSTM network and three feedforward

neural networks. Specifically,

$$
\begin{aligned}
(a_{t_i}, c_{t_i}) =& \varphi(t_i, (W_s)_{t_0 \leq s \leq t_i} | \Phi) \\
& \text{for } \varphi \in \mathbb{LSTM}_{2,h,t_i} \text{ and } \Phi = (A_f, A_i, A_o, A_c, U_f, U_i, U_o, U_c, b_f, b_i, b_o, b_c), \\
Y_{t_i} \approx& \psi^Y(a_{t_i} | \Psi^Y) \text{ for } \psi^Y \in \mathbb{NN}_{h,1}^l \text{ and } \Psi^Y = (A_0^Y, b_0^Y, \ldots, A_{l+1}^Y, b_{l+1}^Y), \\
E[Y_{t_{i+D}} | \mathcal{F}_{t_i}] \approx& \psi^{EY}(a_{t_i} | \Psi^{EY}) \text{ for } \psi^{EY} \in \mathbb{NN}_{h,1}^l \text{ and } \Psi^{EY} = (A_0^{EY}, b_0^{EY}, \ldots, A_{l+1}^{EY}, b_{l+1}^{EY}), \\
Z_{t_i} \approx& \psi^Z(a_{t_i} | \Psi^Z) \text{ for } \psi^Z \in \mathbb{NN}_{h_Y,1} \text{ and } \Psi^Z = (A_0^Z, b_0^Z, \ldots, A_{l+1}^Z, b_{l+1}^Z).
\end{aligned} \tag{3.15}
$$

In words, the algorithm works as follows. We first initialize the parameters $(\Theta^Y, \Theta^{EY}, \Theta^Z) = ((\Phi^Y, \Psi^Y), (\Phi^{EY}, \Psi^{EY}), (\Phi^Z, \Psi^Z))$ in either (3.14) or $(\Theta^Y, \Theta^{EY}, \Theta^Z) = ((\Phi, \Psi^Y), \Psi^{EY}), \Psi^Z)$ in (3.15). At time 0, $X_0 = 0$, $(Y_0, \mathbb{E}[Y_{t_D} | \mathcal{F}_0], Z_0) \approx (\varphi_0^Y(\Theta^Y), \varphi_0^{EY}(\Theta^{EY}), \varphi_0^Z(\Theta^Z))$ for some network $(\varphi^Y, \varphi^{EY}, \varphi^Z)$ given by either in (3.14) or in (3.15), and $\alpha_0 = -Y_0 + \mathbb{E}[Y_{t_D} | \mathcal{F}_0]$. Next, we update $X_{t_{i+1}}$ and $Y_{t_{i+1}}$ according to (3.12), and the solution to the backward equation at t_{i+1} is denoted by $\tilde{Y}_{t_{i+1}}$. In the mean time, $Y_{t_{i+1}}$ is also approximated by a neural network. In such case, we refer $\tilde{Y}_.$ as the label, and $Y_.$ given by the neural network as the prediction. We would like to minimize the mean square error between these two. At time T, Y_{t_N} is also supposed to match $c(X_{t_N} - \bar{X}_{t_N})$, from the terminal condition of (3.11). In addition, the conditional expectation $\mathbb{E}[Y_{t_{i+D}} | \mathcal{F}_{t_i}]$ given by a neural network should be the best predictor of $\tilde{Y}_{t_{i+D}}$. This implies that we would like to find the set of parameters Θ^{EY} such that $\mathbb{E}[(\tilde{Y}_{t_{i+D}} - \varphi_{t_i}^{EY}(\Theta^{EY})]$ is minimized for all $t_i \in \{t_0, \ldots, t_{N-D}\}$. Therefore, for M samples, we would like to minimize two objective functions L_1 and L_2 defined as

$$
\begin{aligned}
L_1(\Theta^Y, \Theta^Z) =& \frac{1}{M} \left[\sum_{j=1}^{M} \sum_{i=1}^{N} (\varphi_{t_i}^{Y,(j)} - \tilde{Y}_{t_i}^{(j)})^2 + \sum_{j=1}^{M} \left(\varphi_{t_N}^{Y,(j)} - cX_{t_N}^{(j)} \right)^2 \right], \\
L_2(\Theta^Z) =& \frac{1}{M} \sum_{j=1}^{M} \sum_{i=0}^{N-D} \left(\varphi_{t_i}^{EY,(j)} - \tilde{Y}_{t_{i+D}}^{(j)} \right)^2.
\end{aligned} \tag{3.16}
$$

The algorithm works as follows,

Algorithm 2: Algorithms for solving mean field control problem with delay according to MV-FABSDE

Initialization of parameters $(\Theta^Y, \Theta^{EY}, \Theta^Z)$ for approach 3 as in (3.14) or approach 4 as in (3.15);

for *each epoch* $e = 1, 2, \ldots$ **do**

    • Generate $\Delta W \in \mathbb{R}^{M \times N}$ for $\Delta W_{t_i}^{(j)} := W_{ji} \sim N(0,1)$, $j \in \{1, \ldots, M\}$ and $i \in \{1, \ldots, N\}$;

    • $\alpha_{t_i}^{(j)} = 0$ for $i \in \{-D, \ldots, -1\}$;

    • $X_0^{(j)} = \bar{X}_0 = x_0$;
$\left(Y_0^{(j)}, \mathbb{E}[Y_{t_D}^{(j)}|\mathcal{F}_0], Z_0^{(j)} \right) \approx \left(\varphi_0^{Y,(j)}(\Theta_e^Y), \varphi_0^{EY,(j)}(\Theta_e^{EY}), \varphi_0^{Z,(j)}(\Theta_e^Z) \right)$ given by (3.14) or by (3.15) ; $\alpha_0^{(j)} \approx -\varphi_0^{Y,(j)} + \varphi_0^{EY,(j)}$ at t_0 ;

    • $L_1(\Theta_e^Y, \Theta_e^Z) = 0$, $L_2(\Theta_e^{EY}) = 0$;

    for $(i = 0, \ldots, N-1)$ **do**

        • $X_{t_{i+1}}^{(j)} = X_{t_i}^{(j)} + (\alpha_{t_i}^{(j)} - \alpha_{t_{i-D}}^{(j)})\Delta t + \sigma \sqrt{\Delta t}\Delta W_{t_i}^{(j)}$;

        • $\bar{X}_{t_{i+1}} = \frac{1}{M}\sum_{j=1}^M X_{t_{i+1}}^{(j)}$;

        • $\tilde{Y}_{t_{i+1}}^{(j)} = Y_{t_k}^{(j)} - \epsilon(X_{t_i}^{(j)} - \bar{X}_{t_i})\Delta t + Z_{t_k}\Delta W_{t_i}^{(j)}$;

        if $(i \leq N-D)$ **then**

            • $\left(Y_{t_{i+1}}^{(j)}, \mathbb{E}[Y_{t_{i+1+D}}^{(j)}|\mathcal{F}_{t_{i+1}}], Z_{t_{i+1}}^{(j)} \right) \approx \left(\varphi_{t_{i+1}}^{Y,(j)}(\Theta_e^Y), \varphi_{t_{i+1}}^{EY,(j)}(\Theta_e^{EY}), \varphi_{t_{i+1}}^{Z,(j)}(\Theta_e^Z) \right)$ given by (3.14) or by(3.15) at t_{i+1} ;

            • $L_1 + = \frac{1}{M}\sum_{j=1}^M \left(\varphi_{t_{i+1}}^{Y,(j)} - \tilde{Y}_{t_{i+1}}^{(j)} \right)^2$;

        else

            • $\left(Y_{t_{i+1}}^{(j)}, \mathbb{E}[Y_{t_{i+1+D}}^{(j)}|\mathcal{F}_{t_{i+1}}], Z_{t_{i+1}}^{(j)} \right) \approx \left(\varphi_{t_{i+1}}^{Y,(j)}(\Theta_e^Y), 0, \varphi_{t_{i+1}}^{Z,(j)}(\Theta_e^Z) \right)$ given by (3.14) or by(3.15) at t_{i+1} ;

            • $L_1 + = \frac{1}{M}\sum_{j=1}^M \left(\varphi_{t_{i+1}}^{Y,(j)}\tilde{Y}_{t_{i+1}}^{(j)} \right)^2$;

        end

    end

    for $(i = 0, 1, \ldots, N-D)$ **do**

        • $L_2 + = \frac{1}{M}\sum_{j=1}^M \left(\varphi_{t_i}^{EY,(j)} - \tilde{Y}_{t_{i+D}}^{(j)} \right)^2$;

    end

    • $L_1 + = \frac{1}{M}\sum_{j=1}^M \left(\varphi_{t_N}^{Y,(j)} - c(X_{t_N}^{(j)} - \bar{X}_{t_N}) \right)^2$;

    • Compute the gradient $\nabla L_1(\Theta^Y, \Theta^Z)$by backpropagation through time;

    • Update Θ_{e+1}^Y and Θ_{e+1}^Z according to SGD algorithm;

    • Compute the gradient $\nabla L_2(\Theta^{EY})$by backpropagation through time;

    • Update Θ_{e+1}^{EY} according to SGD algorithm;

    • Stop if $L^1(\Theta^Y, \Theta^Z)$ are close to 0, and $L_2(\Theta^{EY})$ converges, return $(\Theta^Y, \Theta^{EY}, \Theta^Z)$;

end

12

Again, in the following graphics, we choose $c = 1, \sigma = 1, T = 10, \tau = 4, \Delta t = 0.1, M = 2000$. For a specific representative path, the underlying Brownian motion paths are the same for different approaches. Figure 3.3 compares one representative optimal trajectory of the dynamic and the control, and they coincide. Figure 3.4 plot the sample average of optimal trajectory of the dynamic and the control, which are trajectories of 0. This is the same as the theoretical mean. Comparing to Figure 3.1 and Figure 3.2, as well as based on numerous experiments, we find that given a path of Brownian motion, the two algorithms would yield the same optimal trajectory of state dynamic and the same path for the optimal control. From Figure 3.6, the loss L_1 as defined in (3.16) are minimized to 0 for both approach 3 and approach 4. This can also be observed from Figure 3.5, the red dash line and the blue solid line coincide for both left and right graphs. In addition, from the righthand side of Figure 3.6, we observe the loss L_2 as defined in (3.16) converges. This is due to the fact that the conditional expectation can be understand as an orthogonal projection. Figure 3.7 plot 64 sample paths of the process $(Z_t)_{0 \leq t \leq T}$, which seems to be a deterministic function since σ is constant in this example. Finally, Figure (3.8) show the convergence of the value function as number of epoches increases. Although, two algorithms would arrive approximately at the same optimal value with is about 6, the first algorithm converges faster than the second one, since it directly paramerizes the control, instead of solving the MV-FABSDE system.
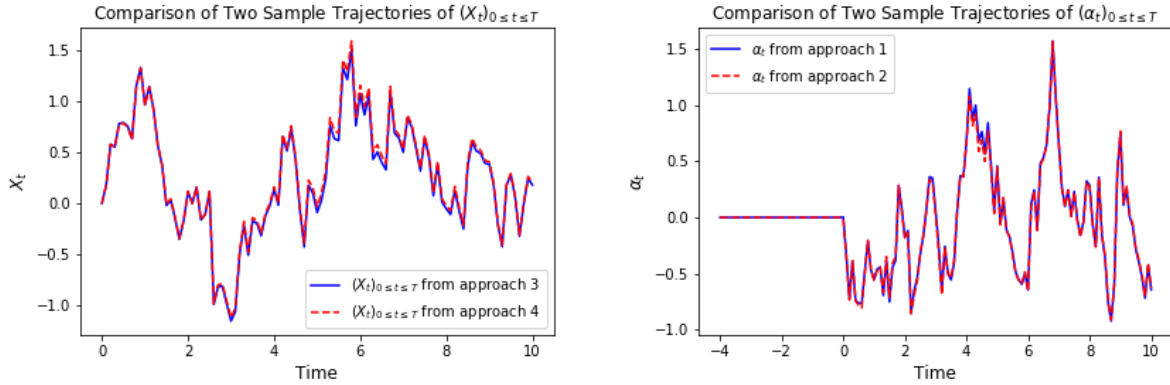

Figure 3.3: On the left, we compare one representative optimal trajectory of $(X_t)_{0 \leq t \leq T}$. The plot on the right show the comparison of one representative optimal trajectory of $(\alpha_t)_{0 \leq t \leq T}$ between approach 3 and approach 4.
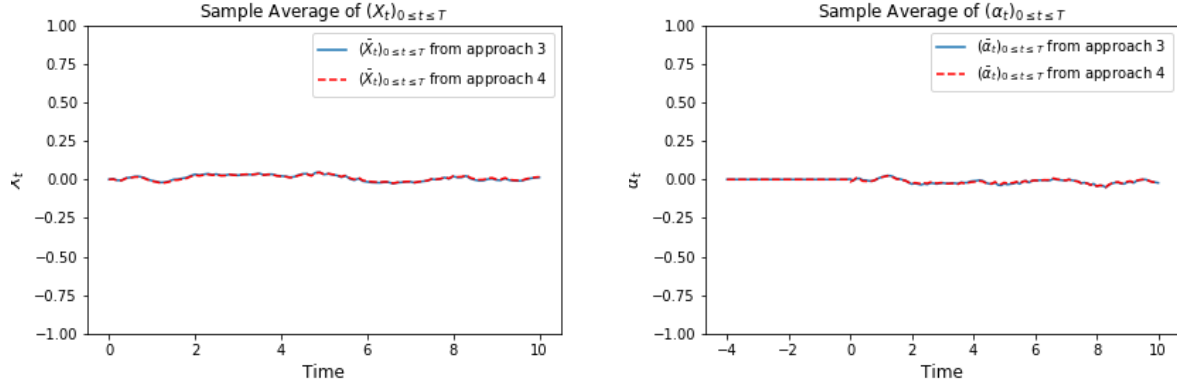
Figure 3.4: On the left, we compare the sample mean of optimal trajectories of $(X_t)_{0 \le t \le T}$. The plot on the right show the comparison of sample mean of trajectories of optimal control $(\alpha_t)_{0 \le t \le T}$ between approach 3 and approach 4.
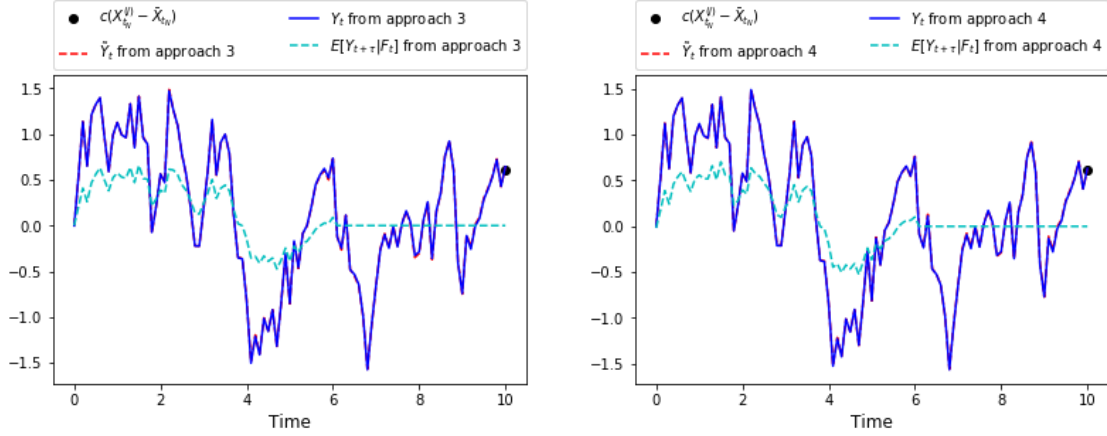

Figure 3.5: Plots of representative trajectories of $((Y_t)_{0 \le t \le T}, (\tilde{Y}_t)_{0 \le t \le T}, (\mathbb{E}[Y_{t+\tau}|\mathcal{F}_t])_{0 \le t \le T})$, from approach 3 (one the left) and from approach 4 (on the right).

4 Stochastic Maximum Principle for Optimality

In this section, we derive the adjoint equation associated to our mean field stochastic control problem (2.3) and (2.4). The necessary and sufficient parts of stochastic maximum principle have been proved for optimality. We assume

(H4.1) b, σ are differentiable with respect to $(X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau})$; f is differentiable with respect to $(X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha)$; g is differentiable with respect to (X_T, μ_T), and their derivatives are bounded.

In order to simplify our notations, let $\theta_t = (X_t, \mu_t, \alpha_t)$. For $0 < \epsilon < 1$, we denote α^ϵ the
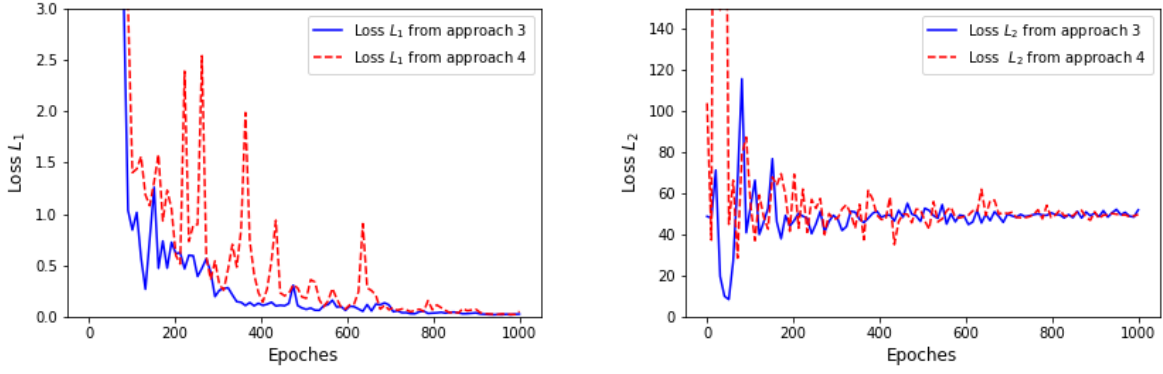
14

Figure 3.6: Loss L_1 (on the left) and L_2 (on the right) as defined in (3.16) from approach 3 and approach 4.
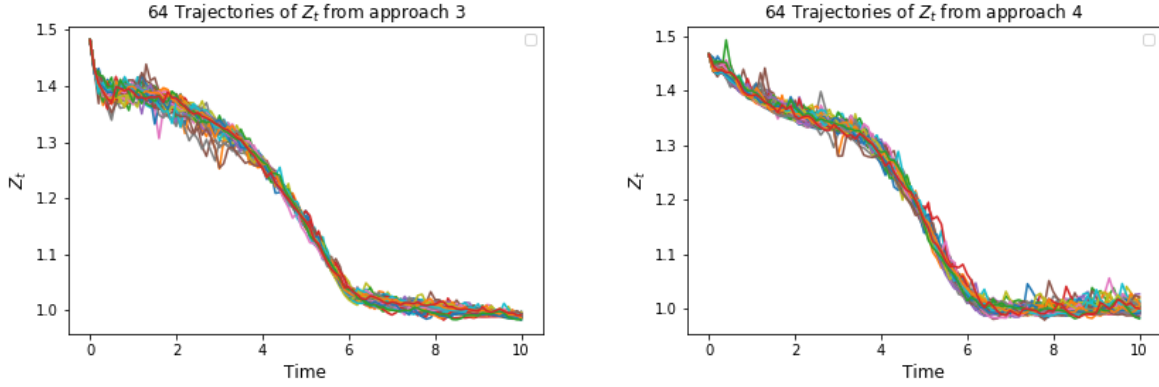

Figure 3.7: 64 trajectories of $(Z_t)_{0 \le t \le T}$ based on approach 3 (on the left) and approach 5 (on the right).

admissible control defined by

$$\alpha_t^\epsilon := \alpha_t + \epsilon(\beta_t - \alpha_t) := \alpha_t + \epsilon \Delta \alpha_t,$$

for any $(\alpha)_{0 \le t \le T}$ and $(\beta)_{0 \le t \le T} \in \mathbb{A}$. $X_t^\epsilon := X_t^{\alpha^\epsilon}$ is the corresponding controlled process. We define

$$\nabla X_t := \lim_{\epsilon \to 0} \frac{X_t^\epsilon - X_t^\alpha}{\epsilon}$$
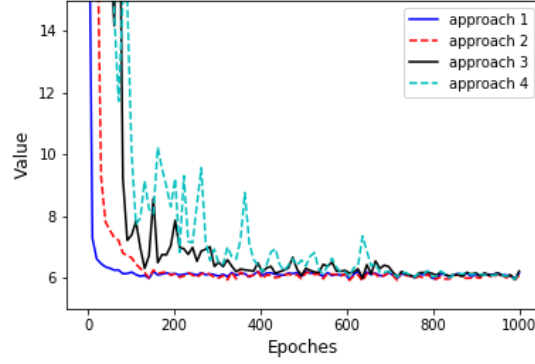
15

Figure 3.8: Comparison convergence of objective values as in (3.10) among four approaches.

to be the variation process, which should follow the following dynamic for $t \in (0, T]$,

$$
\begin{aligned}
d\nabla X_t =& \Big[\partial_x b(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} b(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t] \\
&+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_t} b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dt \\
&+ \Big[\partial_x \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t] \\
&+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dW_t
\end{aligned}
$$
(4.1)

with initial condition $\nabla X_t = \Delta \alpha_t = 0$, $t \in [-\tau, 0]$. $(\tilde{X}_t, \nabla \tilde{X}_t)$ is a copy of $(X_t, \nabla X_t)$ defined on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, where we apply differential calculus on functions of measure, see [***Carmona Book ***] for detail. $\partial_x b, \partial_{x_\tau} b, \partial_\alpha b, \partial_{\alpha_\tau} b$ are derivatives of b with respect to $(X_t, X_{t-\tau}, \alpha_t, \alpha_{t-\tau})$ respectively, and we use the same notation for $\partial.\sigma$.

In the meantime, the Gateaux derivative of functional $(\alpha) \to J(\alpha)$ is given by

$$
\begin{aligned}
&\lim_{\epsilon \to 0} \frac{J(\alpha^\epsilon) - J(\alpha)}{\epsilon} \\
=& \mathbb{E}\Big[\partial_x g(X_T, \mu_T) \nabla X_T + \tilde{\mathbb{E}}[\partial_\mu g(X_T, \mu_T)(\tilde{X}_T) \nabla \tilde{X}_T] \Big] \\
&+ \mathbb{E} \int_0^T \Big[\partial_x f(\theta_t, X_{t-\tau}, \mu_{t-\tau}) \nabla X_t + \tilde{\mathbb{E}}[\partial_\mu f(\theta_t, X_{t-\tau}, \mu_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t] \\
&+ \partial_{x_\tau} f(\theta_t, X_{t-\tau}, \mu_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_{\mu_\tau} f(\theta_t, X_{t-\tau}, \mu_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] \\
&+ \partial_\alpha f(\theta_t, X_{t-\tau}, \mu_{t-\tau})(\Delta \alpha_t) \Big] dt
\end{aligned}
$$
(4.2)

In order to determine the adjoint backward equation of $(Y_t, Z_t)_{0 \le t \le T}$ associated to (2.3), we

16

assume it is of the following form:

$$dY_t = -\varphi_t dt + Z_t dW_t, \quad t \in [0, T],$$
$$Y_T = \partial_x g(X_T, \mu_T) + \tilde{\mathbb{E}}[\partial_\mu g(X_T, \mu_T)(\tilde{X}_T)], \tag{4.3}$$
$$Y_t = Z_t = 0, \quad t \in (T, T + \tau]$$

Next, we apply integration by part to ∇X_t and Y_t. It yields

$$d(\nabla X_t Y_t)$$
$$= Y_t \Big[\partial_x b(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} b(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t]$$
$$+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_\tau} b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dt$$
$$+ Y_t \Big[\partial_x \sigma(t, \theta_t, \theta_{t-\tau}) + \partial_{x_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t]$$
$$+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dW_t$$
$$- \varphi \nabla X_t dt + \nabla X_t Z_t dW_t$$
$$+ Z_t \Big[\partial_x \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t]$$
$$+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dt$$

We integrate from 0 to T, and take expectation to get

$$\mathbb{E}[\nabla X_T Y_T] = \mathbb{E} \int_0^T Y_t \Big[\partial_x b(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} b(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t]$$
$$+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} b(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha + \partial_{\alpha_\tau} b(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dt$$
$$- \mathbb{E} \int_0^T \varphi_t \nabla X_t dt$$
$$+ \mathbb{E} \int_0^T Z_t \Big[\partial_x \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_t + \partial_{x_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \nabla X_{t-\tau} + \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_t) \nabla \tilde{X}_t]$$
$$+ \tilde{\mathbb{E}}[\partial_{\mu_\tau} \sigma(t, \theta_t, \theta_{t-\tau})(\tilde{X}_{t-\tau}) \nabla \tilde{X}_{t-\tau}] + \partial_\alpha \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_t + \partial_{\alpha_\tau} \sigma(t, \theta_t, \theta_{t-\tau}) \Delta \alpha_{t-\tau} \Big] dt$$
$$\tag{4.4}$$

Using the fact that $Y_t = Z_t = 0$ for $t \in (T, T + \tau]$, we are able to make a change of time, and by

17

Fubini's theorem, so that (4.4) becomes

$$
\begin{aligned}
\mathbb{E}[\nabla X_T Y_T] =& \mathbb{E} \int_0^T \bigg(Y_t \partial_x b(t, \theta_t, \theta_{t-\tau}) + Y_{t+\tau} \partial_{x_\tau} b(t+\tau, \theta_{t+\tau}, \theta_t) + \tilde{\mathbb{E}}[\partial_\mu b(t, \tilde{\theta}_t, \tilde{\theta}_{t-\tau})(X_t)] \\
& + \tilde{\mathbb{E}}[\partial_{\mu_\tau} b(t+\tau, \tilde{\theta}_{t+\tau}, \tilde{\theta}_t)(X_t)] \bigg) \nabla X_t dt \\
& + \mathbb{E} \int_0^T \bigg(\partial_\alpha b(t, \theta_t, \theta_{t-\tau}) + \partial_{\alpha_\tau} b(t+\tau, \theta_{t+\tau}, \theta_t) \bigg) \Delta \alpha_t dt \\
& - \mathbb{E}\bigg[\int_0^T \varphi_t \nabla X_t \bigg] dt \\
& + \mathbb{E} \int_0^T \bigg(Z_t \partial_x \sigma(t, \theta_t, \theta_{t-\tau}) + Z_{t+\tau} \partial_{x_\tau} \sigma(t+\tau, \theta_{t+\tau}, \theta_t) + \tilde{\mathbb{E}}[\partial_\mu \sigma(t, \tilde{\theta}_t, \tilde{\theta}_{t-\tau})(X_t)] \\
& + \tilde{\mathbb{E}}[\partial_{\mu_\tau} \sigma(t+\tau, \tilde{\theta}_{t+\tau}, \tilde{\theta}_t)(X_t)] \bigg) \nabla X_t dt \\
& + \mathbb{E} \int_0^T \bigg(\partial_\alpha \sigma(t, \theta_t, \theta_{t-\tau}) + \partial_{\alpha_\tau} \sigma(t+\tau, \theta_{t+\tau}, \theta_t) \bigg) \Delta \alpha_t dt
\end{aligned}
\tag{4.5}
$$

Now we define the Hamiltonian H for $(t, x, \mu, x_\tau, \mu_\tau, y, z, \alpha, \alpha_\tau) \in [0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \times A \times A \times \mathbb{R} \times \mathbb{R}$ as

$$
\begin{aligned}
H(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau, y, z) =& b(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau)y + \sigma(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau)y \\
& + f(t, x, \mu, x_\tau, \mu_\tau, \alpha)
\end{aligned}
\tag{4.6}
$$

Using the terminal condition of Y_T, and plug (4.5) into (4.2), and set the integrand containing ∇X_t to zero, we are able to obtain the adjoint equation is of the following form

$$
\begin{aligned}
dY_t = -& \bigg\{ \partial_x H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{X}_t, \mu_t, \tilde{X}_{t-\tau}, \mu_{t-\tau}, \tilde{\alpha}_t, \tilde{\alpha}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)] \\
& + \mathbb{E}[\partial_{x_\tau} H(t+\tau, X_{t+\tau}, \mu_{t+\tau}, X_t, \mu_t, \alpha_{t+\tau}, \alpha_t, Y_{t+\tau}, Z_{t+\tau}) | \mathcal{F}_t] \\
& + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t+\tau, \tilde{X}_{t+\tau}, \mu_{t+\tau}, \tilde{X}_t, \mu_t, \tilde{\alpha}_t, \tilde{\alpha}_{t-\tau}, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(X_{t+\tau})] | \mathcal{F}_t] \bigg\} dt + Z_t dW_t \\
Y_T =& \partial_x g(X_T, \mu_T) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mu_T)(X_T)].
\end{aligned}
\tag{4.7}
$$

Theorem 4.1. *Let $(\alpha_t)_{0 \le t \le T} \in \mathbb{A}$ be optimal, $(X_t)_{0 \le t \le T}$ be the associated controlled state, and*

$(Y_t, Z_t)_{0 \le t \le T}$ *be the associated adjoint processes defined in* (4.7).*For any* $\beta \in A$, *and* $t \in [0, T]$,

$$\Bigg(\partial_\alpha H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau}, Y_t, Z_t)$$

$$+ \mathbb{E}[\partial_{\alpha_\tau} H(t + \tau, X_{t+\tau}, \mu_{t+\tau}, X_t, \mu_t, \alpha_{t+\tau}, \alpha_t, Y_{t+\tau}, Z_{t+\tau}) | \mathcal{F}_t] \Bigg)(\beta - \alpha_t) \ge 0 \ a.e. \tag{4.8}$$

Proof. Given any $(\beta_t)_{0 \le t \le T} \in \mathbb{A}$, we perturbate α_t by $\alpha_t^\epsilon := \alpha_t + \epsilon(\beta_t - \alpha_t)$ for $0 \le \epsilon \le 1$. Using the adjoint process (4.7), and apply integration by parts formula to $(\nabla X_t Y_t)$. Then plug the result into (4.2), and the Hamiltonian H is defined in (4.6). Also, we use the fact that α is optimal, then

$$0 \le \lim_{\epsilon \to 0} \frac{J(\alpha^\epsilon) - J(\alpha)}{\epsilon}$$

$$= \mathbb{E} \int_0^T \Bigg([\partial_\alpha H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \mathbb{E}[\partial_{\alpha_\tau} H(t + \tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau}) | \mathcal{F}_t] \Bigg)(\beta_t - \alpha_t) dt \tag{4.9}$$

Now, let $C \in \mathcal{F}_t$ be an arbitrary progressively measurable set, and denote C' the complement of C. We choose β_t to be $\beta_t := \beta \mathbb{1}_C + \alpha_t \mathbb{1}_{C'}$ for any given $\beta in A$. Then,

$$\mathbb{E} \int_0^T \Bigg([\partial_\alpha H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \mathbb{E}[\partial_{\alpha_\tau} H(t + \tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau}) | \mathcal{F}_t] \Bigg)(\beta_t - \alpha_t) \mathbb{1}_C dt \ge 0, \tag{4.10}$$

which implies,

$$(\partial_\alpha H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \mathbb{E}[\partial_{\alpha_\tau} H(t + \tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau}) | \mathcal{F}_t])(\beta - \alpha_t) \ge 0. \ a.e. \tag{4.11}$$

$\square$

Remark 4.2. When we further assume that H is convex in $(\alpha_t, \alpha_{t-\tau})$, then for any $\beta, \beta_\tau \in A$, we have

$$H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau}, Y_t, Z_t) \le H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \beta, \beta_\tau, Y_t, Z_t), \ \text{a.e.}$$

as a direct consequence of (4.8).

Theorem 4.3. *Let* $(\alpha_t)_{0 \le t \le T} \in \mathbb{A}$ *be an admissible control. Let* $(X_t)_{0 \le t \le T}$ *be the controlled state, and* $(Y_t, Z_t)_{0 \le t \le T}$ *be the corresponding adjoint processes. We further assume that for each* t, *given* Y_t *and* Z_t, *the function* $(x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau) \to H(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau, Y_t, Z_t)$, *and the function*

$(x, \mu) \rightarrow g(x, \mu)$ *are convex. If*

$$H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau}, Y_t, Z_t) = \inf_{\alpha' \in \mathbb{A}} H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha'_t, \alpha'_{t-\tau}, Y_t, Z_t), \quad (4.12)$$

then $(\alpha_t)_{0 \leq t \leq T}$ *is an optimal control.*

Proof. Let $(\alpha'_t)_{0 \leq t \leq T} \in \mathbb{A}$ be a admissible control, and let $(X'_t)_{0 \leq t \leq T} = (X_t^{\alpha'})_{0 \leq t \leq T}$ be the corresponding controlled state. From the definition of the objective function as in (2.4), we first use convexity of g, and the terminal condition of the adjoint process Y_t in (4.7), then use the fact that H is convex, and because of (4.12), we have the following

$$J(\alpha) - J(\alpha')$$

$$=\mathbb{E}[g(X_T, \mu_T) - g(X'_T, \mu'_T)] + \mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \alpha_{t-\tau}) - f(t, \theta'_t, X'_{t-\tau}, \alpha'_{t-\tau})]dt$$

$$\leq \mathbb{E}[\partial_x g(X_T, \mu_T)(X_T - X'_T) + \tilde{\mathbb{E}}[\partial_\mu g(X_T, \mu_T)(\tilde{X}_T)(\tilde{X}_T - \tilde{X}'_T)]]$$

$$+ \mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \alpha_{t-\tau}) - f(t, \theta'_t, X'_{t-\tau}, \alpha'_{t-\tau})]dt$$

$$=\mathbb{E}[(\partial_x g(X_T, \mu_T) + \tilde{E}[\partial_\mu g(\tilde{X}_T, \mu_T)(X_T)])(X_T - X'_T)]$$

$$+ \mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \alpha_{t-\tau}) - f(t, \theta'_t, X'_{t-\tau}, \alpha'_{t-\tau})]dt$$

$$=\mathbb{E}[Y_T(X_T - X'_T)] + \mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \alpha_{t-\tau}) - f(t, \theta'_t, X'_{t-\tau}, \alpha'_{t-\tau})]dt$$

$$=\mathbb{E}\int_0^T \left[(b(t, \theta_t, \theta_{t-\tau}) - b(t, \theta'_t, \theta'_{t-\tau}))Y_t + (\sigma(t, \theta_t, \theta_{t-\tau}) - \sigma(t, \theta'_t, \theta'_{t-\tau}))Z_t\right]dt$$

$$- \mathbb{E}\int_0^T \left[\partial_x H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\theta}_t, \tilde{\theta}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)](X_t - X'_t)\right]dt$$

$$- \mathbb{E}\int_0^T \left[\mathbb{E}[\partial_{x_\tau} H(t + \tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau})|\mathcal{F}_t]\right.$$

$$\left. + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t + \tau, \tilde{\theta}_{t+\tau}, \tilde{\theta}_t, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(X_t)]|\mathcal{F}_t]\}(X_t - X'_t)\right]dt$$

$$+ \mathbb{E}\int_0^T \left[H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) - H(t, \theta'_t, \theta'_{t-\tau}, Y_t, Z_t)\right]dt$$

$$+ \mathbb{E}\int_0^T \left[(b(t, \theta_t, \theta_{t-\tau}) - b(t, \theta_t, \theta_{t-\tau}))Y_t + (\sigma(t, \theta_t, \theta_{t-\tau}) - \sigma(t, \theta_t, \theta_{t-\tau}))Z_t\right]dt$$

$$=- \mathbb{E}\int_0^T \left[\partial_x H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(X_t - X'_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(\tilde{X}_t)(\tilde{X}_t - \tilde{X}'_t)]\right]dt$$

$$- \mathbb{E}\int_0^T \left[\partial_{x_\tau} H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(X_{t-\tau} - X'_{t-\tau}) + \tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(\tilde{X}_{t-\tau})(\tilde{X}_{t-\tau} - \tilde{X}'_{t-\tau})]\right]dt$$

$$+ \mathbb{E}\int_0^T \left[H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) - H(t, \theta'_t, \theta'_{t-\tau}, Y_t, Z_t)\right]dt$$

$$\leq \mathbb{E}\int_0^T \left[\partial_\alpha H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(\alpha_t - \alpha'_t) + \partial_{\alpha_\tau} H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t)(\alpha_{t-\tau} - \alpha'_{t-\tau})\right]dt$$

$$\leq \mathbb{E}\int_0^T \left(\partial_\alpha H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \mathbb{E}[\partial_{\alpha_\tau} H(t + \tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau})]\right)(\alpha_t - \alpha'_t)dt$$

$$\leq 0.$$

$$(4.13)$$

$$\square$$

5 Existence and Uniqueness Result

Given the necessary and sufficient conditions proven in Section 4, we use the optimal control $(\hat{\alpha}_t)_{0 \leq t \leq T}$ defined by

$$
\begin{aligned}
&\hat{\alpha}(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, Y_t, Z_t, \mathbb{E}[Y_{t+\tau}|\mathcal{F}_t], \mathbb{E}[Z_{t+\tau}|\mathcal{F}_t]) \\
&= \arg\min_{\alpha \in \mathbb{A}} H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau}, Y_t, Z_t),
\end{aligned}
\tag{5.1}
$$

to establish the solvability result of the McKean-Vlasov FABSDE (2.3) and (4.7) for $t \in [0, T]$:

$$
\begin{aligned}
dX_t =& b(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \alpha_t, \alpha_{t-\tau})dt + \sigma(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \hat{\alpha}_t, \hat{\alpha}_{t-\tau})dW_t, \\
dY_t =& -\Bigg\{ \partial_x H(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, \hat{\alpha}_t, \hat{\alpha}_{t-\tau}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{X}_t, \mu_t, \tilde{X}_{t-\tau}, \mu_{t-\tau}, \tilde{\hat{\alpha}}_t, \tilde{\hat{\alpha}}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)] \\
& + \mathbb{E}[\partial_{x_\tau} H(t+\tau, X_{t+\tau}, \mu_{t+\tau}, X_t, \mu_t, \hat{\alpha}_{t+\tau}, \hat{\alpha}_t, Y_{t+\tau}, Z_{t+\tau})|\mathcal{F}_t] \\
& + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t+\tau, \tilde{X}_{t+\tau}, \mu_{t+\tau}, \tilde{X}_t, \mu_t, \tilde{\hat{\alpha}}_{t+\tau}, \tilde{\hat{\alpha}}_t, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(X_t)]|\mathcal{F}_t] \Bigg\}dt + Z_t dW_t
\end{aligned}
\tag{5.2}
$$

with initial condition $X_0 = x_0$; $X_t = \hat{\alpha}_t = 0$ for $t \in [-\tau, 0)$ and terminal condition $Y_T = \partial_x g(X_T, \mu_T) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mu_T)(X_T)]$. In addition to assumption (H 4.1), we further assume

(H5.1) The drift and volatility function b and σ are linear in $x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau$. For all $(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau) \in [0, T] \times \mathbb{R} \times \mathcal{P}_2(\mathbb{R}) \times \mathcal{P}_2(\mathbb{R}) \times A \times A$, we assume that

$$
\begin{aligned}
b(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau) =& b_0(t) + b_1(t)x + \bar{b}_1(t)m + b_2(t)x_\tau + \bar{b}_2(t)m_\tau + b_3(t)\alpha + b_4(t)\alpha_\tau, \\
\sigma(t, x, \mu, x_\tau, \mu_\tau, \alpha, \alpha_\tau) =& \sigma_0(t) + \sigma_1(t)x + \bar{\sigma}_1(t)m + \sigma_2(t)x_\tau + \bar{\sigma}_2(t)m_\tau + \sigma_3(t)\alpha + \sigma_4(t)\alpha_\tau,
\end{aligned}
\tag{5.3}
$$

for some measurable deterministic functions $b_0, b_1, \bar{b}_1, b_2, \bar{b}_2, b_3, b_4, \sigma_0, \sigma_1, \bar{\sigma}_1, \sigma_2, \bar{\sigma}_2, \sigma_3, \sigma_4$ with values in $\mathbb{R}$ bounded by R, and we have used the notation $m = \int x d\mu(x)$ and $m_\tau = \int x d\mu_\tau(x)$ for the mean of measures μ and μ_τ respectively.

(H5.2) The derivatives of f and g with respect to $(x, x_\tau, \mu, \mu_\tau, \alpha)$ and (x, μ) are Lipschitz continuous with Lipschitz constant L.

(H5.3) The function f is strongly L-convex, which means that for any $t \in [0, T]$, any $x, x', x_\tau, x'_\tau \in \mathbb{R}$, any $\alpha, \alpha' \in A$, any $\mu, \mu' \in \mathcal{P}_2(\mathbb{R})$, any random variables X and X' having μ and μ' as

distribution, and any random variables X_τ and X'_τ having μ_τ and μ'_τ as distribution, then

$$
\begin{aligned}
&f(t, x', \mu', x'_\tau, \mu'_\tau, \alpha') - f(t, x, \mu, x_\tau, \mu_\tau, \alpha) \\
&\quad - \partial_x f(t, x, \mu, x_\tau, \mu_\tau, \alpha)(x' - x) - \partial_{x_\tau} f(t, x, \mu, x_\tau, \mu_\tau, \alpha)(x'_\tau - x_\tau) \\
&\quad - \mathbb{E}[\partial_\mu f(t, x, \mu, x_\tau, \mu_\tau, \alpha)(X) \cdot (X' - X)] - \mathbb{E}[\partial_{\mu_\tau} f(t, x, \mu, x_\tau, \mu_\tau, \alpha)(X_\tau) \cdot (X'_\tau - X_\tau)] \\
&\quad - \partial_\alpha f(t, x, \mu, x_\tau, \mu_\tau, \alpha)(\alpha' - \alpha) \geq \kappa |\alpha' - \alpha|^2.
\end{aligned}
\tag{5.4}
$$

The function g is also assumed to be L-convex in (x, μ).

Theorem 5.1. *Under assumptions (H5.1-5.3), the McKean-Vlasov FABSDE (5.2) is uniquely solvable.*

The proof is based on continuation methods. Let $\lambda \in [0, 1]$, consider the following class of McKean-Vlasov FABSDEs, denoted by MV-FABSDE(λ), for $t \in [0, T]$:

$$
\begin{aligned}
dX_t =& (\lambda b(t, \theta_t, \theta_{t-\tau}) + I_t^b)dt + (\lambda \sigma(t, \theta_t, \theta_{t-\tau}) + I_t^\sigma)dW_t, \\
dY_t =& - \Bigg\{ \lambda \bigg(\partial_x H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\theta}_t, \tilde{\theta}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)] \\
& + \mathbb{E}[\partial_{x_\tau} H(t+\tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau})|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t+\tau, \tilde{\theta}_{t+\tau}, \tilde{\theta}_t, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(X_t)]|\mathcal{F}_t] \bigg) + I_t^f \Bigg\}dt \\
& + Z_t dW_t,
\end{aligned}
\tag{5.5}
$$

where we denote $\theta_t = (X_t, \mu_t, \alpha_t)$, with optimality condition

$$
\alpha_t = \hat{\alpha}(t, X_t, \mu_t, X_{t-\tau}, \mu_{t-\tau}, Y_t, Z_t, \mathbb{E}[Y_{t+\tau}|\mathcal{F}_t], \mathbb{E}[Z_{t+\tau}|\mathcal{F}_t]), t \in [0, T],
$$

and with initial condition $X_0 = x_0; X_t = \alpha_t = 0$ for $t \in [-\tau, 0)$ and terminal condition $Y_T = \lambda \Big\{ \partial_x g(X_T, \mu_T) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mu_T)(X_T)] \Big\} + I_T^g$, and $Y_t = 0$ for $t \in (T, T+\tau]$, where $(I_t^b, I_t^\sigma, I_t^f)_{0 \leq t \leq T}$ are some square-integrable progressively measurable processes with values in $\mathbb{R}$, and $I_T^g \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ is a square integrable $\mathcal{F}_T$-measurable random variable with value in $\mathbb{R}$.

Observe that when $\lambda = 0$, system (5.5) becomes decoupled standard SDE and BSDE, which has an unique solution. When setting $\lambda = 1$, $I_t^b = I_t^\sigma = I_t^f = 0$ for $0 \leq t \leq T$, and $I_T^g = 0$, we are able to recover the system of (5.2).

Lemma 5.2. *Given $\lambda_0 \in [0, 1)$, for any square-integrable progressively measurable processes $(I_t^b, I_t^\sigma, I_t^f)_{0 \leq t \leq T}$, and $I_T^g \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, such that system FABSDE(λ_0) admits a unique solution, then there exists $\delta_0 \in (0, 1)$ such that the system MV-FABSDE(λ) admits a unique solution for any $\lambda \in [\lambda_0, \lambda_0 + \delta_0]$.*

Proof. Assuming that $(\check{X}, \check{Y}, \check{Z}, \check{\alpha})$ are given as an input, for any $\lambda \in [\lambda_0, \lambda_0 + \delta_0]$, where $\delta_0 > 0$ to be determined, denoting $\delta := \lambda - \lambda_0 \leq \delta_0$, we take

$$
\begin{aligned}
I_t^b &\leftarrow \delta[b(t, \check{\theta}_t, \check{\theta}_{t-\tau})] + I_t^b, \\
I_t^\sigma &\leftarrow \delta[\sigma(t, \check{\theta}_t, \check{\theta}_{t-\tau})] + I_t^\sigma, \\
I_t^f &\leftarrow \delta\bigg[\partial_x H(t, \check{\theta}_t, \check{\theta}_{t-\tau}, Y_t, Z_t) + \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\check{\theta}}_t, \tilde{\check{\theta}}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)] \\
&\qquad + \mathbb{E}[\partial_{x_\tau} H(t+\tau, \check{\theta}_{t+\tau}, \check{\theta}_t, \check{Y}_{t+\tau}, \check{Z}_{t+\tau})|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t+\tau, \tilde{\check{\theta}}_{t+\tau}, \tilde{\check{\theta}}_t, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(\check{X}_t)]|\mathcal{F}_t]\bigg] \\
&\qquad + I_t^f, \\
I_T^g &\leftarrow \delta\bigg[\partial_x g(\check{X}_T, \mu_T) + \tilde{\mathbb{E}}[\partial_\mu g(\tilde{\check{X}}_T, \mu_T)(\check{X}_T)]\bigg] + I_T^g.
\end{aligned}
$$
(5.6)

According to the assumption, let (X, Y, Z) be the solutions of MV-FABSDE(λ_0) corresponding to inputs $(\check{X}, \check{Y}, \check{Z})$, i.e., for $t \in [0, T]$

$$
\begin{aligned}
dX_t =&(\lambda_0 b_t + \delta \check{b}_t + I_t^b)dt + (\lambda_0 \sigma_t + \delta \check{\sigma}_t + I_t^\sigma)dW_t, \\
dY_t =& -\bigg\{\lambda_0(\partial_x H_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{H}_t(X_t)] + \mathbb{E}[\partial_{x_\tau} H_{t+\tau}|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t]) \\
&+ \delta(\partial_x \check{H}_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}(X_t)]|\mathcal{F}_t]) + I_t^f\bigg\}dt \\
&+ Z_t dW_t,
\end{aligned}
$$
(5.7)

with initial condition, $X_0 = x_0$, $X_s = \alpha_s = 0$ for $s \in [-\tau, 0)$, and terminal condition

$$
Y_T = \lambda_0\bigg(\partial_x g_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(X_T)]\bigg) + \delta\bigg(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{g}}_T(\check{X}_T)]\bigg) + I_T^g,
$$
(5.8)

and $Y_t = Z_t = 0$ for $t \in (T, T+\tau]$, where we have used simplified notations,

$$
\begin{aligned}
b_t :=&b(t, \theta_t, \theta_{t-\tau}); \quad \check{b}_t := b(t, \check{\theta}_t, \check{\theta}_{t-\tau}); \quad \sigma_t := \sigma(t, \theta_t, \theta_{t-\tau}); \quad \check{\sigma}_t := \sigma(t, \check{\theta}_t, \check{\theta}_{t-\tau}); \\
\partial_x H_t :=&\partial_x H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t); \quad \tilde{\mathbb{E}}[\partial_\mu \tilde{H}_t(X_t)] := \tilde{\mathbb{E}}[\partial_\mu H(t, \tilde{\theta}_t, \tilde{\theta}_{t-\tau}, \tilde{Y}_t, \tilde{Z}_t)(X_t)] \\
\mathbb{E}[\partial_{x_\tau} H_{t+\tau}|\mathcal{F}_t] :=&\mathbb{E}[\partial_{x_\tau} H(t+\tau, \theta_{t+\tau}, \theta_t, Y_{t+\tau}, Z_{t+\tau})|\mathcal{F}_t]; \\
\mathbb{E}[\tilde{\mathbb{E}}[\partial_\mu \tilde{H}_t(X_t)]|\mathcal{F}_t] :=&\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} H(t+\tau, \tilde{\theta}_{t+\tau}, \tilde{\theta}_t, \tilde{Y}_{t+\tau}, \tilde{Z}_{t+\tau})(X_t)]|\mathcal{F}_t]; \\
\partial_x g_T :=&\partial_x g(X_T, \mu_T); \quad \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(X_T)]] := \tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T, \mu_T)(X_T)]
\end{aligned}
$$

similar notation for $\partial_x \check{H}_t$, $\tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t(X_t)]$, $\mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}|\mathcal{F}_t]$, and $\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}(X_t)]|\mathcal{F}_t]$.
(5.9)

We would like to show that the map $\Phi : (\check{X}, \check{Y}, \check{Z}, \check{\alpha}) \rightarrow \Phi(\check{X}, \check{Y}, \check{Z}, \check{\alpha}) = (X, Y, Z, \alpha)$ is a

contraction.

Applying integration by parts to $\Delta X_t Y_t$, we have

$$
\begin{aligned}
&d(\Delta X_t Y_t) \\
={}&Y_t\bigg\{[\lambda_0\Delta b_t + \delta\Delta\check{b}_t]dt + [\lambda_0\Delta\sigma_t + \delta\Delta\check{\sigma}_t]dW_t\bigg\} \\
&- \Delta X_t\bigg\{\lambda_0(\partial_x H_t + \tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)] + \mathbb{E}[\partial_{x_\tau}H_{t+\tau}|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t]) \\
&+ \delta(\partial_x\check{H}_t + \tilde{\mathbb{E}}[\partial_\mu\check{\tilde{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t] + \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\check{\tilde{H}}_{t+\tau}(\check{X}_t)]|\mathcal{F}_t])\bigg\}dt \\
&+ \Delta X_t Z_t dW_t + (\lambda_0\Delta\sigma_t + \delta\Delta\check{\sigma}_t)Z_t dt.
\end{aligned} \tag{5.10}
$$

After integrating from 0 to T, and taking expectation on both sides, we obtain

$$
\begin{aligned}
&\mathbb{E}[\Delta X_T Y_T] \\
={}&\lambda_0\mathbb{E}\int_0^T\bigg(\Delta b_t Y_t + \Delta\sigma_t Z_t - \Delta X_t(\partial_x H_t + \tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)] + \mathbb{E}[\partial_{x_\tau}H_{t+\tau}|\mathcal{F}_t] \\
&+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t])\bigg)dt \\
&+ \delta\mathbb{E}\int_0^T\bigg(\Delta\check{b}_t Y_t + \Delta\check{\sigma}Z_t - \Delta X_t(\partial_x\check{H}_t + \tilde{\mathbb{E}}[\partial_\mu\check{\tilde{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t] \\
&+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\check{\tilde{H}}_{t+\tau}(\check{X}_t)]|\mathcal{F}_t])\bigg)dt
\end{aligned} \tag{5.11}
$$

In the meantime, from the terminal condition of Y_T given in (5.8), and since g is convex, we also have

$$
\begin{aligned}
&\mathbb{E}[\Delta X_T Y_T] \\
={}&\mathbb{E}\left[\Delta X_T\bigg(\lambda_0(\partial_x g_T + \tilde{\mathbb{E}}[\partial_\mu\tilde{g}_T(X_T)]) + \delta(\partial_x\check{g}_T + \tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}_T(X_T)]) + I_T^g\bigg)\right] \\
\geq{}&\lambda_0\mathbb{E}[g(X_T,\mu_T) - g(X_T' - \mu_T')] + \delta\Delta X_T(\partial_x\check{g}_T + \tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}_T(X_T)]) + \Delta X_T I_T^g
\end{aligned} \tag{5.12}
$$

Following the proof of sufficient part of maximum principle and using 5.11, and 5.12, we find

$$\lambda_0(J(\alpha) - J(\alpha'))$$

$$=\lambda_0\mathbb{E}[g(X_T, \mu_T) - g(X_T', \mu_T')] + \lambda_0\mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \mu_{t-\tau}) - f(t, \theta_t', X_{t-\tau}', \mu_{t-\tau}')]dt$$

$$\leq \mathbb{E}[\Delta X_T Y_T] - \delta\Delta X_T(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(\check{X}_T)]) - \Delta X_T I_T^g$$

$$+ \lambda_0\mathbb{E}\int_0^T [f(t, \theta_t, X_{t-\tau}, \mu_{t-\tau}) - f(t, \theta_t', X_{t-\tau}', \mu_{t-\tau}')]dt$$

$$=\lambda_0\mathbb{E}\int_0^T \Bigg[\Delta b_t Y_t + \Delta\sigma_t Z_t - \Delta X_t(\partial_x H_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{H}_t(X_t)] + \mathbb{E}[\partial_{x_\tau} H_{t+\tau}|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t])\Bigg]dt$$

$$+ \delta\mathbb{E}\int_0^T \Bigg[\Delta\check{b}_t Y_t + \Delta\check{\sigma}_t Z_t - \Delta X_t(\partial_x \check{H}_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}(X_t)]|\mathcal{F}_t])\Bigg]dt$$

$$+ \lambda_0\mathbb{E}\int_0^T [H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) - H(t, \theta_t', \theta_{t-\tau}', Y_t, Z_t)]dt$$

$$- \lambda_0\mathbb{E}\int_0^T (\Delta b_t Y_t + \Delta_t\sigma Z_t)dt - \delta\Delta X_T(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(\check{X}_T)]) - \Delta X_T I_T^g$$

$$=\lambda_0\mathbb{E}\int_0^T \Bigg[H(t, \theta_t, \theta_{t-\tau}, Y_t, Z_t) - H(t, \theta_t', \theta_{t-\tau}', Y_t, Z_t) - \Delta X_t(\partial_x H_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{H}_t(X_t)] + \mathbb{E}[\partial_{x_\tau} H_{t+\tau}|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t])\Bigg]dt + \delta\mathbb{E}\int_0^T \Bigg[\Delta\check{b}_t Y_t + \Delta\check{\sigma} Z_t - \Delta X_t(\partial_x \check{H}_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}(\check{X}_t)]|\mathcal{F}_t])\Bigg]dt - \delta\Delta X_T(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(\check{X}_T)]) - \Delta X_T I_T^g$$

$$\leq - \mathbb{E}\int_0^T \lambda_0\kappa|\Delta\alpha_t|^2 dt + \delta\mathbb{E}\int_0^T \Bigg[\Delta\check{b}_t Y_t + \Delta\check{\sigma}_t Z_t - \Delta X_t(\partial_x \check{H}_t + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t(\check{X}_t)] + \mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}(\check{X}_t)]|\mathcal{F}_t])\Bigg]dt - \delta\Delta X_T(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(\check{X}_T)]) - \Delta X_T I_T^g$$

$$(5.13)$$

Reverse the role of α and α', we also have

$$\lambda_0(J(\alpha') - J(\alpha))$$

$$\leq - \mathbb{E}\int_0^T \lambda_0\kappa|\Delta\alpha_t'|^2 dt + \delta\mathbb{E}\int_0^T \Bigg[\Delta\check{b}_t' Y_t' + \Delta\check{\sigma}_t' Z_t' - \Delta X_t'(\partial_x \check{H}_t' + \tilde{\mathbb{E}}[\partial_\mu \tilde{\check{H}}_t'(\check{X}_t')] + \mathbb{E}[\partial_{x_\tau} \check{H}_{t+\tau}'|\mathcal{F}_t]$$

$$+ \mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau} \tilde{\check{H}}_{t+\tau}'(\check{X}_t')]|\mathcal{F}_t])\Bigg]dt - \delta\Delta X_T'(\partial_x \check{g}_T + \tilde{\mathbb{E}}[\partial_\mu \tilde{g}_T(\check{X}_T')]) - \Delta X_T' I_T^g$$

$$(5.14)$$

26

Summing (5.13) and (5.14), using the fact that b and σ have the linear form, using change of time and Lipschitz assumption, it yields

$$2\lambda_0\kappa\mathbb{E}\int_0^T|\Delta\alpha_t|^2dt$$

$$\leq\delta\mathbb{E}\int_0^T\Bigg[\Delta\check{b}_t\Delta Y_t+\Delta\check{\sigma}\Delta Z_t-\Delta X_t(\Delta\partial_x\check{H}_t+\Delta\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{H}}_t(X_t)]+\Delta\mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t]$$

$$+\Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\check{\tilde{H}}_{t+\tau}(X_t)]|\mathcal{F}_t])\Bigg]dt+\delta\Delta X_T(\partial_{x'}\check{g}'_T-\partial_x\check{g}_T+\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}'_T(\check{X}_T)]-\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}_T(\check{X}_T)])$$

$$\leq\frac{1}{2}\mathbb{E}\int_0^T\Bigg[\epsilon(|\Delta X_t|^2+|\Delta Y_t|^2+|\Delta Z_t|^2)+\frac{1}{\epsilon}\delta^2\Big(|\Delta\check{b}_t|^2+|\Delta\check{\sigma}|^2$$

$$+|\Delta\partial_x\check{H}_t+\Delta\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{H}}_t(X_t)]+\Delta\mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t]+\Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\check{\tilde{H}}_{t+\tau}(X_t)]|\mathcal{F}_t]|^2\Big)\Bigg]dt$$

$$+\frac{1}{2}\Bigg(\epsilon|\Delta X_T|^2+\frac{1}{\epsilon}\delta^2\Big|\partial_{x'}\check{g}'_T-\partial_x\check{g}_T+\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}'_T(\check{X}_T)]-\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{g}}_T(\check{X}_T)]\Big|^2\Bigg)$$

$$\leq\frac{1}{2}\epsilon\mathbb{E}\Bigg[\int_0^T\epsilon(|\Delta X_t|^2+|\Delta Y_t|^2+|\Delta Z_t|^2+|\Delta\alpha_t|^2)dt+|\Delta X_T|^2\Bigg]$$

$$+\frac{1}{2}\delta\frac{C}{\epsilon}\mathbb{E}\Bigg[\int_0^T(|\Delta\check{X}_t|^2+|\Delta\check{Y}_t|^2+|\Delta\check{Z}_t|^2+|\Delta\check{\alpha}_t|^2)]dt+|\Delta\check{X}_T|^2\Bigg],$$

(5.15)

where we have used simplified notation:

$$\Delta b_t:=b(t,\theta_t,\theta_{t-\tau})-b(t,\theta'_t,\theta'_{t-\tau});\quad\Delta\check{b}_t:=b(t,\check{\theta}_t,\check{\theta}_{t-\tau})-b(t,\check{\theta}'_t,\check{\theta}'_{t-\tau});$$

$$\Delta\sigma_t:=\sigma(t,\theta_t,\theta_{t-\tau})-\sigma(t,\theta'_t,\theta'_{t-\tau});\quad\Delta\check{\sigma}_t:=\sigma(t,\check{\theta}_t,\check{\theta}_{t-\tau})-\sigma(t,\check{\theta}'_t,\check{\theta}'_{t-\tau})$$

$$\partial_xg_T:=\partial_xg(X_T,\mu_T)-\partial_xg(X'_T,\mu_T);$$

$$\Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{g}_T(X_T)]:=\tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}_T,\mu_T)(X_T)]-\tilde{\mathbb{E}}[\partial_\mu g(\tilde{X}'_T,\mu_T)(X'_T)]$$

$$\Delta\partial_xH_t:=\partial_xH(t,\theta_t,\theta_{t-\tau},Y_t,Z_t)-\partial_xH(t,\theta'_t,\theta'_{t-\tau},Y_t,Z_t)$$

$$\Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)]:=\tilde{\mathbb{E}}[\partial_\mu H(t,\tilde{\theta}_t,\tilde{\theta}_{t-\tau},\tilde{Y}_t,\tilde{Z}_t)(X_t)]-\tilde{\mathbb{E}}[\partial_\mu H(t,\tilde{\theta}'_t,\tilde{\theta}'_{t-\tau},\tilde{Y}_t,\tilde{Z}_t)(X'_t)]$$

$$\Delta\mathbb{E}[\partial_{x_\tau}H_{t+\tau}|\mathcal{F}_t]:=\mathbb{E}[\partial_{x_\tau}H(t+\tau,\theta_{t+\tau},\theta_t,Y_{t+\tau},Z_{t+\tau})|\mathcal{F}_t]-\mathbb{E}[\partial_{x_\tau}H(t+\tau,\theta'_{t+\tau},\theta'_t,Y_{t+\tau},Z_{t+\tau})|\mathcal{F}_t]$$

$$\Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)]|\mathcal{F}_t]:=\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}H(t+\tau,\tilde{\theta}_{t+\tau},\tilde{\theta}_t,\tilde{Y}_{t+\tau},\tilde{Z}_{t+\tau})(X_t)]|\mathcal{F}_t]$$

$$-\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}H(t+\tau,\tilde{\theta}'_{t+\tau},\tilde{\theta}'_t,\tilde{Y}_{t+\tau},\tilde{Z}_{t+\tau})(X'_t)]|\mathcal{F}_t]$$

similar notation for $\Delta\partial_x\check{H}_t,\ \Delta\tilde{\mathbb{E}}[\partial_\mu\check{\tilde{H}}_t(X_t)],\ \Delta\mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t],\ $ and $\Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\check{\tilde{H}}_{t+\tau}(X_t)]|\mathcal{F}_t].$

(5.16)

Next, we apply Ito's formula to ΔX_t^2,

$$
\begin{aligned}
&d\Delta X_t^2 \\
=&2\Delta X_t dX_t + d\langle X, X\rangle_t \\
=&2\Delta X_t(\lambda_0\Delta b_t + \delta\Delta\check{b}_t)dt + 2\Delta X_t(\lambda_0\Delta\sigma_t + \delta\Delta\check{\sigma}_t)dW_t + \left(\lambda_0\Delta\sigma_t + \delta\Delta\check{\sigma}_t\right)^2 dt
\end{aligned}
\tag{5.17}
$$

Then integrate from 0 to T, and take expectation,

$$
\begin{aligned}
&\mathbb{E}[|\Delta X_t|^2] \\
=&2\lambda_0\mathbb{E}\int_0^t |\Delta X_s\Delta b_s|ds + 2\delta\mathbb{E}\int_0^t |\Delta X_s\Delta\check{b}_s|ds + \mathbb{E}\int_0^t |\lambda_0\Delta\sigma_s + \delta\Delta\check{\sigma}_s|^2 ds \\
\leq&\lambda_0\mathbb{E}\int_0^t (|\Delta X_s|^2 + |\Delta b_s|^2)ds + \mathbb{E}\int_0^t (|\Delta X_s|^2 + \delta^2|\Delta\check{b}_s|^2)ds \\
&+ \mathbb{E}\int_0^t (2\lambda_0^2|\Delta\sigma_s|^2 + 2\delta^2|\Delta\check{\sigma}_s|^2)ds \\
\leq&C\mathbb{E}\int_0^{t+\tau} (|\Delta X_s|^2 + |\Delta\alpha_s|^2)ds + \delta C\mathbb{E}\int_0^{t+\tau} (|\Delta\check{X}_s|^2 + |\Delta\check{\alpha}_s|^2)ds
\end{aligned}
\tag{5.18}
$$

From Gronwall's inequality, we can obtain

$$
\sup_{0\leq t\leq T}\mathbb{E}[|X_t|^2] \leq C\mathbb{E}\int_0^T |\Delta\alpha_t|^2 dt + \delta C\mathbb{E}\int_0^T (|\Delta\check{X}_t|^2 + |\Delta\check{\alpha}_t|^2)dt
\tag{5.19}
$$

Similarly, applying Ito's formula to $|\Delta Y_t|^2$, and taking expectation, we have

$$
\begin{aligned}
&\mathbb{E}\left[|\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2 ds\right] \\
=&2\lambda_0\mathbb{E}\int_t^T \left|\Delta Y_t\left(\Delta\partial_x H_t + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)] + \Delta\mathbb{E}[\partial_{x_\tau}H_{t+\tau}|\mathcal{F}_t] + \Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t]\right)\right| \\
&+ 2\delta\mathbb{E}\int_t^T \left|\Delta Y_t\left(\Delta\partial_x\check{H}_t + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{\check{H}}_t(\check{X}_t)] + \Delta\mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t] + \Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{\check{H}}_{t+\tau}(\check{X}_t)]|\mathcal{F}_t]\right)\right| \\
&+ \mathbb{E}|\Delta Y_T|^2 \\
\leq&\mathbb{E}\int_t^T \left(\frac{1}{\epsilon}|\Delta Y_t|^2 + \epsilon\lambda_0^2\left|\Delta\partial_x H_t + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{H}_t(X_t)] + \Delta\mathbb{E}[\partial_{x_\tau}H_{t+\tau}|\mathcal{F}_t] + \Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{H}_{t+\tau}(X_t)]|\mathcal{F}_t]\right|^2\right)dt \\
&+ \mathbb{E}\int_t^T \left(|\Delta Y_t|^2 + \delta^2\left|\Delta\partial_x\check{H}_t + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{\check{H}}_t(X_t)] + \Delta\mathbb{E}[\partial_{x_\tau}\check{H}_{t+\tau}|\mathcal{F}_t] + \Delta\mathbb{E}[\tilde{\mathbb{E}}[\partial_{\mu_\tau}\tilde{\check{H}}_{t+\tau}(X_t)]|\mathcal{F}_t]\right|^2\right)dt \\
&+ \mathbb{E}\left|\lambda_0\left(\Delta\partial_x g_T + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{g}_T(X_T)]\right) + \delta\left(\Delta\partial_x\check{g}_T + \Delta\tilde{\mathbb{E}}[\partial_\mu\tilde{\check{g}}_T(X_T)]\right)\right|^2
\end{aligned}
\tag{5.20}
$$

Choose $\epsilon = 96 \max\{R^2, L\}$, and from assumption (H5.1 - H5.2) and Gronwall's inequality, we obtain a bound for $\sup_{0 \le t \le T} \mathbb{E}|\Delta Y_t|^2$; and then substitute the it back to the same inequality, we are able to obtain the bound for $\int_0^T E|Z_t|^2 dt$. By combining these two bounds, we deduce that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0 \le t \le T} |Y_t|^2 + \int_0^T |Z_t|^2 dt\right] \\
&\le C\mathbb{E}\left(\sup_{0 \le t \le T} |\Delta X_t|^2 + \int_0^T |\Delta \alpha_t|^2 dt\right) + \delta C \mathbb{E}\left[\sup_{0 \le t \le T}\left(|\Delta \check{X}_t|^2 + |\Delta \check{Y}_t|^2\right) + \int_0^T \left(|\Delta \check{Z}_t|^2 + |\Delta \check{\alpha}_t|^2\right) dt\right]
\end{aligned}
\tag{5.21}
$$

Finally, combining (5.19) and (5.21), and (5.15), we deduce

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0 \le t \le T} |\Delta X_t|^2 + \sup_{0 \le t \le T} |\Delta Y_t|^2 + \int_0^T \left(|\Delta Z_t^2| + |\Delta \alpha_t|^2\right) dt\right] \\
&\le \delta C \mathbb{E}\left[\sup_{0 \le t \le T} |\Delta \check{X}_t|^2 + \sup_{0 \le t \le T} |\Delta \check{Y}_t|^2 + \int_0^T \left(|\Delta \check{Z}_t|^2 + |\Delta \check{\alpha}_t|^2\right) dt\right]
\end{aligned}
\tag{5.22}
$$

Let $\delta_0 = \frac{1}{2C}$, it is clear that the mapping Φ is a contraction for all $\delta \in (0, \delta_0)$. It follows that there is a unique fixed point which is the solution of MV-FABSDE(λ) for $\lambda = \lambda_0 + \delta, \ \delta \in (0, \delta_0)$. $\qquad \square$

Proof of Theorem 5.1. For $\lambda = 0$, FABSDE(0) has a unique solution. Using Lemma 5.2, there exists a $\delta_0 > 0$ such that FBSDE(δ) has a unique solution for $\delta \in [0, \delta_0]$, assuming $(n-1)\delta_0 < 1 \le n\delta_0$. Following by a induction argument, we repeat Lemma 5.2 for n times, which gives us the existence of the unique solution of FABSDE(1). $\qquad \square$

References

[1] A. Bensoussan, S.C.P. Yam, and Z. Zhang. Well-posedness of mean-field type forward–backward stochastic differential equations. *Stochastic Processes and their Applications*, 125:3327–3354, 2015.

[2] Alain Bensoussan, Jens Frehse, and Phillip Yam. *Mean Field Games and Mean Field Type Control Theory*. Springer-Verlag New York, 2013.

[3] Rene Carmona and Francois Delarue. *Probabilistic Theory of Mean Field Games with Applications I & II*. Springer International Publishing, 2018.

[4] Rene Carmona, Francois Delarue, and Aime Lachapelle. Control of mckean–vlasov dynamics versus mean field games. *Mathematics and Financial Economics*, 7(2):131–166, 2013.

[5] Rene Carmona, Jean-Pierre Fouque, Seyyed Mostafa Mousavi, and Li-Hsien Sun. Systemic risk and stochastic games with delay. *Journal of Optimization and Applications (JOTA)*, 2018.

[6] Rene Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. Mean field games and systemic risk. *Communications in Mathematical Sciences*, 13(4):911–933, 2015.

[7] W. E, J. Han, and A. Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *arXiv e-prints*, June 2017.

[8] Jean-Pierre Fouque and Zhaoyu Zhang. Mean field game with delay: a toy model. *Risks*, 6,90:1–17, 2018.

[9] Sepp Hochreiter and Jurgen Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[10] Jin Ma, Philip Protter, and Jiongmin Yong. Solving forward-backward stochastic differential equations explicitly — a four step scheme. *Probab. Th. Rel. Fields*, 98:339–359, 1994.

[11] Jin Ma and Jiongmin Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. Springer-Verlag Berlin Heidelberg, 2007.

[12] Shige Peng and Zhen Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3):825–843, 1999.

[13] Shige Peng and Zhe Yang. Anticipated backward stochastic differential equations. *Ann. Probab.*, 37:877–902, 2009.

[14] M. Raissi. Forward-Backward Stochastic Neural Networks: Deep Learning of High-dimensional Partial Differential Equations. *arXiv e-prints*, April 2018.

[15] R.S. Sutton, D.A. McAllester, S.P. Singh, and Y. Mansour. Policy gradient methods for reinforcement learning with function approximation. *In Advances in Neural Information Processing Systems (NIPS)*, pages 1057–1063, 1999.

[16] Jianfeng Zhang. *Backward Stochastic Differential Equations*. Springer-Verlag New York, 2017.