



# Deep Neural Network Solution for Finite State Mean Field Game with Error Estimation

Jialiang Luo<sup>1</sup> · Harry Zheng<sup>1</sup>

Accepted: 27 September 2022 / Published online: 25 October 2022  
© The Author(s) 2022

## Abstract

We discuss the numerical solution to a class of continuous time finite state mean field games. We apply the deep neural network (DNN) approach to solving the fully coupled forward and backward ordinary differential equation system that characterizes the equilibrium value function and probability measure of the finite state mean field game. We prove that the error between the true solution and the approximate solution is linear to the square root of DNN loss function. We give an example of applying the DNN method to solve the optimal market making problem with terminal rank-based trading volume reward.

**Keywords** Finite state mean field game · Forward backward ODE · Deep neural network · Error estimation

**Mathematics Subject Classification** 93E20 · 90C39

**JEL Classification** C7 · G1

## 1 Introduction

There has been extensive research in mean field game (MFG) since its introduction by Lasry and Lions [21] and Huang et al. [18] as a limit of symmetric nonzero sum noncooperative  $N$ -player dynamic games when the number of players tends to infinite, see Carmona and Delarue [4] and Guéant et al. [16] for excellent introduction to MFG. The main research focus nowadays is in two areas. One is to study the existence and uniqueness of MFG equilibrium with the partial differential equation (PDE) system that characterizes the equilibrium value function and mean field state, see Lasry and Lions [21]. The other is to analyze the convergence from the stochastic differential game among large but finite number of players to

---

✉ Harry Zheng  
h.zheng@imperial.ac.uk

Jialiang Luo  
jjialiang.luo14@imperial.ac.uk

<sup>1</sup> Department of Mathematics, Imperial College, London SW7 2AZ, UK

the MFG limit when the number of players tends to infinity and the numerical approximation for MFG, see Achdou et al. [1] and Achdou and Capuzzo-Dolcetta [2]. The MFG theory has been applied to many modeling problems in economics, politics, social welfare and other areas, see, for example, Guéant ([14] and Lasry et al. [22]).

In this paper, we focus on MFG with finite time horizon and continuous time state dynamic of each agent taking values in a finite set under fully symmetric payoff and complete information. Gomes et al. [12] first study finite state MFG and prove the existence and uniqueness of Nash equilibrium with the coupled forward and backward ordinary differential equation (FBODE) system and show the convergence of  $N$ -player game's Nash equilibrium to that of the limiting MFG when  $N$  tends to infinite and time horizon is small. [8] analyze the MFG with a probabilistic approach. Carmona and Wang [7] tackle both the mean field of states and that of controls and prove the existence of equilibrium with backward stochastic differential equation and the uniqueness of equilibrium when the Hamiltonian does not depend on mean field controls. Carmona and Wang [6] analyze finite state MFG between one major player and infinite number of minor players. Cardaliaguet et al. [3] make the breakthrough in convergence analysis for a diffusion model with common noise and characterize the equilibrium with the master equation and its regular solution. Cecchin and Pelino [9] apply the master equation to obtain the convergence of feedback Nash equilibrium in the finite state space, which extends the convergence result in Gomes et al. [12] without requiring the time horizon being sufficiently small.

Despite the progress in existence, uniqueness and convergence for Nash equilibrium of the finite state MFG, there is still a considerable obstacle to approximate the  $N$ -player game with a simpler MFG. One main difficulty is that the Nash equilibrium of finite state MFG is characterized by a FBODE system with both the initial and terminal conditions, which in generally has no analytical solution and is difficult to solve numerically. One commonly used method for solving FBODE is the shooting method but it tends to work better when the dimension is low and the boundary condition is simple. The shooting method fails to work in our case. Gomes and Saude [13] propose a numerical scheme to solve finite state MFG under some monotone conditions that do not hold in many applications.

There has been active research in recent years on using the deep neural network (DNN) to solve PDEs and ODEs with different boundary conditions, see, e.g., Lagaris et al. [19], Malek and Beidokhti [27], Lee and Kang [24], Lagaris et al. [20]. Given that the feature of FBODE system is similar to that of a PDE, we are motivated to use DNN to numerically solve the FBODE system in the finite state MFG problem. Sirignano and Spiliopoulos [30] propose DGM (deep Galerkin method) to solve high-dimensional PDEs with a mesh-free DNN and show the convergence of approximate solutions to the true solution under some conditions, which is similar in spirit to the Galerkin method except that the solution is approximated by a neural network instead of a linear combination of basis functions. [5] provide a comprehensive literature review on deep learning method applied on MFG. Many papers apply the DGM approach to numerically solve high-dimensional PDEs derived from different types of MFG (see, e.g., Han et al. [17], Ruthotto et al. [29]), while others apply the DNN to solve the corresponding BSDEs (see, e.g., Fouque and Zhang [11], Lauriere [23]). Most of these papers only provide numerical results without rigorous proof for the numerical solutions. There is no guarantee that the neural network approximation can converge to the true solution, and the approximation may not be accurate enough albeit the loss function is already small as there is no relation established between the loss function and the error between approximate and true solutions. Li et al. [25] and Li et al. [26] prove the strong and uniform convergence of the DGM approach. Parallel to this paper, Mishra and Molinaro [28] focus on error estimation of physical informed neural network (PINN), which is the other name of

DGM. For PDEs satisfying certain conditions, they provide the abstract framework to relate the loss function of neural network to the error between the true solution and the approximate solution generated by neural network and prove the error bound for several specific types of PDEs. Their assumptions on the regularity of PDEs are strong (see Assumption 2.1 in Mishra and Molinaro [28]) and are not necessarily satisfied by the FBODE system derived in this paper. To our best knowledge, there is no existing literature addressing the error between the approximation and the true solution via the loss function for FBODEs derived from continuous time finite state MFG problems. We provide the error bound estimation to fill the gap.

The main result of the paper, Theorem 2.6, states that the error between the true solution  $(\theta, p)$  of the FBODE system and the DNN approximate solution  $(\tilde{\theta}, \tilde{p})$  is linear to the square root of the loss function in the DNN method, which provides the magnitude of the error bound for the DNN approximation as well as the convergence result. To bridge  $\theta$  and  $\tilde{\theta}$ , we use the master equation for  $\theta$  in Cecchin and Pelino [9] and prove that  $\tilde{\theta}$  satisfies a similar equation. Cecchin and Pelino [9] prove the equilibrium of finite players finite state game converges to that of the corresponding MFG with the former satisfying a backward ODE while the latter a FBODE which is equivalent to a backward PDE (master equation) and can be compared with the backward ODE system. In contrast, we want to estimate the error between the true solution and the DNN approximation to MFG with both satisfying FBODE systems and the one for the DNN approximation having extra error terms compared with the one for the true solution. We leverage the master equations to connect the two FBODE systems and do error analysis. Due to perturbation terms in the FBODE system, we need to address the issue of negative  $\tilde{p}$ , prohibited in [9, Theorem 6] and find a new way to bypass that difficulty.

As an application, we apply the DNN to numerically solve an optimal market making problem with the same framework as that of Guéant [15], except that the terminal reward depends on the trading volume ranking that is determined in a so-called market maker incentive program contract designed by the exchange to encourage market maker to provide more liquidity (i.e., trading volume). El Euch et al. [10] discuss the market maker incentive contract and analyze how exchange should optimally decide the commission fee schedule for market makers. The trading volume ranking-related reward, commonly seen in market incentive programs from various exchanges, is not considered in El Euch et al. [10]. In this paper, we use a finite state MFG to model the competition between market makers in the presence of the trading volume ranking reward and solve the Nash equilibrium using the DNN approach. The results may help exchanges design better market incentive program by better understanding market makers' behavior responding to the contract.

The rest of the paper is organized as follows. In Sect. 2, we formulate the finite state MFG model and state the main result of the paper, Theorem 2.6, on the error estimation of DNN approximate solution. In Sect. 3, we discuss an application in the optimal market making with rank-based trading volume reward. Section 4 contains the proofs of Theorems 2.4, 2.5, 2.6 and Proposition 3.1. Section 5 concludes.

## 2 Model and Main Results

Define a finite state MFG in continuous time similar to the one in Cecchin and Pelino [9]. The finite state space is  $\Sigma = \{1, \dots, K\}$ , and the reference game player's state is denoted by  $Z$ , which is a Markov chain. The player at state  $z$  can decide the switching intensities with feedback controls  $\lambda : [0, T] \times \Sigma \rightarrow (\mathbb{R}^+)^K$  from  $\Sigma$  to  $(\mathbb{R}^+)^K$ . The dynamic for the player

is given by

$$dZ_t = \sum_{z \in \Sigma} dN_t^z,$$

where  $N_t^z$  is a Poisson process with controlled intensity  $\lambda_z(t, Z_t)$ .

If there are some states that state  $z$  cannot access, then we can simply set the corresponding components in the intensity vector to be zero. The probability measure on mean field of states is a function  $p : [0, T] \rightarrow P(\Sigma)$ , where

$$P(\Sigma) = \{(p_1, \dots, p_K) : \sum_{z=1}^K p_z = 1, p_z \geq 0\}.$$

Start at time  $t \in [0, T]$ , given any probability measure  $p$  on the mean field of state, game player with controlled state process  $Z_t$  that start at state  $z$  wants to optimize

$$\theta_z(t) := \sup_{\lambda \in \mathcal{A}} \mathbb{E}_t \left[ \int_t^T F(Z_t, \lambda(t, Z_t)) dt + G(Z_T, p(T)) \right], \tag{2.1}$$

where  $\mathbb{E}_t[\cdot]$  is the conditional expectation given the initial state  $Z_t = z$  at time  $t$ ,  $F$  the running payoff,  $G$  the terminal payoff and  $\mathcal{A}$  the admissible control set containing all measurable functions  $\lambda : [0, T] \times \Sigma \rightarrow (\mathbb{R}^+)^K$ . We assume for any  $z \in \Sigma$ ,  $F(z, \lambda)$  is an upper bounded function which does not depend on  $\lambda_z$ , the  $z$ th component of  $\lambda$ . Define  $\theta : [0, T] \rightarrow \mathbb{R}^K$  by  $\theta(t) = (\theta_1(t), \dots, \theta_K(t))$ . According to Cecchin and Pelino [9], the equilibrium value function  $\theta$  and the mean field probability  $p$  satisfy the following FBODE system:

$$\begin{aligned} \frac{d\theta_z(t)}{dt} &= -H(z, \Delta^z \theta(t)), \quad \theta_z(T) = G(z, p(T)), \\ \frac{dp_z(t)}{dt} &= \sum_y p_y(t) \lambda_z^*(y, \Delta^y \theta(t)), \quad p_z(t_0) = p_{z,0}, \end{aligned} \tag{2.2}$$

where  $\Delta^z$  is the difference operator, defined as

$$\Delta^z \theta(t) := (\theta_1(t) - \theta_z(t), \dots, \theta_K(t) - \theta_z(t))$$

and  $H : \Sigma \times \mathbb{R}^K \rightarrow \mathbb{R}$  is the Hamiltonian function, defined for any  $\mu \in \mathbb{R}^K$  satisfying  $\mu_z = 0$  as

$$H(z, \mu) := \sup_{\lambda \in (\mathbb{R}^+)^K} \{\lambda \cdot \mu + F(z, \lambda)\}$$

and  $\lambda^*(z, \mu) = (\lambda_1^*(z, \mu), \dots, \lambda_K^*(z, \mu))$  is the optimizer of  $H(z, \mu)$  except for  $\lambda_z^*(z, \mu)$ , which can be any value since in the proof of our main result we always let  $\mu_z = [\Delta^z \theta(t)]_z = \theta_z(t) - \theta_z(t) = 0$  and  $F(z, \lambda)$  is independent to  $\lambda_z$ . For notation convenience, we define

$$\lambda_z^*(z, \mu) := - \sum_{y \neq z} \lambda_y^*(z, \mu). \tag{2.3}$$

The backward equation in (2.2) comes from the optimization problem (2.1) given  $p$  and the forward equations comes from the consistent condition for probability measure  $p$  on mean field of states when everyone follows the equilibrium strategy. According to [12, Proposition 1], if  $H$  is differentiable and  $\lambda^*(z, \mu)$  is positive except the  $z$ th element, for  $y \neq z$ , we have

$$\lambda_y^*(z, \mu) = [D_\mu H(z, \mu)]_y,$$

where  $\lambda_y^*(z, \mu)$  is the intensity from state  $z$  to state  $y$  and  $[D_\mu H(z, \mu)]_y$  the  $y$ th component of the gradient  $D_\mu H(z, \mu)$ . In the proofs of main results, we always have  $\mu_z = 0$  when we use  $H(z, \mu)$ ,  $D_\mu H(z, \mu)$  or  $D_{\mu\mu}^2 H(z, \mu)$ , the Hessian matrix. For proof simplicity, with a little abuse of notation, we follow Cecchin and Pelino [9] to define artificially that

$$[D_\mu H(z, \mu)]_z = \lambda_z^*(z, \mu).$$

Then we can conclude that

$$\lambda^*(z, \mu) = D_\mu H(z, \mu),$$

and the feedback control  $\lambda(t, z) = \lambda^*(z, \Delta^z \theta(t))$  under the equilibrium.

We next assume  $H$ ,  $G$  and  $\lambda^*$  satisfy the following assumptions.

**Assumption 2.1** Assume that, under (2.3),  $H(z, \mu)$  has a unique optimizer  $\lambda^*(z, \mu)$  for every  $\mu$ .  $H$  is  $C^2$  in  $\mu$  on bounded set,  $H$ ,  $D_\mu H$  and  $D_{\mu\mu}^2 H$  are locally Lipschitz in  $\mu$ , and the second derivatives is bounded away from 0 on bounded set, i.e., there exists a constant  $C$  such that for any  $\mu$  in that bounded set satisfying  $\mu_z = 0$ , we have

$$C^{-1} \|\mu\|^2 \leq \mu \cdot D_{\mu\mu}^2 H(z, \mu) \cdot \mu \leq C \|\mu\|^2. \tag{2.4}$$

Moreover, assume that  $G$  is differentiable, and there exists a constant  $C$  such that when  $p$  is bounded, its directional derivative in any vector  $w$  satisfies

$$\left| \frac{\partial G}{\partial w}(z, p + \Delta p) - \frac{\partial G}{\partial w}(z, p) \right| \leq C \|\Delta p\| \|w\| \tag{2.5}$$

and that  $G$  is decreasing in  $p$ , i.e., for all  $p, \bar{p} \in \mathbb{R}^K$ ,

$$\sum_{z \in \Sigma} (G(z, p) - G(z, \bar{p}))(p_z - \bar{p}_z) \leq 0. \tag{2.6}$$

**Remark 2.2**  $H$  satisfies  $H(z, \mu) \geq H(z, \bar{\mu})$  for any  $z \in \Sigma$  if two vectors  $\mu = (\mu_1, \dots, \mu_K)$  and  $\bar{\mu} = (\bar{\mu}_1, \dots, \bar{\mu}_K)$  satisfy

$$\mu_i \geq \bar{\mu}_i, \quad i \in \Sigma, i \neq z.$$

Then from [12, Proposition 2], solution to (2.2) has a prior bound  $C_{GH}$  as long as  $H$  satisfies Remark 2.2 and  $G$  is bounded for all  $p(T)$  in compact set  $[0, 1]^K$ .  $C_{GH}$  is defined as,

$$\|\theta\| \leq C_{GH} := \max_{z \in \Sigma, p \in [0, 1]^K} \{G(z, p)\} + 2 \max_{z \in \Sigma} H(z, 0)T, \tag{2.7}$$

where the norm  $\|\cdot\|$  is the  $\infty$  norm.  $G$  is bounded because it is continuous and defined on a compact set. For given  $H$  and  $G$ , as  $\theta$  satisfies ODE system (2.2), and both  $H$  is Lipschitz continuous in Assumption 2.1,  $\frac{d\theta_z(t)}{dt}$  is also bounded. Similarly, as  $D_\mu H$  and  $\frac{d\theta_z(t)}{dt}$  are bounded, we can further see that  $\frac{d^2\theta_z(t)}{dt^2}$  is bounded. From similar argument on  $p$  and  $\lambda^*$ ,  $\frac{dp_z(t)}{dt}$  and  $\frac{d^2p_z(t)}{dt^2}$  are also bounded. It means for given  $H$  and  $G$ , there exists constants  $C_{\theta GH}$  and  $C_{pGH}$ , such that

$$\begin{aligned} \left\| \frac{d\theta_z(t)}{dt} \right\| &\leq C_{\theta GH}, & \left\| \frac{d^2\theta_z(t)}{dt^2} \right\| &\leq C_{\theta GH}, \\ \left\| \frac{dp_z(t)}{dt} \right\| &\leq C_{pGH}, & \left\| \frac{d^2p_z(t)}{dt^2} \right\| &\leq C_{pGH}. \end{aligned} \tag{2.8}$$

We summarize [9, Theorem 1], [12, Theorem 2], and state the following theorem without proof.

**Theorem 2.3** *Under Assumption 2.1, ODE system (2.2) has a unique solution  $(\theta, p)$  for any initial value  $p(t_0) \in P(\Sigma)$  and the MFG has a unique Nash equilibrium point.*

We assume in the rest of the paper that Assumption 2.1 holds, which guarantees the existence, uniqueness and convergence of the finite state MFG. However, to find the equilibrium, we need to solve (2.2), which generally does not have analytical solution. As (2.2) is a FBODE system, we cannot solve it numerically by simple discretization. We apply the deep neural network approach in Sirignano and Spiliopoulos [30] to numerically solve (2.2).

Define two sets of neural network functions as

$$\begin{aligned} \Theta^n(v_1, v) &:= \left\{ \tilde{\theta} : [0, T] \rightarrow \mathbb{R}^K; \tilde{\theta}(t) \right. \\ &= \left. \left( v_1 \left( \sum_{i=1}^n \beta_{1,i} v(\alpha_i t + c_i) \right) \right), \dots, v_1 \left( \sum_{i=1}^n \beta_{K,i} v(\alpha_i t + c_i) \right) \right\}, \\ \mathbf{P}^n(v_2, v) &:= \left\{ \tilde{p} : [0, T] \rightarrow \mathbb{R}^{K-1}; \tilde{p}(t) \right. \\ &= \left. \left( v_2 \left( \sum_{i=n+1}^{2n} \beta_{1,i} v(\alpha_i t + c_i) \right) \right), \dots, v_2 \left( \sum_{i=n+1}^{2n} \beta_{K-1,i} v(\alpha_i t + c_i) \right) \right\}, \end{aligned}$$

where  $v : \mathbb{R} \rightarrow \mathbb{R}$  is the triple continuously differentiable activation function, and two strictly increasing triple continuously differentiable activation functions  $v_1, v_2 : \mathbb{R} \rightarrow \mathbb{R}$  have twice continuously differentiable inverse functions  $v_1^{-1}$  and  $v_2^{-1}$ . They satisfy

$$\sup |v_1| = C_{GH} + e, \quad \inf v_2 = -e, \quad \sup v_2 = 1 + e, \tag{2.9}$$

where  $e$  is any constant that is small.

In the numerical tests of this paper, we use hyperbolic tangent functions  $\tanh$  for activation functions, in particular,  $v_1(x) = a \tanh x + b$  for some constants  $a, b$  and  $v_2(x) = (\tanh x + 1)/2$ . We approximate the solution  $(\theta, p)$  to (2.2) numerically by  $(\tilde{\theta}^{(N)}, \tilde{p}^{(N)})$  which satisfies

$$(\tilde{\theta}_1^{(N)}, \dots, \tilde{\theta}_K^{(N)}) \in \Theta^N(v_1, v), \quad (\tilde{p}_1^{(N)}, \dots, \tilde{p}_{K-1}^{(N)}) \in \mathbf{P}^N(v_2, v), \quad \tilde{p}_K^{(N)} = 1 - \sum_{i \neq K} \tilde{p}_i^{(N)}. \tag{2.10}$$

By considering both the differential operators and boundary conditions in (2.2), we define the loss function  $\Psi$  for any approximate solution  $(\tilde{\theta}, \tilde{p})$  as

$$\begin{aligned} \Psi(\tilde{\theta}, \tilde{p}) &:= \sum_{z \in \Sigma} \left\{ \int_{t_0}^T \left( \frac{d\tilde{\theta}_z(t)}{dt} + H(z, \Delta^z \tilde{\theta}(t)) \right)^2 dt \right. \\ &+ \int_{t_0}^T \left( \frac{d\tilde{p}_z(t)}{dt} - \sum_y \tilde{p}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \right)^2 dt + \int_{t_0}^T \left( \sum_z (\tilde{p}_z(t))^- \right)^2 dt \\ &+ (\tilde{p}_z(t_0) - \tilde{p}_{z,0})^2 + (\tilde{\theta}_z(T) - G(z, \tilde{p}(T)))^2 \\ &\left. + \sum_{z \in \Sigma} \left( B_\theta - \max_{t \in [0, T]} \left| \frac{d^2 \tilde{\theta}_z(t)}{dt^2} \right| \right)^- + \sum_{z \in \Sigma} \left( B_p - \max_{t \in [0, T]} \left| \frac{d^2 \tilde{p}_z(t)}{dt^2} \right| \right)^- \right\}, \end{aligned} \tag{2.11}$$

where  $(\tilde{p}_K(t))^- := -\tilde{p}_K(t)\mathbb{1}_{\{p_K(t) \leq 0\}}$  and  $B_\theta, B_p$  can be any constants that satisfy

$$B_\theta > C_{\theta GH} \geq \max_{t \in [0, T]} \left| \frac{d^2 \theta_z(t)}{dt^2} \right|,$$

$$B_p > C_{pGH} \geq \max_{t \in [0, T]} \left| \frac{d^2 p_z(t)}{dt^2} \right|,$$

where constants  $C_{\theta GH}$  and  $C_{pGH}$  are from (2.8). Then it follows

$$\sum_{z \in \Sigma} \left( B_\theta - \max_{t \in [0, T]} \left| \frac{d^2 \theta_z(t)}{dt^2} \right| \right)^- + \sum_{z \in \Sigma} \left( B_p - \max_{t \in [0, T]} \left| \frac{d^2 p_z(t)}{dt^2} \right| \right)^- = 0.$$

Both the integral term and maximum term in (2.11) can be calculated via Monte Carlo simulation. Practically, we use similar approach as in Sirignano and Spiliopoulos [30] to calculate these two to increase the robustness of training. Given  $N$ , the structure of the neural network has been determined. We train the network by finding the optimal values of  $\{\beta_{j,i}\}_{i,j=1}^{2K-1, 2n}, \{\alpha_i\}_{i=1}^{2n}$  and  $\{c_i\}_{i=1}^{2n}$  that determine  $(\tilde{\theta}^{(N)}, \tilde{p}^{(N)})$  such that they minimize  $\Psi$ . For the true solution  $(\theta, p)$ ,  $\Psi(\theta, p) = 0$ . Since  $(\theta, p)$  exists and is unique,  $\Psi$  has unique minimal point  $\Psi(\theta, p) = 0$ . We provide the convergence result Theorem 2.4 similar to the Theorem 7.1 in Sirignano and Spiliopoulos [30].

**Theorem 2.4** *There exists a sequence of  $(\tilde{\theta}^{(N)}, \tilde{p}^{(N)})$  defined in (2.10) such that*

$$\lim_{N \rightarrow +\infty} \Psi(\tilde{\theta}^{(N)}, \tilde{p}^{(N)}) = 0.$$

**Proof** See Sect. 4.

When the Loss function  $\Psi$  is smaller than certain value, because of the uniform bounds on the first and second derivatives of the approximation function, the maximum error on the time interval is also smaller than certain value.

**Theorem 2.5** *There exists constant  $\varepsilon_0$ , such that for any  $\varepsilon < \varepsilon_0$ , if  $\Psi < \varepsilon$  then there exists constant  $C$  such that for all  $t \in [t_0, T]$  and  $z \in \Sigma$ , we have*

$$\begin{aligned} \left\| \frac{d\tilde{\theta}_z(t)}{dt} + H(z, \Delta^z \tilde{\theta}(t)) \right\| &< C\varepsilon^{\frac{1}{3}}, \quad \left\| \tilde{\theta}_z(T) - G(z, \tilde{p}(T)) \right\| < C\varepsilon^{\frac{1}{2}}, \\ \left\| \frac{d\tilde{p}_z(t)}{dt} - \sum_y \tilde{p}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \right\| &< C\varepsilon^{\frac{1}{3}}, \quad \left\| \tilde{p}_z(t_0) - p_{z,0} \right\| < C\varepsilon^{\frac{1}{2}}, \end{aligned} \tag{2.12}$$

where  $\|\cdot\|$  is the infinity norm.

**Proof** See Sect. 4.

Note that the constant  $C$  in Theorem 2.5 depends on the FBODE system and the bound of its true solution, but is independent of the DNN structure. Theorem 2.5 is an algorithm independent result. We now state our main result on the error estimation for numerical solution to the finite state mean field game.

**Theorem 2.6** *For every  $t \in [t_0, T]$  and  $z \in \Sigma$ , assume  $\tilde{\theta}(t)$  and  $\tilde{p}(t)$  satisfy:*

$$\begin{aligned} \frac{d\tilde{\theta}_z(t)}{dt} &= -H(z, \Delta^z \tilde{\theta}(t)) + \epsilon_1(t, z), \quad \tilde{\theta}_z(T) = G(z, \tilde{p}(T)) + \epsilon_3(z), \\ \frac{d\tilde{p}_z(t)}{dt} &= \sum_y \tilde{p}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) + \epsilon_2(t, z), \quad \tilde{p}_z(t_0) = p_{z,0} + \epsilon_4(z), \end{aligned} \tag{2.13}$$

where  $p_0 \in P(\Sigma)$ ,  $\tilde{p}_K(t) = 1 - \sum_{z \neq K} \tilde{p}_z(t)$  and  $\tilde{p}_z(t) \in [0, 1]$  for  $z < K$ . Then there exists uniform constant  $B$  and  $N_0$ , such that when  $N > N_0$  and

$$\sum_{i=1}^2 |\epsilon_i(t, z)| + \sum_{i=3}^4 |\epsilon_i(z)| + (\tilde{p}_K(t))^- \leq \frac{1}{N}, \quad \forall (t, z) \in [t_0, T] \times \Sigma,$$

we have for all  $t \in [t_0, T]$  and  $z \in \Sigma$ ,

$$|\theta_z(t) - \tilde{\theta}_z(t)| + |p_z(t) - \tilde{p}_z(t)| \leq \frac{B}{N}.$$

**Proof** See Sect. 4.

Note that the constant  $B$  depends on the FBODE system and the bound of its true solution, but is independent of the DNN structure, which implies that Theorem 2.6 is an algorithm independent result.

Combining Theorems 2.4 and 2.6, we immediately have the following result.

**Theorem 2.7** *The sequence  $(\tilde{\theta}^{(N)}, \tilde{p}^{(N)})$  defined in (2.10) converges uniformly to the true solution  $(\theta, p)$  of the FBODE (2.2), that is, for  $t \in [0, T]$ ,*

$$\lim_{N \rightarrow +\infty} \tilde{\theta}^{(N)}(t) = \theta(t), \quad \lim_{N \rightarrow +\infty} \tilde{p}^{(N)}(t) = p(t).$$

**Remark 2.8** Although we only prove the convergence and error estimation results for a two-layer neural network structure characterized by  $\Theta^n(v_1, \nu)$  and  $\mathbf{P}^n(v_2, \nu)$ , the idea and the proof can be easily adapted to more sophisticated neural network models (more layers, LSTM, etc.) as they share similar structures.

### 3 Application: Optimal Market Making with Rank-Based Reward

The model setting is similar to Guéant [15], except that exchange provides incentive reward for market making. The terminal payoffs of market makers depend on their trading volumes and rankings in the market. The optimization problems of different market makers are coupled. It is in general difficult to solve a finite players game due to high dimensionality, but MFG provides a good approximation, see [8], [9] and [12]. We therefore use MFG as a proxy to solve the optimal market marking problem.

Consider a continuum family of market makers  $\Omega_{mm}$  who keep quoting bid/ask limit orders. Select one of them as a reference market maker. Assume asset price  $S_t$  follows a Brownian motion with initial value  $S$ ,

$$dS_t = \sigma dW_t,$$

where  $W_t$  is a standard Brownian adapted to the natural filtration  $\{\mathcal{F}_t^W\}_{t \in \mathbb{R}_+}$  and  $\sigma$  the volatility of the stock. Assume  $\delta_t^a$  and  $\delta_t^b$  are ask/bid spreads, which are controls determined by the reference market maker. Denote by  $N_t^a$  and  $N_t^b$  the jump processes for buy/sell market order arrivals to the reference market maker, with intensities  $\Lambda(\delta_t^a)$  and  $\Lambda(\delta_t^b)$ , respectively. Assume  $\Lambda : \mathbb{R} \rightarrow \mathbb{R}$  has continuous inverse function, and is decreasing, continuously differentiable satisfying:

$$\frac{\partial^2 \Lambda}{\partial \delta^2}(\delta) \Lambda(\delta) < 2 \left( \frac{\partial \Lambda}{\partial \delta}(\delta) \right)^2. \tag{3.1}$$



The reference market maker has state variables  $(X_t, q_t, v_t)$  in which  $X_t$  is cash account,  $q_t$  inventory,  $v_t$  accumulated trading volume, with initial value  $(x, q, v)$ . We assume  $q_t$  can only take values in a finite set  $\mathbf{Q} = \{-Q, \dots, Q\}$  and  $v_t$  can only take values in a finite set  $\mathbf{V} := \{0, \dots, v_{\max}\}$  and any trading volume above  $v_{\max}$  is not counted in the reward calculation. Denote by  $I^b := \mathbb{1}_{q+1 \in \mathbf{Q}}$  and  $I^a := \mathbb{1}_{q-1 \in \mathbf{Q}}$  the indicators of market maker’s buying and selling capabilities.

The dynamic of  $X_t$  is given by

$$dX_t = (S_t + \delta_t^a)I^a(q_t)dN_t^a - (S_t - \delta_t^b)I^b(q_t)dN_t^b,$$

that of  $q_t$  by

$$dq_t = I^b(q_t)dN_t^b - I^a(q_t)dN_t^a,$$

and that of  $v_t$  by

$$dv_t = (I^b(q_t)dN_t^b + I^a(q_t)dN_t^a)\mathbb{1}_{\{v_t < v_{\max}\}}.$$

Denote by  $p(t, q_t, v_t)$  the probability measure on the mean field of discrete states  $(q_t, v_t)$  for all market makers. The reference market maker wants to solve the following optimization problem:

$$\sup_{\delta^a, \delta^b} \mathbb{E}_0 \left[ X_T + q_T S_T - l(|q_T|) + R(v_T) - \frac{1}{2} \gamma \sigma^2 \int_0^T q_t^2 dt \right], \tag{3.2}$$

where  $X_T + q_T S_T$  is the cash value at time  $T$ ,  $l(|q_T|)$  the terminal inventory holding penalty with  $l$  an increasing convex function on  $R_+$  with  $l(0) = 0$ ,  $\gamma \sigma^2 \int_0^T q_t^2 dt$  the accumulated running inventory holding penalty with  $\gamma$  a positive constant representing the risk adverse level, and  $R(v_T)$  the cash reward given by the exchange as incentive for market makers to increase trading volume  $v_T$ . We consider the rank-based trading volume reward, defined by

$$R(v_T) := R \left( 1 - \sum_{j=v_T}^{v_{\max}} \sum_{i=-Q}^Q p(T, i, j) \right), \tag{3.3}$$

where  $1 - \sum_{j=v_T}^{v_{\max}} \sum_{i=-Q}^Q p(T, i, j)$  is the percentage of market makers that the reference market maker exceeds in trading volumes and  $R$  a positive constant representing the maximum reward set by the exchange.

Using the martingale property, (3.2) can be reduced to  $x + qS + \theta(0, q, v)$ , where  $\theta$  is the value function defined by

$$\begin{aligned} \theta(t, q, v) := & \sup_{\delta^a, \delta^b} \mathbb{E}_t \left[ \int_t^T \left[ \delta_s^a \Lambda(\delta_s^a) + \delta_s^b \Lambda(\delta_s^b) - \frac{1}{2} \gamma \sigma^2 q_s^2 \right] ds \right. \\ & \left. - l(|q_T|) + R \left( 1 - \sum_{j=v_T}^{v_{\max}} \sum_{i=-Q}^Q p(T, i, j) \right) \right] \end{aligned}$$

and  $\mathbb{E}_t[\cdot]$  is the conditional expectation given  $q_t = q, v_t = v$ .

We assume market maker takes closed loop feedback control, i.e., when market maker has state  $(q, v)$ ,

$$\delta_t^a = \delta^a(t, q, v), \quad \delta_t^b = \delta^b(t, q, v).$$

Since the only relevant states are  $q_t$  and  $v_t$  that both take values in finite sets, the problem can be reduced to a continuous time finite state MFG discussed in Cecchin and Pelino [9] by reformulating some notations as follows. Define  $K := (2Q + 1)(v_{\max} + 1)$  and  $\Sigma := \{1, \dots, K\}$ . There is a one-to-one mapping  $Z : \mathbf{Q} \times \mathbf{V} \rightarrow \Sigma$ : For every  $(q, v) \in \mathbf{Q} \times \mathbf{V}$ , there exists  $z \in \Sigma$  such that

$$z = Z(q, v)$$

and for every  $z \in \Sigma$ , there exists  $(q, v) \in \mathbf{Q} \times \mathbf{V}$  such that

$$(q, v) = Z^{-1}(z) := (Z_1^{-1}(z), Z_2^{-1}(z)).$$

The state  $(q, v)$  is then reformulated by state  $z$ . The value function  $\theta$  and probability measure on mean field of state  $p$  are reformulated as  $\theta, p : [0, T] \rightarrow \mathbb{R}^K$ , where

$$\begin{aligned} \theta(t) &:= (\theta_1(t), \dots, \theta_K(t)), & \theta_z(t) &= \theta(t, Z_1^{-1}(z), Z_2^{-1}(z)) \\ p(t) &:= (p_1(t), \dots, p_K(t)), & p_z(t) &= p(t, Z_1^{-1}(z), Z_2^{-1}(z)). \end{aligned}$$

Define  $\lambda$  as

$$\lambda(t, z) := (\lambda_1(t, z), \dots, \lambda_K(t, z)),$$

where  $\lambda$  satisfy

$$\begin{aligned} \lambda_{\beta^a(z)}(t, z) &:= \lambda(\delta_t^a) > 0; & \lambda_{\beta^b(z)}(t, z) &:= \lambda(\delta_t^b) > 0; \\ \lambda_z(t, z) &:= -\sum_{y \neq z} \lambda_y(t, z); & \lambda_y(t, z) &:= 0 \quad y \neq \beta^a(z), \beta^b(z), z. \end{aligned} \tag{3.4}$$

$\beta^a(z)$  and  $\beta^b(z)$  are defined as the two accessible states from state  $z$ ,

$$\begin{aligned} \beta^a(z) &= \begin{cases} Z(Z_1^{-1}(z) - 1, Z_2^{-1}(z) + 1) & Z_1^{-1}(z) > -Q, Z_2^{-1}(z) < v_{\max} \\ Z(Z_1^{-1}(z) - 1, v_{\max}) & Z_1^{-1}(z) > -Q, Z_2^{-1}(z) = v_{\max} \\ z & Z_1^{-1}(z) = -Q \end{cases} \\ \beta^b(z) &= \begin{cases} Z(Z_1^{-1}(z) + 1, Z_2^{-1}(z) + 1) & Z_1^{-1}(z) < Q, Z_2^{-1}(z) < v_{\max} \\ Z(Z_1^{-1}(z) + 1, v_{\max}) & Z_1^{-1}(z) < Q, Z_2^{-1}(z) = v_{\max} \\ z & Z_1^{-1}(z) = Q \end{cases} \end{aligned}$$

Define  $F$  and  $G$  as

$$\begin{aligned} F(t, z, \lambda(t, z)) &:= \Lambda^{-1}(\lambda_{\beta^a(z)}(t, z))\lambda_{\beta^a(z)}(t, z) \\ &\quad + \Lambda^{-1}(\lambda_{\beta^b(z)}(t, z))\lambda_{\beta^b(z)}(t, z) - \frac{1}{2}\gamma\sigma^2 Z_1^{-1}(z)^2 \\ G(z, p) &:= -I(|Z_1^{-1}(z)|) + R \left( 1 - \sum_{j=v}^{v_{\max}} \sum_{i=-Q}^Q pZ(i, j) \right). \end{aligned} \tag{3.5}$$

The optimal market making problem is now reduced to a continuous time finite state MFG discussed in Sect. 2 of this paper. Denote the game as  $G_R$ . We have the following result.

**Proposition 3.1**  $G_R$  satisfies Assumption 2.1.

**Proof** See Sect. 4.

According to Cecchin and Pelino [9], the Nash equilibrium of MFG  $G_R$  and that of finite players game exist and are unique, and the game with  $N$  players converges to the limiting MFG case in the order  $O(\frac{1}{N})$ . We can numerically solve the FBODE system corresponding to the finite state MFG in (3.5) and find the equilibrium value function and probability of mean field by the DNN technique.

We next do some numerical tests. We use a LSTM (long short-term memory) neural network to approximate the solution  $(\theta, p)$ . Denote the function constructed by LSTM neural network as  $(\tilde{\theta}(t, \beta), \tilde{p}(t, \beta))$ , where  $\beta$  is the parameters set for neural network, designed by the following steps: Layer 0 is the input  $t \in [0, T]$  and layer  $k$  with output  $h_k$  is designed as follows:

$$\begin{aligned} f_k &= \sigma_g(W_f t + U_f h_{k-1} + b_f) \\ i_k &= \sigma_g(W_i t + U_i h_{k-1} + b_i) \\ o_k &= \sigma_g(W_o t + U_o h_{k-1} + b_o) \\ \tilde{c}_t &= \sigma_c(W_c t + U_c h_{k-1} + b_c) \\ c_k &= f_k \circ c_{k-1} + i_k \circ \tilde{c}_k \\ h_k &= o_k \circ \sigma_h(c_k) \end{aligned}$$

with the initial values  $c_0 = h_0 = 0$ , where the operator  $\circ$  is the element-wise product, functions  $\sigma$  some scaled tanh activation functions (hence satisfying all assumptions in our main results),  $t \in [0, T]$  input to the LSTM network,  $f_k \in \mathbb{R}^h$  forget gate’s activation vector,  $i_k \in \mathbb{R}^h$  input/update gate’s activation vector,  $o_k \in \mathbb{R}^h$  output gate’s activation vector,  $h_k \in \mathbb{R}^h$  hidden state vector,  $\tilde{c}_k \in \mathbb{R}^h$  cell input activation vector,  $c_k \in \mathbb{R}^h$  cell state vector,  $W \in \mathbb{R}^{h \times 1}$ ,  $U \in \mathbb{R}^{h \times h}$ ,  $b \in \mathbb{R}^h$  weight matrices and bias vector parameters which need to be learned during training, and  $h$  the number of hidden units. LSTM network is an advanced version of a traditional neural network and provides more accurate approximation for complicated functions. For our model, this specific structure performs better than that of traditional neural network.

We use a LSTM network with 3 layers and 32 nodes per layer. The network is trained by stochastic gradient approach with mesh-free randomly sampling points in  $[0, T]$ . This randomness adds to the robustness of the network. The detailed training procedure is similar to that in Sirignano and Spiliopoulos [30].

The market order arrival intensity function is given by  $\Lambda(\delta) := Ae^{-k\delta}$  and the liquidity penalty function  $l(q) := aq^2$ . We assume all market makers start at 0 inventory and 0 trading volume. The benchmark data used are  $S = 20$ , hourly volatility  $\sigma = 0.01$ ,  $\gamma = 1$ ,  $T = 10$  hours, capacity  $Q = 1$ ,  $v_{\max} = 10$ ,  $k = 2$ ,  $a = 2$ ,  $A = 0.5$ , and  $R = 2$ .

There are two typical schemes of trading volume rewards in most of exchanges’ incentive programs. One is the rank-based trading volume reward as in (3.3), and the other is the linear trading volume reward, defined by

$$R_L(v_T) := R \frac{v_T}{v_{\max}}. \tag{3.6}$$

Since  $R_L(v_T)$  is independent of the mean field of state, the FBODE system for the value function and the probability of mean field of state is decoupled and can be solved numerically with a standard Euler scheme. We next do numerical tests and compare the value functions, optimal bid/ask spreads and probability distributions of trading volumes under three different trading volume reward schemes: 1. no trading volume reward ( $R = 0$ , benchmark case), 2. linear trading volume reward ( $R$  in (3.6)), and 3. rank-based trading volume reward ( $R$  in (3.3)). The rank-based reward introduces competition between market makers, whereas the

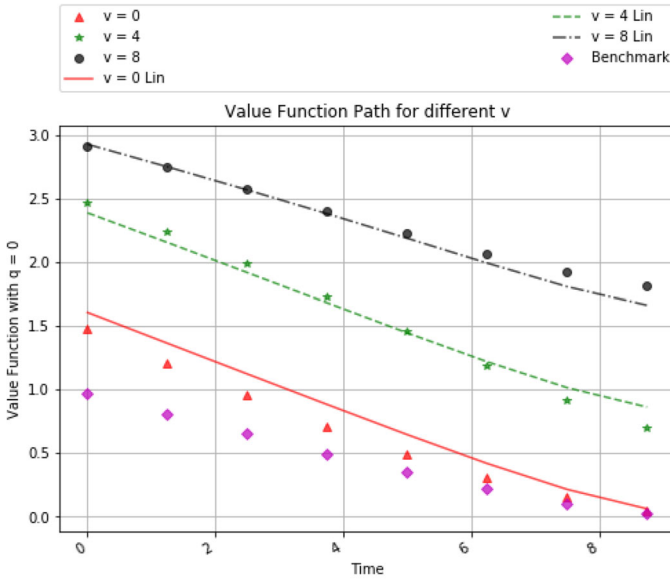


Fig. 1  $\theta(t, 0, v)$  Path for two schemes

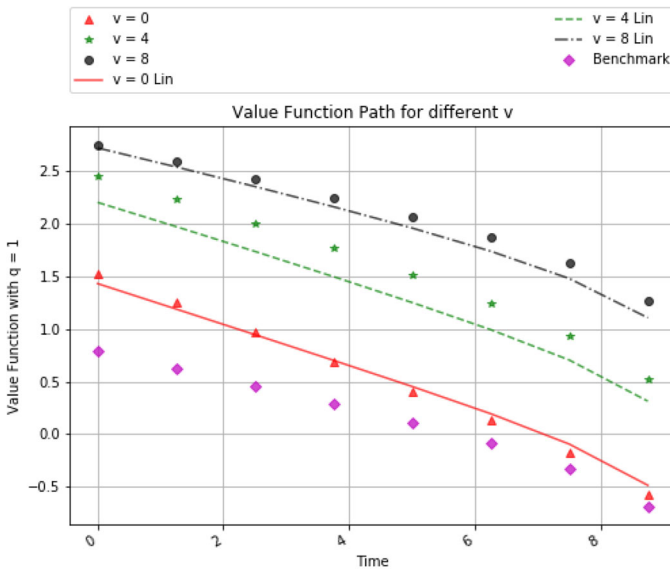


Fig. 2  $\theta(t, 1, v)$  Path for two schemes

linear reward does not. The training result is satisfactory and the average loss is less than 0.003.

Figures 1 and 2 show the value functions  $\theta(t, q, v)$  for fixed  $q = 0, 1$  and varied  $v$  with the ‘Benchmark’ representing no trading volume reward, ‘ $v = 0$  lin’ path the linear reward and initial trading volume  $v = 0$ , and ‘ $v = 0$ ’ the rank-based reward. It is clear that the

introduction of market incentive  $R$  increases the value functions for market makers, and the higher the initial trading volume  $v$ , the higher the value function  $\theta$ . Even for market makers with initial trading volume  $v = 0$ , the value functions are still higher than the benchmark one as they benefit from their potential market incentive gains, which explains the convergence of the curves for  $v = 0$  to the benchmark one as  $t$  tends to  $T$ . The value functions for linear and rank awards are largely the same.

## 4 Proofs

### 4.1 Proof of Theorem 2.4

**Proof** According to Theorem 2.3, there exists unique solution  $(\theta, p)$  to ODE system (2.2), which is also the unique minimal point for  $\Psi$  such that

$$\Psi(\theta, p) = 0.$$

We use  $(v_i^{-1})'$  to denote the first-order derivative of  $v_i^{-1}$  for  $i = 1, 2$ . From (2.7) we know  $\theta$  is bounded by  $C_{GH}$ . Hence,  $\frac{d}{dt}\theta_z(t)$  is also bounded uniformly for  $t$  and  $z$ . Moreover,  $p(t) \in P(\Sigma)$  for any  $t \in [0, T]$ , and hence, is also bounded. From the assumption on  $v_1, v_2$ , we know

$$\begin{aligned} \theta_z(t) &< \sup \|v_1\|, \\ \inf v_2 &< p_z(t) < \sup v_2. \end{aligned}$$

It means  $\theta_z$ 's image is bounded and a strict subset of  $v_1^{-1}$ 's domain. It is similar for  $p_z$  and  $v_2^{-1}$ . Combining with the continuously differentiability of  $v_1^{-1}$  and  $v_2^{-1}$ , we know  $v_1^{-1}(\theta_z(t))$ ,  $(v_1^{-1})'(\theta_z(t))$ ,  $v_2^{-1}(p_z(t))$  and  $(v_2^{-1})'(p_z(t))$  are bounded by some constant  $C$  uniformly for  $t$  and  $z$ .  $(v_i)'$  and  $(v_i)''$  are Lipschitz continuous on  $[-2C, 2C]$  with coefficient  $L$  for  $i = 1, 2$ . Define  $C^N(v) := \{\zeta : [0, T] \rightarrow \mathbb{R}; \zeta(t) = \sum_{i=1}^N \beta_i v(\alpha_i t + c_i)\}$ . According to the proof of Theorem 7.1 in Sirignano and Spiliopoulos [30], for any  $0 < \varepsilon < C$ , there exists  $N > 0$  and  $y_z \in C^N(v)$  such that

$$\begin{aligned} &\|y_z(t) - v_1^{-1}(\theta_z(t))\| + \left\| \frac{d}{dt}y_z(t) - \frac{d}{dt}v_1^{-1}(\theta_z(t)) \right\| \\ &+ \left\| \frac{d^2}{dt^2}y_z(t) - \frac{d^2}{dt^2}v_1^{-1}(\theta_z(t)) \right\| \leq \varepsilon. \end{aligned} \tag{4.1}$$

Hence, we have

$$\|v_1(y_z(t)) - \theta_z(t)\| \leq C \|y_z(t) - v_1^{-1}(\theta_z(t))\| \leq C\varepsilon.$$

On the other hand,

$$\begin{aligned} \frac{d}{dt}v_1(y_z(t)) - \frac{d}{dt}\theta_z(t) &= \frac{d}{dt}v_1(y_z(t)) - \frac{d}{dt}v_1(v_1^{-1}(\theta_z(t))) \\ &= (v_1)'(y_z(t))\frac{d}{dt}y_z(t) - (v_1)'(v_1^{-1}(\theta_z(t)))\frac{d}{dt}v_1^{-1}(\theta_z(t)) \\ &= (v_1)'(y_z(t)) \left[ \frac{d}{dt}y_z(t) - \frac{d}{dt}v_1^{-1}(\theta_z(t)) \right] \\ &+ \frac{d}{dt}v_1^{-1}(\theta_z(t))[(v_1)'(y_z(t)) - (v_1)'(v_1^{-1}(\theta_z(t)))]. \end{aligned}$$

As  $y_z(t) \in [-2C, 2C]$ , there exists constant  $C_1$  such that  $(v_1)'(y_z(t))$  is bounded by it uniformly. Moreover, we have

$$\left\| \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right\| \leq \|(v_1^{-1})'(\theta_z(t))\| \left\| \frac{d}{dt} \theta_z(t) \right\| \leq C^2.$$

Hence, we have

$$\begin{aligned} \left\| \frac{d}{dt} v_1(y_z(t)) - \frac{d}{dt} \theta_z(t) \right\| &\leq \left\| \frac{d}{dt} y_z(t) - \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right\| \|v_1'(y_z(t))\| \\ &\quad + \left\| v_1'(y_z(t)) - v_1'(v_1^{-1}(\theta_z(t))) \right\| \|(v_1^{-1})'(\theta_z(t))\| \left\| \frac{d}{dt} \theta_z(t) \right\| \\ &\leq C_1 \left\| \frac{d}{dt} y_z(t) - \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right\| \\ &\quad + C^2 L \|y_z(t) - v_1^{-1}(\theta_z(t))\| \leq (C_1 + C^2 L)\varepsilon. \end{aligned}$$

The first inequality above comes from the boundedness and Lipschitz continuity of  $v_1'$ , as well as the boundedness of  $(v_1^{-1})'(\theta_z(t))$ . Moreover, for second-order derivatives, we have

$$\left\| \frac{d^2}{dt^2} v_1^{-1}(\theta_z(t)) \right\| = \left\| (v_1^{-1})''(\theta_z(t)) \frac{d}{dt} \theta_z(t) + (v_1^{-1})'(\theta_z(t)) \left( \frac{d}{dt} \theta_z(t) \right)^2 \right\|.$$

As  $\theta_z(t)$  and  $\frac{d}{dt} \theta_z(t)$  are bounded and  $(v_1^{-1})$  is twice continuously differentiable,  $\frac{d^2}{dt^2} v_1^{-1}(\theta_z(t))$  is bounded. To estimate the difference of second-order derivatives between approximation function and true function, we have

$$\begin{aligned} \frac{d^2}{dt^2} v_1(y_z(t)) - \frac{d^2}{dt^2} \theta_z(t) &= \left( \frac{d}{dt} y_z(t) \right)^2 v_1''(y_z(t)) + v_1'(y_z(t)) \frac{d^2}{dt^2} y_z(t) \\ &\quad - \left( \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right)^2 v_1''(v_1^{-1}(\theta_z(t))) - v_1'(v_1^{-1}(\theta_z(t))) \frac{d^2}{dt^2} v_1^{-1}(\theta_z(t)). \end{aligned}$$

Define

$$\begin{aligned} a &:= \left\| \left( \frac{d}{dt} y_z(t) \right)^2 - \left( \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right)^2 \right\| \|v_1''(y_z(t))\| \\ b &:= \left\| v_1''(y_z(t)) - v_1''(v_1^{-1}(\theta_z(t))) \right\| \left\| \left( \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right)^2 \right\| \\ c &:= \left\| \frac{d^2}{dt^2} y_z(t) - \frac{d^2}{dt^2} v_1^{-1}(\theta_z(t)) \right\| \|v_1'(y_z(t))\| \\ d &:= \left\| v_1'(y_z(t)) - v_1'(v_1^{-1}(\theta_z(t))) \right\| \left\| \frac{d^2}{dt^2} v_1^{-1}(\theta_z(t)) \right\|, \end{aligned}$$

and we have

$$\left\| \frac{d^2}{dt^2} v_1(y_z(t)) - \frac{d^2}{dt^2} \theta_z(t) \right\| \leq a + b + c + d.$$

As  $y_z(t)$  and  $\frac{d}{dt} v_1^{-1}(\theta_z(t))$  are bounded from previous proof, and  $v_1$  is triple continuously differentiable function by definition,  $v_1''(y_z(t))$ ,  $v_1'(y_z(t))$ ,  $(\frac{d}{dt} v_1^{-1}(\theta_z(t)))^2$  and  $\frac{d^2}{dt^2} v_1^{-1}(\theta_z(t))$

are also bounded. Moreover,  $\left\| \frac{d}{dt} y_z(t) \right\| \leq \left\| \frac{d}{dt} v_1^{-1}(\theta_z(t)) \right\| + \varepsilon$ , and hence bounded. According to the Lipschitz continuity of  $v_1'$  and  $v_1''$ , as well as (4.1), we know there exists constants  $C_2$ , such that

$$\left\| \frac{d^2}{dt^2} v_1(y_z(t)) - \frac{d^2}{dt^2} \theta_z(t) \right\| \leq a + b + c + d \leq C_2 \varepsilon.$$

By making transformation on  $\varepsilon$  in above proof, we know for any  $0 < \varepsilon < C$ , there exists  $N > 0$  and  $y_z \in C^N(\nu)$  such that

$$\|v_1(y_z(t)) - \theta_z(t)\| + \left\| \frac{d}{dt} v_1(y_z(t)) - \frac{d}{dt} \theta_z(t) \right\| + \left\| \frac{d^2}{dt^2} v_1(y_z(t)) - \frac{d^2}{dt^2} \theta_z(t) \right\| \leq \varepsilon.$$

Hence, we know there exists  $N > 0$  and  $y_z \in C^N(\nu)$  such that

$$\left\| \frac{d^2}{dt^2} v_1(y_z(t)) \right\| \leq \varepsilon + \left\| \frac{d^2}{dt^2} \theta_z(t) \right\| \leq \varepsilon + C_{\theta GH} < B_{\theta}.$$

Similarly, we know  $\left\| \frac{d}{dt} v_2(y_z(t)) \right\| \leq C_{pGH} < B_p$ . Then we get

$$\left( B_{\theta} - \max_{t \in [0, T]} \left| \frac{d\tilde{\theta}_z(t)}{dt} \right| \right)^{-} = \left( B_p - \max_{t \in [0, T]} \left| \frac{d\tilde{p}_z(t)}{dt} \right| \right)^{-} = 0$$

If we define

$$\begin{aligned} \hat{\Theta}^N(v_1, \nu) &:= \left\{ \zeta : [0, T] \rightarrow \mathbb{R}^K; \quad \zeta(t) \right. \\ &= \left( v_1 \left( \sum_{i=1}^N \beta_{1,i} \nu(\alpha_{1,i} t + c_{1,i}) \right) \right), \dots, v_1 \left( \sum_{i=1}^n \beta_{K,i} \nu(\alpha_{K,i} t + c_{K,i}) \right) \left. \right\} \end{aligned}$$

Then from proof above we know for any  $0 < \varepsilon < C$ , there exists  $N > 0$  and  $\tilde{\theta}^{(N)} \in \hat{\Theta}^N(v_1, \nu)$  such that

$$\left\| \tilde{\theta}_z^{(N)}(t) - \theta_z(t) \right\| + \left\| \frac{d}{dt} \tilde{\theta}_z^{(N)}(t) - \frac{d}{dt} \theta_z(t) \right\| \leq \varepsilon.$$

On the other hand, notice that any function  $f_N \in \hat{\Theta}^N(v_1, \nu)$ , there should exists  $f_{KN} \in \Theta^{KN}(v_1, \nu)$  such that  $f_{KN} = f_N$ , by letting some  $\beta_{j,i} = 0$ . It means  $\hat{\Theta}^N(v_1, \nu) \subset \Theta^{KN}(v_1, \nu)$ , and  $\tilde{\theta}^{(N)} \in \Theta^{KN}(v_1, \nu)$ . For  $p$  and  $\mathbf{P}^n(v_2, \nu)$ , we can have similar argument. Hence, we conclude that for any  $0 < \varepsilon < C$ , there exists  $N > 0$  and  $\tilde{\theta}^{(N)} \in \Theta^N(v_1, \nu)$ ,  $\tilde{p}^{(N)} \in \mathbf{P}^N(v_2, \nu)$  such that

$$\begin{aligned} \left\| \tilde{\theta}_z^{(N)}(t) - \theta_z(t) \right\| + \left\| \frac{d}{dt} \tilde{\theta}_z^{(N)}(t) - \frac{d}{dt} \theta_z(t) \right\| &\leq \varepsilon \\ \left\| \tilde{p}_z^{(N)}(t) - p_z(t) \right\| + \left\| \frac{d}{dt} \tilde{p}_z^{(N)}(t) - \frac{d}{dt} p_z(t) \right\| &\leq \varepsilon. \end{aligned}$$

Then similar to the proof for Theorem 7.1 in Sirignano and Spiliopoulos [30], we know there exists a uniform constant  $M$  which only depends on boundedness of  $\theta$ ,  $\lambda^*$  and Lipschitz coefficient of  $\lambda^*$  and  $H$ , such that

$$\Psi(\tilde{\theta}^{(N)}, \tilde{p}^{(N)}) \leq M\varepsilon.$$

It concludes the proof. □

### 4.2 Proof of Theorem 2.5

**Proof** Since for different  $t_0$ , the following proof is the same, we assume  $t_0 = 0$  for the ease of notation. We first focus on proving the first inequality in (2.12). Define

$$e(t, z) := \frac{d\tilde{\theta}_z(t)}{dt} + H(z, \Delta^z \tilde{\theta}(t)).$$

As  $\Psi < \varepsilon$ , from the definition of  $\Psi$  we know  $\frac{d^2\theta_z(t)}{dt^2}$  is uniform bounded on  $[0, T]$ . Furthermore,  $H$  is Lipschitz continuous, and  $\theta$  has bounded first-order derivative to  $t$ . Hence,  $e(t, z)$  is a Lipschitz continuous function on  $[0, T]$ . Denote its Lipschitz coefficient as  $L$ . There exists  $\hat{t} \in [0, T]$  such that

$$|e(\hat{t}, z)| = \sup_{t \in [0, T]} |e(t, z)|.$$

For any  $\Delta t$  such that

$$\frac{|e(\hat{t}, z)|}{16L} < \Delta t < \frac{|e(\hat{t}, z)|}{2L}. \tag{4.2}$$

From Lipschitz continuity of  $e$ , we have

$$||e(\hat{t} \pm \Delta t, z)| - |e(\hat{t}, z)|| \leq |e(\hat{t} \pm \Delta t, z) - e(\hat{t}, z)| \leq L\Delta t.$$

It means

$$|e(\hat{t}, z)| - L\Delta t \leq |e(\hat{t} \pm \Delta t, z)| \leq |e(\hat{t}, z)| + L\Delta t.$$

From (4.2), we know  $L\Delta t \leq \frac{|e(\hat{t}, z)|}{2}$ ; therefore,

$$|e(\hat{t}, z)| - L\Delta t \geq \frac{1}{2}|e(\hat{t}, z)| \geq 0.$$

Combining the above two inequalities, we know for any  $t \in [\hat{t} - \Delta t, \hat{t} + \Delta t]$ ,

$$e^2(t, z) \geq (|e(\hat{t}, z)| - L\Delta t)^2 > \frac{1}{4}|e(\hat{t}, z)|^2.$$

We can have the following estimation. Without loss of generality, we can assume  $T - \hat{t} \geq \frac{T}{2}$  (otherwise we can use the other side  $[\hat{t} - \Delta t, \hat{t}]$  as the limit of integration on the second inequality in the following),

$$\begin{aligned} \Psi &\geq \int_{\hat{t}}^{\hat{t}+\Delta t} e^2(t, z) dt \geq \min(\Delta t(|e(\hat{t}, z)| - L\Delta t)^2, (T - \hat{t})(|e(\hat{t}, z)| - L\Delta t)^2) \\ &> \min\left(\frac{|e(\hat{t}, z)|^3}{64L}, \frac{|e(\hat{t}, z)|^2}{8} T\right), \end{aligned}$$

which implies that for any  $t \in [0, T]$ ,

$$|e(t, z)| < \max\left(4L^{\frac{1}{3}}\Psi^{\frac{1}{3}}, \frac{2\sqrt{2T}}{T}\Psi^{\frac{1}{2}}\right).$$

There exists  $\Psi_0$  such that for all  $\Psi < \Psi_0$ ,  $4L^{\frac{1}{3}}\Psi^{\frac{1}{3}} \geq \frac{2\sqrt{2T}}{T}\Psi^{\frac{1}{2}}$ . Hence, there exists constant  $C$  such that

$$|e(t, z)| < 4L^{\frac{1}{3}}\Psi^{\frac{1}{3}} < C\varepsilon^{\frac{1}{3}}.$$



Using the similar arguments, the third inequality in (2.12) can be proved. The second and fourth inequalities are trivial.  $\square$

### 4.3 Proof of Theorem 2.6

The general structure of the proof is similar to those in Cecchin and Pelino [9] with one key difference: In Cecchin and Pelino [9],  $p$  satisfies a non-perturbed Kolmogorov forward equation and has initial value in  $P(\Sigma)$  with nonnegative components, whereas  $\tilde{p}$  satisfies a perturbed Kolmogorov forward equation (2.13) and its initial value is not necessarily in  $P(\Sigma)$  and may be negative, which makes some prior estimations in Cecchin and Pelino [9] not applicable for our case. We need to provide extra modifications by adding and subtracting an extra term  $M_1$  such that  $\tilde{p}(t) + M_1$  is nonnegative, and need to modify every step in the proof to estimate the extra terms introduced by  $M_1$ . For the completeness, we give the whole proof.

Note that the solution pair  $(\tilde{\theta}, \tilde{p})$  to (2.13) is determined only by initial time  $t_0$  and initial value  $\tilde{p}(t_0)$ . We first prove that  $\tilde{\theta}$  is well defined (Proposition 4.5), continuous at  $\tilde{p}(t_0)$  (Proposition 4.2), and continuously differentiable at  $\tilde{p}(t_0)$  (Theorem 4.7) by discussing the linearized system (4.11), we then prove that  $\tilde{\theta}$  satisfies a PDE similar to the master equation in Cecchin and Pelino [9] (Theorem 4.8) and that the master equation on some discrete grids of  $P(\Sigma)$  can be approximated by a backward ODE with extra error terms (Proposition 4.9), and we finally estimate the difference between  $\theta$  and  $\tilde{\theta}$  by comparing the two backward ODE systems and conclude the proof.

Denote by  $\|x\| := \max_{1 \leq z \leq K} |x_z|$ , the norm of  $x$  in  $\mathbb{R}^K$  and  $\|f\| := \max_{t \in [0, T]} \max_{1 \leq z \leq K} |f_z(t)|$ , the norm of  $f$  in  $\mathcal{C}^0([0, T]; \mathbb{R}^K)$ . Due to the introduction of perturbation terms in ODE system (2.13), the existence and uniqueness of its solution can no longer be guaranteed for every initial value  $\tilde{p}(t_0)$ . However, under certain conditions on (2.13), we can have the existence and prior bound estimation of solution to (2.13).

**Proposition 4.1** *Given constant  $M > 0$ , define  $I_{p, M} := [-M, 1 + M]^K$  and*

$$C_G(M) := \max_{z \in \Sigma, p \in I_{p, M}} |G(z, p)| + \|\epsilon_3\| + 2\|\epsilon_1\|T + 2 \max_{z \in \Sigma} H(z, 0)T,$$

$$A_G(M) := [-2C_G(M), 2C_G(M)]^K, \quad \Lambda(M) := \max_{y, z \in \Sigma, \mu \in A_G(M)} |\lambda_y^*(z, \mu)|.$$

If functions  $\epsilon_i, i = 2, 4$  satisfy

$$\|\epsilon_2\| + \|\epsilon_4\| < \frac{1}{N_0}, \tag{4.3}$$

where  $\frac{1}{N_0} := \frac{1}{3}Me^{-\Lambda(M)T}$ . Then for any initial time  $t_0 \in [0, T]$  and  $\tilde{p}(t_0) \in \bar{B}(P(\Sigma), \frac{1}{N_0})$ , ODE system (2.13) has solution  $(\tilde{\theta}, \tilde{p})$ , where

$$\bar{B}\left(P(\Sigma), \frac{1}{N_0}\right) = \left\{ \tilde{p} \in \mathbb{R}^K, \text{ s.t. } \min_{p \in P(\Sigma)} \|\tilde{p} - p\| \leq \frac{1}{N_0} \right\}.$$

Moreover,  $(\tilde{\theta}, \tilde{p})$  satisfy the following on  $[t_0, T]$  uniformly for any initial time  $t_0 \in [0, T]$  and initial value  $\tilde{p}(t_0)$ .

$$\tilde{\theta}_z(t) \in [-C_G(M), C_G(M)], \quad \tilde{p}_z(t) \in [-M, 1 + M].$$

**Proof** Given a prior  $\bar{p}$  such that  $\bar{p}(t) \in [-M, 1+M]^K$  for all  $t \in [t_0, T]$ , Lipschitz continuous with Lipschitz coefficient bounded by  $L(M)$ , where

$$\left\| \frac{d\bar{p}}{dt} \right\| \leq L(M) = K(2M + 1)\Lambda(M) + \frac{1}{N_0},$$

and starts with the same  $\bar{p}(t_0) = \tilde{p}(t_0) \in \bar{B}(P(\Sigma), \frac{1}{N_0})$ , with which we solve the backward ODE in (2.13):

$$\frac{d\tilde{\theta}_z(t)}{dt} = -H(z, \Delta^z \tilde{\theta}(t)) + \epsilon_1(t, z), \quad \tilde{\theta}_z(T) = G(z, \bar{p}(T)) + \epsilon_3(z).$$

We know function  $\tilde{\theta}(t)$  is bounded by constant  $C_G(M)$  following a similar proof as [12, Proposition 2]. Note that  $C_G(M)$  is monotonically non-decreasing w.r.t  $M$ , hence  $\Lambda(M)$  is also non-decreasing w.r.t  $M$ . Since  $\bar{p}(t_0) \in \bar{B}(P(\Sigma), \frac{1}{N_0})$ , there exists  $p_0 \in P(\Sigma)$  such that  $\bar{p}(t_0) - p_0 = \epsilon_4$  where  $\|\epsilon_4\| \leq \frac{1}{N_0}$ . Consider two functions  $\tilde{p}$  and  $p$  satisfying

$$\begin{aligned} \frac{d\tilde{p}_z(t)}{dt} &= \sum_y \tilde{p}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) + \epsilon_2(t, z), \quad \tilde{p}_z(t_0) = p_{z,0} + \epsilon_4(z), \\ \frac{dp_z(t)}{dt} &= \sum_y p_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)), \quad p_z(t_0) = p_{z,0} \end{aligned}$$

Integrating both side and subtracting  $\tilde{p}$  and  $p$ , we get

$$\|\tilde{p}(t) - p(t)\| \leq \Lambda(M) \int_{t_0}^t \|\tilde{p}(s) - p(s)\| ds + \|\epsilon_2\| + \|\epsilon_4\|.$$

By Gronwall inequality, we have

$$\|\tilde{p}(t) - p(t)\| \leq (\|\epsilon_2\| + \|\epsilon_4\|) e^{\Lambda(M)T} < M.$$

As  $p$  is the solution to a Kolmogorov equation,  $p(t) \in P(\Sigma)$ . Hence, the solution  $\tilde{p}(t) \in [-M, 1+M]^K$  for all  $t \in [t_0, T]$ , and  $\tilde{p}$  is also Lipschitz continuous with Lipschitz coefficient bounded by  $L(M)$ , as  $\left\| \frac{d\tilde{p}}{dt} \right\| \leq L(M)$ .

Let  $\mathcal{F}([t_0, T])$  be the set of Lipschitz continuous functions defined on  $[t_0, T]$ , with Lipschitz coefficient bounded by  $L(M)$ , taking values in  $[-M, 1+M]^K$  and starting at the same initial value  $\tilde{p}(t_0)$  at  $t_0$ . We can define mapping  $\xi : \mathcal{F}([t_0, T]) \rightarrow \mathcal{F}([t_0, T])$  in the following way: given  $\tilde{p} \in \mathcal{F}([t_0, T])$ , let  $\tilde{\theta}$  be the solution of terminal value problem in (2.13). Then  $\tilde{\theta}(t)$  is bounded by  $C_G(M)$ . Let  $\xi(\tilde{p})$  be the solution to the initial value problem in (2.13).  $\xi(\tilde{p}) \in \mathcal{F}([t_0, T])$  from the above argument. Following the proof of [12, Proposition 4],  $\mathcal{F}([t_0, T])$  is a set of uniformly bounded and equicontinuous functions. Thus, by Arzela–Ascoli theorem, it is a relatively compact set. It is also clear that it is a convex set. Hence, by Brouwer fixed point theorem, we know there exists fixed point for  $\xi$ , which proves the existence of solution to (2.13).  $\square$

We next prove that under certain condition,  $(\tilde{\theta}, \tilde{p})$  is unique and continuous w.r.t initial condition.

**Proposition 4.2** *There exist positive constants  $N_0$  and  $C$ , such that if we have condition (4.3), then for any  $t_0 \in [0, T]$  and initial condition  $\tilde{p}(t_0) \in \bar{B}(P(\Sigma), \frac{1}{N_0})$ , the solution to (2.13) is*

unique. Moreover, let  $(\tilde{\theta}, \tilde{p})$  and  $(\hat{\theta}, \hat{p})$  be two solutions to ODE system (2.13) with different initial conditions  $\tilde{p}(t_0), \hat{p}(t_0) \in \bar{B}(P(\Sigma), \frac{1}{N_0})$ , then

$$\|\tilde{\theta} - \hat{\theta}\| + \|\tilde{p} - \hat{p}\| \leq C \|\tilde{p}(t_0) - \hat{p}(t_0)\| \tag{4.4}$$

**Proof** Start with any  $M$  and the corresponding  $N_0$  defined in Proposition 4.1. Then both  $\tilde{\theta}$  and  $\hat{\theta}$  uniform bounded by  $C_G(M)$ . Let us first assume  $\tilde{p}_z(t), \hat{p}_z(t) \geq -M_1$  uniformly, and we will decide later the value for  $M_1$  and prove the condition for it. Similarly to the proof for [9, Proposition 5], we first try to obtain estimation on LHS of (4.7) given later. Define  $\phi := \tilde{\theta} - \hat{\theta}$  and  $\pi := \tilde{p} - \hat{p}$ . Then the couple  $(\phi, \pi)$  solves

$$\begin{aligned} \frac{d\phi_z(t)}{dt} &= -H(z, \Delta^z \tilde{\theta}(t)) + H(z, \Delta^z \hat{\theta}(t)), \quad \phi_z(T) = G(z, \tilde{p}(T)) - G(z, \hat{p}(T)), \\ \frac{d\pi_z(t)}{dt} &= \sum_y \{ \tilde{p}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) - \hat{p}_y(t) \lambda_z^*(y, \Delta^y \hat{\theta}(t)) \}, \quad \pi_z(t_0) = \tilde{p}_z(t_0) - \hat{p}_z(t_0), \end{aligned} \tag{4.5}$$

Integrating  $\frac{d}{dt} \sum_{z \in \Sigma} \phi_z(t) \pi_z(t)$  over the interval  $[t_0, T]$ , using the product rule and (4.5), also noting that  $\sum_z \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \phi_y(t) = 0$ , after some simple calculus, we have

$$\begin{aligned} \sum_{z \in \Sigma} \phi_z(t_0) \pi_z(t_0) &= \int_{t_0}^T \sum_{z \in \Sigma} [H(z, \Delta^z \tilde{\theta}(t)) - H(z, \Delta^z \hat{\theta}(t)) \\ &\quad - \Delta^z \phi(t) \cdot \lambda^*(z, \Delta^z \tilde{\theta}(t))] (\tilde{p}_z(t) + M_1) dt \\ &\quad + \int_{t_0}^T \sum_{z \in \Sigma} [H(z, \Delta^z \hat{\theta}(t)) - H(z, \Delta^z \tilde{\theta}(t)) \\ &\quad + \Delta^z \phi(t) \cdot \lambda^*(z, \Delta^z \hat{\theta}(t))] (\hat{p}_z(t) + M_1) dt + \sum_{z \in \Sigma} \phi_z(T) \pi_z(T) \\ &\quad + M_1 \int_{t_0}^T \sum_{z \in \Sigma} \Delta^z \phi(t) \cdot [\lambda^*(z, \Delta^z \tilde{\theta}(t)) - \lambda^*(z, \Delta^z \hat{\theta}(t))] dt. \end{aligned} \tag{4.6}$$

As  $\lambda^*(z, \mu) = D_\mu H(z, \mu)$ , by Taylor theorem, there exists point  $a$  on the line between  $\Delta^z \tilde{\theta}(t)$  and  $\Delta^z \hat{\theta}(t)$  such that

$$\begin{aligned} H(z, \Delta^z \tilde{\theta}(t)) - H(z, \Delta^z \hat{\theta}(t)) - \Delta^z \phi(t) \cdot \lambda^*(z, \Delta^z \tilde{\theta}(t)) \\ = -\Delta^z \phi(t) \cdot D_{\mu\mu}^2 H(z, a) \cdot \Delta^z \phi(t). \end{aligned}$$

Then from assumption (2.4), and do above similar on another way round, we have the following estimations:

$$H(z, \Delta^z \tilde{\theta}(t)) - H(z, \Delta^z \hat{\theta}(t)) - \Delta^z \phi(t) \cdot \lambda^*(z, \Delta^z \tilde{\theta}(t)) \leq -C^{-1} \|\Delta^z \phi(t)\|^2.$$

Similarly, we have

$$H(z, \Delta^z \hat{\theta}(t)) - H(z, \Delta^z \tilde{\theta}(t)) + \Delta^z \phi(t) \cdot \lambda^*(z, \Delta^z \hat{\theta}(t)) \leq -C^{-1} \|\Delta^z \phi(t)\|^2.$$

Since  $\sum_{z \in \Sigma} \phi_z(T) \pi_z(T) \leq 0$  by (2.6),  $\tilde{p}_z(t), \hat{p}_z(t) > -M_1$  by the choice of  $M_1$ , using (4.6) and the inequalities above, we have

$$\phi(t_0) \cdot \pi(t_0) + C^{-1} \int_{t_0}^T \sum_{z \in \Sigma} \|\Delta^z \phi(t)\|^2 (\tilde{p}_z(t) + \hat{p}_z(t) + 2M_1) dt$$

$$\leq M_1 \int_{t_0}^T \sum_{z \in \Sigma} \Delta^z \phi(t) \cdot [\lambda^*(z, \Delta^z \tilde{\theta}(t)) - \lambda^*(z, \Delta^z \hat{\theta}(t))] dt.$$

By Lipschitz continuity of  $\lambda^*$ , there exists  $C$  such that

$$|\int_{t_0}^T \sum_{z \in \Sigma} \|\Delta^z \phi(t)\|^2 (\tilde{p}_z(t) + \hat{p}_z(t) + 2M_1) dt| \leq C(\|\pi(t_0)\| \|\phi\| + M_1 \|\phi\|^2). \tag{4.7}$$

Note that unlike the proof for [9, Proposition 5], both  $\tilde{p}_z$  and  $\hat{p}_z$  can be negative in our setting. If they were nonnegative, then we could choose  $M_1 = 0$ , the same technique in [9, Proposition 5] would work. In the case here, we have to introduce  $M_1$  that requires us to do many more prior estimations.

We next derive the bound for  $\pi$ . Integrating the second equation in (4.5) over  $[t_0, t]$ , we have

$$\pi_z(t) = \pi_z(t_0) + \int_{t_0}^t \sum_y \{ \tilde{p}_y(s) \lambda_z^*(y, \Delta^y \tilde{\theta}(s)) - \hat{p}_y(s) \lambda_z^*(y, \Delta^y \hat{\theta}(s)) \} ds.$$

As  $\lambda^*$  is both bounded and Lipschitz continuous, there exists  $C$  such that

$$\begin{aligned} \max_{z \in \Sigma} |\pi_z(t)| &\leq \max_{z \in \Sigma} |\pi_z(t_0)| + C \int_{t_0}^t \max_{z \in \Sigma} |\pi_z(s)| ds + C \int_{t_0}^t \sum_{z \in \Sigma} \|\Delta^z \phi(s)\| |\tilde{p}_z(s)| ds \\ &\leq \max_{z \in \Sigma} |\pi_z(t_0)| + C \int_{t_0}^t \max_{z \in \Sigma} |\pi_z(s)| ds \\ &\quad + C \int_{t_0}^t \sum_{z \in \Sigma} \|\Delta^z \phi(s)\| (\tilde{p}_z(s) + M_1) ds + M_1 C \int_{t_0}^t \sum_{z \in \Sigma} \|\Delta^z \phi(s)\| ds, \end{aligned}$$

where the second line holds because  $\tilde{p}_z(s) + M_1 > 0$ . Moreover, we have

$$\begin{aligned} &\int_{t_0}^t \sum_{z \in \Sigma} \|\Delta^z \phi(s)\| (\tilde{p}_z(s) + M_1) ds \\ &\leq \sqrt{\int_{t_0}^t \sum_{z \in \Sigma} \|\Delta^z \phi(s)\|^2 (\tilde{p}_z(s) + M_1) ds} \sqrt{\int_{t_0}^t \sum_{z \in \Sigma} (\tilde{p}_z(s) + M_1) ds}. \end{aligned}$$

Applying Gronwall inequality, as  $\tilde{p}_z(s) \in [-M, 1 + M]$ , there exists  $C$  such that

$$\begin{aligned} \|\pi\| &\leq C \|\pi(t_0)\| + C \sqrt{\int_{t_0}^T \sum_{z \in \Sigma} \|\Delta^z \phi(t)\|^2 (\tilde{p}_z(t) + M_1) dt} + M_1 C \|\phi\| \\ &\leq C \|\pi(t_0)\| + C \sqrt{\|\pi(t_0)\| \|\phi\| + M_1 \|\phi\|^2} + C M_1 \|\phi\| \\ &\leq C \|\pi(t_0)\| + C \|\pi(t_0)\|^{\frac{1}{2}} \|\phi\|^{\frac{1}{2}} + C(M_1 + \sqrt{M_1}) \|\phi\|, \end{aligned} \tag{4.8}$$

where  $C$  also only depends on  $M$  in Proposition 4.1.

We next derive the bound for  $\phi$ . Integrating the first equation in (4.5) over  $[t_0, t]$ , from the Lipschitz continuity of  $G, H$ , there exists  $C$  such that

$$\max_{z \in \Sigma} \phi_z(t) \leq C \max_{z \in \Sigma} |\pi_z(T)| + C \int_{t_0}^t \max_{z \in \Sigma} |\phi_z(s)| ds.$$

Applying Gronwall inequality, there exists constant  $C$  such that

$$\|\phi\| \leq C\|\pi\| \tag{4.9}$$

By combining (4.8) and (4.9), using  $AB \leq \varepsilon A^2 + \frac{1}{\varepsilon} B^2$  for  $A, B > 0$ , there exists  $C$  such that

$$\|\pi\| \leq C\|\pi(t_0)\| + \left[ \frac{1}{4} + C^2(M_1 + \sqrt{M_1}) \right] \|\pi\|.$$

As  $C$  only depend on the boundedness and Lipschitz coefficient of  $H, G, \lambda^*$  and the bound of  $D^2_{\mu\mu} H, \tilde{\theta}, \hat{\theta}$ , which depend on the  $M$  in Proposition 4.1. We only need to select  $M_1$  such that

$$C^2(M_1 + \sqrt{M_1}) < \frac{1}{4},$$

and we can have (4.4). Then it remains to decide the new  $N_0$  such that we have  $\tilde{p}_z(t), \hat{p}(t) > -M_1$  uniformly as we assumed. From Proposition 4.1,  $N_1 := \frac{3e^{\Lambda(M_1)}}{M_1}$  and we can simply define our new  $N_0$  as  $\max N_0, N_1$ . On the other hand, the uniqueness of solution comes directly from (4.4). □

According to Proposition 4.1 and 4.2, take any  $t \in [t_0, T]$  and  $\bar{p}_0 \in \bar{B}(P(\Sigma), \frac{1}{N_0})$  as the initial value for ODE system (2.13), there exists an unique solution  $(\bar{\theta}(s), \bar{p}(s))$  on  $[t, T]$ . Note that  $\bar{\theta}(s)$  might not equal  $\tilde{\theta}(s)$  stated as the solution to (2.13) in Theorem 2.6, since  $\tilde{\theta}$  depends on the values of initial time  $t$  and initial condition  $\bar{p}_0$  chosen above. And we can define a function  $\tilde{U}$  on  $t \in [t_0, T]$  and  $\tilde{p}_0 \in \bar{B}(P(\Sigma), \frac{1}{N_0})$  by the corresponding  $\bar{\theta}(t)$  explained above.

$$\tilde{U}(t, z, \tilde{p}_0) := \bar{\theta}(t, z). \tag{4.10}$$

According to Proposition 4.1 and 4.2,  $\tilde{U}$  is well defined and continuous w.r.t  $\tilde{p}_0$ . Moreover, for  $(\tilde{\theta}, \tilde{p})$ , the solution to (2.13) in Theorem 2.6 on  $[t_0, T]$ , which is the approximated solution we got from DNN and want to estimate error on, we have for all  $t \in [t_0, T]$  that:

$$\tilde{U}(t, z, \tilde{p}(t)) := \tilde{\theta}(t, z).$$

It suggests  $\tilde{U}$  has all information of  $\tilde{\theta}$ . If we can compare  $\tilde{U}$  with the  $U$  defined similar in Cecchin and Pelino [9] corresponding to the true solution to (2.2), we can estimate the error of  $\tilde{\theta}$ . To compare  $\tilde{U}$  with the  $U$ , we need to prove that  $\tilde{U}$  also satisfy the master equation similar to  $U$  in Cecchin and Pelino [9]. To achieve this goal, we are to prove the continuously differentiability of  $\tilde{U}$  in the following steps. We first define the derivative of  $\tilde{U}$  w.r.t vector  $\tilde{p}_0$  in a similar way to in Cecchin and Pelino [9], Define operator  $D^y_p$  as follows.

**Definition 4.3** Define operator of a function  $U : \mathbb{R}^K \rightarrow \mathbb{R}$  as  $D^y U : \mathbb{R}^K \rightarrow \mathbb{R}^K$  for  $y \in \Sigma$ .

$$[D^y U(p)]_z := \lim_{s \rightarrow 0} \frac{U(p + s(\delta_z - \delta_y)) - U(p)}{s},$$

where  $D^y U(p) = ([D^y U(p)]_1, \dots, [D^y U(p)]_K)$ , and  $\delta_z \in \mathbb{R}^K$  such that all elements are 0 except the  $z$  element is 1.

By noticing that  $\mu = \sum_{z \neq 1} \mu_z(\delta_z - \delta_1) + (\sum_{z=1}^K \mu_z)\delta_1$ , if  $\tilde{U}$  is differentiable, we have the following lemma from the linearity of directional derivative.

**Lemma 4.4** Define the derivative of function  $U(p)$  along the direction  $\mu \in \mathbb{R}^K$  as a map  $\frac{\partial}{\partial \mu} U : \mathbb{R}^K \rightarrow \mathbb{R}$ ,

$$\frac{\partial}{\partial \mu} U(p) := \lim_{s \rightarrow 0} \frac{U(p + s\mu) - U(p)}{s}.$$

It satisfies

$$\frac{\partial}{\partial \mu} U(p) = D^1 U(p) \cdot \mu + \frac{\partial}{\partial \delta_1} U(p) \left( \sum_{z=1}^K \mu_z \right),$$

where  $\frac{\partial}{\partial \delta_1}$  is in fact the first component of the gradient of  $\tilde{U}$ , and  $DU(p) := D^1 U(p)$  for notation simplicity. When  $\sum_{z=1}^K \mu_z = 0$ , for any  $y \in \Sigma$ , we have

$$D^y U(p) \cdot \mu = DU(p) \cdot \mu = \frac{\partial}{\partial \mu} U(p).$$

In order to characterize the directional derivative of  $\tilde{U}$  w.r.t  $\tilde{p}_0$ , given  $\tilde{\theta}$  and  $\tilde{p}$ , let us define a linear system of ODE for  $(u, \rho)$  similar to [9, Equation (80)], which will be used quite a few times in the following.

$$\begin{aligned} \frac{du_z(t)}{dt} &= -\lambda^*(z, \Delta^z \tilde{\theta}(t)) \cdot \Delta^z u(t) - b(t, z) \\ \frac{d\rho_z(t)}{dt} &= \sum_y \rho_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) + \sum_y \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot \Delta^y u(t) + c(t, z) \\ u_z(T) &= \frac{\partial G}{\partial \rho(T)}(z, \tilde{p}(T)) + u_{T,z} = \nabla G(z, \tilde{p}(T)) \cdot \rho(T) + u_{T,z} \\ \rho_z(t_0) &= \rho_z \cdot 0. \end{aligned} \tag{4.11}$$

Similar to [9, Equation (80)],  $D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t))$  is the gradient of  $\lambda_z^*$  w.r.t its second variable in  $\mathbb{R}^K$ . The unknowns are  $u$  and  $\rho$ , while  $b, c, u_T, \rho_0$  are given measurable functions, with  $c$  satisfying  $\sum_{z=1}^K c(t, z) = 0$ . In fact, (4.11) is generalization of [9, Equation (80)]. In (4.11), it is a general directional derivatives of any direction in the terminal condition of  $u_z(T)$ , while in [9, Equation (80)], it is directional derivatives of specific directions.

We first prove in Proposition 4.5 that the linear system (4.11) has a unique solution, which is linear bounded by its initial and boundary conditions.

**Proposition 4.5** There exist positive constants  $N_0$  and  $C$ , such that if we have (4.3) and  $\tilde{p}(t_0) \in \tilde{B}(P(\Sigma), \frac{1}{N_0})$ , then for any measurable function  $b, c$  and vector  $u_T$ , the linear system (4.11) has a unique solution  $(u, \rho)$ . Moreover, it satisfies

$$\|u\| + \|\rho\| \leq C[\|u_T\| + \|\rho_0\| + \|b\| + \|c\|]. \tag{4.12}$$

**Proof** We only discuss the case when  $t_0 = 0$ , as it can be extended to any  $t_0 \in [0, T]$  by the same argument.

We first let  $N_0$  bigger than the one in Proposition 4.2. And similar to the proof for Proposition 4.2, to cope with the potential negativeness, we first assume  $\tilde{p}_z(t) \geq -M_1$  uniformly and  $M_1 \leq M$ , and we will decide later the value for  $M_1$  small enough and find the  $N_0$  such that it holds. As  $\sum_{z \in \Sigma} \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) = 0$ , we have  $\sum_{z, y \in \Sigma} \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot \Delta^y u(t) = 0$ ,

and  $\sum_{z \in \Sigma} \frac{d\rho_z(t)}{dt} = 0$ . Hence, for any  $t \in [0, T]$ , we have

$$\eta := \sum_{z \in \Sigma} \rho_z(t) = \sum_{z \in \Sigma} \rho_{z,0}. \tag{4.13}$$

Define set  $P_\eta(\Sigma)$  as

$$P_\eta(\Sigma) := \left\{ p \in \mathbb{R}^K, \quad s.t. \quad \sum_{z=1}^K p_z = \eta \right\}.$$

We define map  $\xi$  from  $C^0([0, T]; P_\eta(\Sigma))$  to itself as follows: for a fixed  $\rho \in C^0([0, T]; P_\eta(\Sigma))$ , we consider the solution  $u = u(\rho)$  to the backward ODE for  $u$  in (4.11), and define  $\xi(\rho)$  to be the solution to the forward ODE for  $\rho$  in (4.11) with  $u = u(\rho)$ . From (4.13),  $\xi(\rho)$  is well defined as  $\xi(\rho)(t) \in P_\eta(\Sigma)$  for any  $t$ .

Similar to the proof for [9, Proposition 6], the solution to (4.11) is the fixed point of mapping  $\xi$ , and we prove its existence by Schaefer’s fixed point theorem, which asserts that a continuous and compact mapping  $\xi$  of a Banach space  $X$  into itself has fixed point if the set  $\{\rho \in X : \rho = \omega\xi(\rho), \omega \in [0, 1]\}$  is bounded. Firstly,  $\xi$  is continuous as the system (4.11) is linear in  $u$  and  $\rho$ .  $C^0([0, T]; P_\eta(\Sigma))$  is a convex subset of Banach space  $C^0([0, T]; \mathbb{R}^K)$ . Moreover, from the linearity and bounded coefficients of system (4.11),  $\xi$  maps any bounded set of  $C^0([0, T]; P_\eta(\Sigma))$  into set of bounded and Lipschitz continuous functions with uniform Lipschitz coefficient in  $C^1([0, T]; P_\eta(\Sigma))$ , which by Arzela–Ascoli theorem, is relatively compact. By compact map definition,  $\xi$  is a compact map. Hence, to apply Schaefer’s fixed point theorem, it remains to prove that the set  $\{\rho : \rho = \omega\xi(\rho)\}$  is uniform bounded for  $\forall \omega \in [0, 1]$ . We can restrict to  $\omega > 0$  since otherwise  $\rho = 0$ . Fix a  $\rho$  such that  $\rho = \omega\xi(\rho)$ , which means the couple  $(u(\rho), \xi(\rho))$  solves (for notation simplicity we neglect their dependency on  $\rho$ )

$$\begin{aligned} \frac{du_z(t)}{dt} &= -\lambda^*(z, \Delta^z \tilde{\theta}(t)) \cdot \Delta^z u(t) - b(t, z) \\ \frac{d\xi_z(t)}{dt} &= \sum_y \xi_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \\ &\quad + \sum_y \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot \Delta^y u(t) + c(t, z) \\ u_z(T) &= \nabla G(z, \tilde{p}(T)) \cdot \omega \xi(T) + u_{T,z} \\ \xi_z(t_0) &= \rho_{z,0}. \end{aligned} \tag{4.14}$$

We need to prove the solution  $(u, \xi)$  if existed, are bounded uniformly for any  $\omega \in (0, 1]$ . For notation simplicity, we omit the dependence of  $\lambda^*$  on the second variable. From (4.14),

$$\begin{aligned} \sum_{z \in \Sigma} \frac{d}{dt} (u_z(t) \xi_z(t)) &= - \sum_{z,y \in \Sigma} \xi_z(t) \lambda_y^*(z) (u_y(t) - u_z(t)) + \sum_{z,y \in \Sigma} \xi_y(t) \lambda_z^*(y) u_z(t) \\ &\quad + \sum_{z,y \in \Sigma} u_z(t) \tilde{p}_y(t) D_\mu \lambda_z^*(y) \cdot \Delta^y u(t) + \sum_{z \in \Sigma} c(t, z) u_z(t) - \sum_{z \in \Sigma} \xi_z(t) b(t, z). \end{aligned}$$

The first line is 0 by exchanging  $z$  and  $y$  in the second double sum and using (2.3). Integrating over  $[0, T]$  and using the expression of  $u_z(T)$ , we have

$$\begin{aligned} & \sum_{z \in \Sigma} \xi_z(T) [\nabla G(z, \tilde{p}(T)) \cdot \omega \xi(T) + u_{T,z}] - u(0) \cdot \rho_0 \\ &= \int_0^T \sum_{z \in \Sigma} c(t, z) u_z(t) dt - \int_0^T \sum_{z \in \Sigma} \xi_z(t) b(t, z) dt \\ & \quad + \int_0^T \sum_{z, y \in \Sigma} \tilde{p}_y(t) D_\mu \lambda_z^*(y) \cdot \Delta^y u(t) (u_z(t) - u_y(t)) dt. \end{aligned}$$

Reorganize the terms and we get

$$\begin{aligned} & \int_0^T \sum_{z, y \in \Sigma} \tilde{p}_y(t) D_\mu \lambda_z^*(y) \cdot \Delta^y u(t) (u_z(t) - u_y(t)) dt - \omega \sum_{z \in \Sigma} \xi_z(T) \nabla G(z, \tilde{p}(T)) \cdot \xi(T) \\ &= \int_0^T \sum_{z \in \Sigma} \xi_z(t) b(t, z) dt - \int_0^T \sum_{z \in \Sigma} c(t, z) u_z(t) dt + \sum_{z \in \Sigma} \xi_z(T) u_{T,z} - u(0) \cdot \rho_0. \end{aligned}$$

From assumption on  $G$  in (2.6) and definition of directional derivative, we have

$$-\omega \sum_{z \in \Sigma} \xi_z(T) \nabla G(z, \tilde{p}(T)) \cdot \xi(T) = -\omega \sum_{z \in \Sigma} \xi_z(T) \frac{\partial G}{\partial \xi}(z, \tilde{p}(T)) \geq 0.$$

Moreover, as  $\lambda^*(y) = D_\mu H(y)$  (we also neglect the dependence of  $H$  on the second variable),

$$\begin{aligned} & \int_0^T \sum_{z, y \in \Sigma} \tilde{p}_y(t) D_\mu \lambda_z^*(y) \cdot \Delta^y u(t) (u_z(t) - u_y(t)) dt \\ &= \int_0^T \sum_{y \in \Sigma} \tilde{p}_y(t) \Delta^y u(t) \cdot D_{\mu\mu}^2 H(y) \cdot \Delta^y u(t) dt. \end{aligned}$$

Since  $\tilde{p}$  and  $\hat{p}$  can be negative, the same step in [9, Proposition 6] to obtain estimation on RHS of above is not applicable. However, as  $\tilde{p}_y(t) + M_1 \geq 0$  for all  $y \in \Sigma$ , from (2.4), we can rewrite the RHS of the equation and get the following estimation instead.

$$\begin{aligned} & \int_0^T \sum_{y \in \Sigma} \tilde{p}_y(t) \Delta^y u(t) \cdot D_{\mu\mu}^2 H(y) \cdot \Delta^y u(t) dt \\ &= \int_0^T \sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \Delta^y u(t) \cdot D_{\mu\mu}^2 H(y) \cdot \Delta^y u(t) dt \\ & \quad - M_1 \int_0^T \sum_{y \in \Sigma} \Delta^y u(t) \cdot D_{\mu\mu}^2 H(y) \cdot \Delta^y u(t) dt \\ &\geq C^{-1} \int_0^T \sum_{z \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^z u(t)\|^2 dt \\ & \quad - M_1 C \int_0^T \sum_{z \in \Sigma} \|\Delta^z u(t)\|^2 dt. \end{aligned}$$



So there exists constant  $C$  and  $C_1$  ( $C_1$  only depends on the dimension of  $u$ ) such that

$$\int_0^T \sum_{z \in \Sigma} (\tilde{p}_z(t) + M_1) \|\Delta^z u(t)\|^2 dt \leq C \left( \int_0^T |c(t) \cdot u(t)| dt + \int_0^T |\xi(t) \cdot b(t)| dt + \|\xi(T)\| \|u_T\| + \|u(0)\| \|\rho_0\| + M_1 C_1 \|u\|^2 \right), \tag{4.15}$$

where  $b(t) := (b(t, 1), \dots, b(t, K))$ , and  $c(t)$  is defined similarly. As  $\lambda^*$  and  $D_\mu \lambda^*$  is bounded by constant  $C$ , from ODE for  $\xi$  in (4.14) we have

$$\begin{aligned} |\xi_z(t)| &\leq |\rho_{0,z}| + C \int_0^t \sum_{y \in \Sigma} |\xi_y(s)| ds \\ &+ C \int_0^t \left[ \sum_{y \in \Sigma} (\tilde{p}_y(s) + M_1) \|\Delta^y u(s)\| + |c(s, z)| \right] ds \\ &+ C M_1 \int_0^t \sum_{y \in \Sigma} \|\Delta^y u(s)\| ds. \end{aligned}$$

So that by Gronwall’s inequality, there exists constant  $C$  such that

$$\|\xi\| \leq C(\|\rho_0\| + \|c\|) + C \int_0^T \sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\| dt + C M_1 \int_0^T \sum_{y \in \Sigma} \|\Delta^y u(t)\| dt,$$

where there exists  $C$  such that  $\sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \leq C^2$  and

$$\begin{aligned} \int_0^T \sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\| dt &= \int_0^T \sum_{y \in \Sigma} \sqrt{\tilde{p}_y(t) + M_1} \sqrt{\tilde{p}_y(t) + M_1} \|\Delta^y u(t)\| dt \\ &\leq \int_0^T \sqrt{\sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1)} \sqrt{\sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\|^2} dt \\ &\leq C \int_0^T \sqrt{\sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\|^2} dt \\ &\leq C \sqrt{\int_0^T \sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\|^2 dt}. \end{aligned}$$

From (4.15), there exist different constants  $C$  at each line such that

$$\begin{aligned} \|\xi\| &\leq C(\|\rho_0\| + \|c\|) + C \int_0^T \sum_{y \in \Sigma} (\tilde{p}_y(t) + M_1) \|\Delta^y u(t)\| dt + C M_1 \int_0^T \sum_{y \in \Sigma} \|\Delta^y u(t)\| dt \\ &\leq C(\|\rho_0\| + \|c\|) + C(M_1 + \sqrt{M_1}) \|u\| \\ &\quad + C(\|c\|^{1/2} \|u\|^{1/2} + \|\xi\|^{1/2} \|b\|^{1/2} + \|\xi(T)\|^{1/2} \|u_T\|^{1/2} + \|u(0)\|^{1/2} \|\rho_0\|^{1/2}). \end{aligned}$$

Moreover, using Gronwall inequality on the backward ODE in (4.14) for function  $u$ , there exists  $C$  such that

$$\|u\| \leq C[\|u_T\| + \|\xi(T)\| + \|b\|].$$

Then there exists  $C$  such that

$$\begin{aligned} \|\xi\| \leq & C(\|\rho_0\| + \|c\|) + C(M_1 + \sqrt{M_1})(\|u_T\| + \|\xi(T)\| + \|b\|) + C\|c\|^{\frac{1}{2}}(\|u_T\| + \|\xi(T)\| + \|b\|)^{\frac{1}{2}} \\ & + C(\|\xi\|^{\frac{1}{2}}\|b\|^{\frac{1}{2}} + \|\xi(T)\|^{\frac{1}{2}}\|u_T\|^{\frac{1}{2}} + (\|u_T\|^{\frac{1}{2}} + \|\xi(T)\|^{\frac{1}{2}} + \|b\|^{\frac{1}{2}})\|\rho_0\|^{\frac{1}{2}}) \end{aligned}$$

As  $\|\xi(T)\| \leq \|\xi\|$ , using the inequality  $AB \leq \varepsilon A^2 + \frac{1}{4\varepsilon} B^2$  for  $A, B \geq 0$ , there exists  $C$  such that

$$\|\xi\| \leq C(\|c\| + \|b\| + \|\rho_0\| + \|u_T\|) + \left( C(M_1 + \sqrt{M_1}) + \frac{1}{4} \right) \|\xi\|.$$

Note that the constant  $C$  only depends on the boundedness of  $\tilde{\theta}$ , which depends on  $M$  in Proposition 4.1. If

$$C(M_1 + \sqrt{M_1}) \leq \frac{1}{4},$$

then we have

$$\|\xi\| \leq 2C(\|c\| + \|b\| + \|\rho_0\| + \|u_T\|).$$

Hence, the solution pair  $(u, \xi)$  are bounded for all  $\omega \in [0, 1]$ , which means  $\rho = \omega\xi(\rho)$  are also uniform bounded, and hence prove the existence of solution to (4.11). Meanwhile, let  $\omega = 1$  leads to the uniform bound estimation for solution  $(u, \rho)$  to (4.11), and the uniqueness of it comes directly from (4.12). If  $N_0 > \frac{3e^{\Lambda(M_1)}}{M_1}$ , from Proposition 4.1, we have  $\tilde{p}_y(t) > -M_1$  uniformly, which concludes our proof. Hence, we can just update our  $N_0$  set before to satisfy the inequality. □

Then we can prove the differentiability of  $\tilde{U}$  w.r.t  $\tilde{\rho}_0$  in Proposition 4.6.

**Proposition 4.6** *Let  $(\tilde{\theta}, \tilde{p})$  and  $(\hat{\theta}, \hat{p})$  be the solutions to ODE system (2.13), respectively, starting from  $(t_0, \tilde{p}(t_0))$  and  $(t_0, \hat{p}(t_0))$ , and  $(v, \zeta)$  be the solution to (4.11) starting from  $\rho_0 := \hat{p}(t_0) - \tilde{p}(t_0)$ . There exist positive constants  $N_0$  and  $C$ , such that if we have (4.3), then for any  $t_0 \in [0, T]$  and  $\tilde{p}(t_0), \hat{p}(t_0) \in \tilde{B}(P(\Sigma), \frac{1}{N_0})$ , we have*

$$\|\hat{\theta} - \tilde{\theta} - v\| + \|\hat{p} - \tilde{p} - \zeta\| \leq C\|\hat{p}(t_0) - \tilde{p}(t_0)\|^2.$$

**Proof** Without loss of generality, we assume  $t_0 = 0$ . Similar to the proof of [9, Theorem 7], we can use results from Proposition 4.5 to prove our conclusion. Define  $N_0$  as the one in Proposition 4.5. Then  $\tilde{p}_y(t), \hat{p}_y(t) > -M_1$  uniformly on  $(t, y) \in [0, T] \times \Sigma$ . Define linearized system with  $w := \hat{p}(0) - \tilde{p}(0)$ :

$$\begin{aligned} \frac{dv_z(t)}{dt} &= -\lambda^*(z, \Delta^z \tilde{\theta}(t)) \cdot \Delta^z v(t) \\ \frac{d\zeta_z(t)}{dt} &= \sum_y \zeta_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) + \sum_y \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot \Delta^y v(t) \\ v_z(T) &= \frac{\partial G}{\partial \zeta(T)}(z, \tilde{p}(T)) = D^1 G(z, \tilde{p}(T)) \cdot \zeta(T) + \frac{\partial G}{\partial \delta_1}(z, \tilde{p}(T)) \sum_{z=1}^K w_z \\ \zeta_z(0) &= w_z. \end{aligned} \tag{4.16}$$

From condition in Theorem 2.6, the sum of every component of  $\tilde{p}$  equals 1 for all  $t \in [0, T]$ . Hence, we know  $\sum_{z \in \Sigma} \epsilon_2(t, z) = 0$  and define

$$S(\hat{p}, \tilde{p}) := \sum_{z \in \Sigma} (\hat{p}_z(0) - \tilde{p}_z(0)) = \sum_{z \in \Sigma} (\hat{p}_z(T) - \tilde{p}_z(T)).$$

We know there exists  $C$  such that  $|S(\hat{p}, \tilde{p})| \leq C \|\hat{p}(T) - \tilde{p}(T)\|$ . Set  $u := \hat{\theta} - \tilde{\theta} - v$  and  $\rho := \hat{p} - \tilde{p} - \zeta$ , they solve (4.11), where

$$\begin{aligned} b(t, z) &:= H(z, \Delta^z \hat{\theta}(t)) - H(z, \Delta^z \tilde{\theta}(t)) - \lambda^*(z, \Delta^z \tilde{\theta}(t)) \cdot (\Delta^z \hat{\theta}(t) - \Delta^z \tilde{\theta}(t)) \\ c(t, z) &:= \sum_y \hat{p}_y(t) [\lambda_z^*(y, \Delta^y \hat{\theta}(t)) - \lambda_z^*(y, \Delta^y \tilde{\theta}(t))] \\ &\quad - \sum_y \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot (\Delta^y \hat{\theta}(t) - \Delta^y \tilde{\theta}(t)) \\ u_{T,z} &:= G(z, \hat{p}(T)) - G(z, \tilde{p}(T)) - D^1 G(z, \tilde{p}(T))(\hat{p}(T) - \tilde{p}(T)) \\ &\quad - \frac{\partial G}{\partial \delta_1}(z, \tilde{p}(T)) S(\hat{p}, \tilde{p}). \end{aligned}$$

From (2.3),  $\sum_{z \in \Sigma} c(t, z) = 0$ . The existence and uniqueness of solution to (4.16) is guaranteed by Proposition 4.5. We can simplify  $b$  and  $c$  as

$$\begin{aligned} b(t, z) &= \int_0^1 [D_\mu H(z, \Delta^z \tilde{\theta}(t) + s(\Delta^z \hat{\theta}(t) - \Delta^z \tilde{\theta}(t))) - D_\mu H(z, \Delta^z \tilde{\theta}(t))] \\ &\quad \cdot (\Delta^z \hat{\theta}(t) - \Delta^z \tilde{\theta}(t)) ds \\ c(t, z) &= \sum_y \hat{p}_y(t) \int_0^1 [D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t) + s(\hat{\theta}(t) - \Delta^y \tilde{\theta}(t))) \\ &\quad - D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t))] \cdot (\Delta^z \hat{\theta}(t) - \Delta^z \tilde{\theta}(t)) ds \\ &\quad + \sum_y (\hat{p}_y(t) - \tilde{p}_y(t)) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot (\Delta^y \hat{\theta}(t) - \Delta^y \tilde{\theta}(t)). \end{aligned}$$

Moreover, since

$$\begin{aligned} G(z, \hat{p}(T)) - G(z, \tilde{p}(T)) &= \int_0^1 \frac{\partial G}{\partial (\hat{p}(T) - \tilde{p}(T))}(z, \tilde{p}(T) + s(\hat{p}(T) - \tilde{p}(T))) ds \\ &= \int_0^1 D^1 G(z, \tilde{p}(T) + s(\hat{p}(T) - \tilde{p}(T))) \cdot ((\hat{p}(T) - \tilde{p}(T))) ds \\ &\quad + \int_0^1 \frac{\partial G}{\partial \delta_1}(z, \tilde{p}(T) + s(\hat{p}(T) - \tilde{p}(T))) S(\hat{p}, \tilde{p}) ds \end{aligned}$$

we have

$$\begin{aligned} u_{T,z} &= \int_0^1 (D^1 G(z, \tilde{p}(T) + s(\hat{p}(T) - \tilde{p}(T))) - D^1 G(z, \tilde{p}(T))) \cdot ((\hat{p}(T) - \tilde{p}(T))) ds \\ &\quad + \int_0^1 (\frac{\partial G}{\partial \delta_1}(z, \tilde{p}(T) + s(\hat{p}(T) - \tilde{p}(T))) - \frac{\partial G}{\partial \delta_1}(z, \tilde{p}(T))) S(\hat{p}, \tilde{p}) ds. \end{aligned}$$

From Proposition 4.1,  $\tilde{\theta}$ ,  $\tilde{p}$ ,  $\hat{\theta}$  and  $\hat{p}$  are bounded. From Assumption 2.1, namely the Lipschitz continuity of  $D_\mu H$ ,  $D_\mu \lambda^*$ ,  $\frac{\partial G}{\partial \delta_1}$  and  $D^1 G$  in their second variable, there exists constant  $C$

such that

$$\begin{aligned} \|b\| &\leq C \|\tilde{\theta} - \hat{\theta}\|^2 \\ \|u_{T,z}\| &\leq C \|\tilde{p}(T) - \hat{p}(T)\|^2 \\ \|c\| &\leq C(\|\tilde{\theta} - \hat{\theta}\|^2 + \|\tilde{\theta} - \hat{\theta}\| \cdot \|\tilde{p} - \hat{p}\|). \end{aligned}$$

Applying Proposition 4.5 and then Proposition 4.2, we have there exists  $C$  such that

$$\|u\| + \|\rho\| \leq C \|\hat{p}(0) - \tilde{p}(0)\|^2,$$

which concludes the proof. □

As (4.16) is a linear system.  $v$  and  $\zeta$  in (4.16) can be viewed as a linear map of  $w$ . Hence, Proposition 4.6 suggests that  $\tilde{U}$  is differentiable w.r.t  $\tilde{p}_0$  and the directional derivative  $\frac{\partial}{\partial w} \tilde{U}(t, z, \tilde{p})$  is the solution to ODE system (4.16), with  $\tilde{\theta}_z(t) = \tilde{U}(t, z, \tilde{p}(t))$ .

**Theorem 4.7** *There exist positive constants  $N_0$  and  $C$ , such that if we have (4.3),  $\tilde{U}$  is differentiable on  $B(P(\Sigma), \frac{1}{N_0})$ , and for any vector  $w$ ,  $\frac{\partial}{\partial w} \tilde{U}(t, z, \tilde{p}(t))$  exists and is Lipschitz continuous w.r.t  $\tilde{p}$ , uniformly in  $t, z$ .  $\frac{\partial}{\partial w} \tilde{U}(t, z, \tilde{p}(t))$  is also continuous w.r.t  $t$ .*

**Proof** Define  $N_0$  as the one in Proposition 4.5. Let  $(\tilde{\theta}, \tilde{p})$  and  $(\hat{\theta}, \hat{p})$  be two solutions to (2.13), with initial conditions  $\tilde{p}(t_0), \hat{p}(t_0) \in B(P(\Sigma), \frac{1}{N_0})$ . Let also  $(\tilde{v}, \tilde{\zeta})$  and  $(\hat{v}, \hat{\zeta})$  characterize  $\frac{\partial}{\partial w} \tilde{U}(t_0, z, \tilde{p}(t_0))$  and  $\frac{\partial}{\partial w} \tilde{U}(t_0, z, \hat{p}(t_0))$ , respectively. Then  $(\tilde{v}, \tilde{\zeta})$  satisfies the following.

$$\begin{aligned} \frac{d\tilde{v}_z(t)}{dt} &= -\lambda^*(z, \Delta^z \tilde{\theta}(t)) \cdot \Delta^z \tilde{v}(t) \\ \frac{d\tilde{\zeta}_z(t)}{dt} &= \sum_y \tilde{\zeta}_y(t) \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) + \sum_y \tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) \cdot \Delta^y \tilde{v}(t) \\ \tilde{v}_z(T) &= \frac{\partial G}{\partial \tilde{\zeta}(T)}(z, \tilde{p}(T)) \\ \tilde{\zeta}_z(t_0) &= w_z. \end{aligned} \tag{4.17}$$

From Proposition 4.5, we know the uniform bound of both  $\tilde{v}$  and  $\tilde{\zeta}$  depend linearly on norm of  $w$ . Similar is for  $(\hat{v}, \hat{\zeta})$ , except for replacing  $(\tilde{\theta}, \tilde{p})$  by  $(\hat{\theta}, \hat{p})$ . Set  $u := \tilde{v} - \hat{v}$ ,  $\rho := \tilde{\zeta} - \hat{\zeta}$ . They solve the linear system (4.11) with  $\rho(t_0) = 0$  and

$$\begin{aligned} b(t, z) &:= (\lambda^*(z, \Delta^z \tilde{\theta}(t)) - \lambda^*(z, \Delta^z \hat{\theta}(t))) \cdot \Delta^z \hat{v}(t) \\ c(t, z) &:= \sum_{y \in \Sigma} \hat{\zeta}_y(t) (\lambda_z^*(y, \Delta^y \tilde{\theta}(t)) - \lambda_z^*(y, \Delta^y \hat{\theta}(t))) \\ &\quad + \sum_{y \in \Sigma} [\tilde{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \tilde{\theta}(t)) - \hat{p}_y(t) D_\mu \lambda_z^*(y, \Delta^y \hat{\theta}(t))] \cdot \Delta^z \hat{v}(t) \\ u_{T,z} &:= \frac{\partial G}{\partial \hat{\zeta}(T)}(z, \tilde{p}(T)) - \frac{\partial G}{\partial \hat{\zeta}(T)}(z, \hat{p}(T)). \end{aligned}$$

Using the Lipschitz continuity of  $\lambda^*$ ,  $D_\mu \lambda^*$  and directional derivatives of  $G$ , applying the bounds (4.12) to  $\hat{v}$  and  $\hat{\zeta}$ , and the estimation on  $\|\tilde{\theta} - \hat{\theta}\|, \|\tilde{p} - \hat{p}\|$  from Proposition 4.2, there

exists  $C$  such that

$$\begin{aligned} \|b\| &\leq C\|\tilde{\theta} - \hat{\theta}\|\|\hat{v}\| \leq C\|\tilde{p}(t_0) - \hat{p}(t_0)\|\|w\| \\ \|c\| &\leq C\|\tilde{\theta} - \hat{\theta}\|\|\hat{\zeta}\| + C\|\tilde{\theta} - \hat{\theta}\|\|\hat{v}\| + C\|\tilde{p} - \hat{p}\|\|\hat{v}\| \\ &\leq C\|\tilde{p}(t_0) - \hat{p}(t_0)\|\|w\| \\ \|u_T\| &\leq C\|\tilde{p} - \hat{p}\|\|\hat{\zeta}\| \leq C\|\tilde{p}(t_0) - \hat{p}(t_0)\|\|w\| \end{aligned}$$

From Proposition 4.5, we have

$$\|u\| \leq C(\|b\| + \|c\| + \|u_T\|) \leq C\|\tilde{p}(t_0) - \hat{p}(t_0)\|\|w\|.$$

From Proposition 4.6, we have

$$\tilde{v}_z(t_0) = \frac{\partial \tilde{U}}{\partial w}(t_0, z, \tilde{p}(t_0)), \quad \hat{v}_z(t_0) = \frac{\partial \tilde{U}}{\partial w}(t_0, z, \hat{p}(t_0)).$$

Therefore,  $\frac{\partial \tilde{U}}{\partial w}$  is Lipschitz continuous, uniform w.r.t  $t$  and  $z$ .

On the other hand, for another initial time  $t_1 > t_0$ , we first compare  $\frac{\partial}{\partial w} \tilde{U}(t_0, z, \tilde{p}(t_0))$  and  $\frac{\partial}{\partial w} \tilde{U}(t_1, z, \tilde{p}(t_1))$ , where  $(t_1, \tilde{p}(t_1))$  is on the path  $(t, \tilde{p}(t))$  start from  $t_0$  to  $T$ . They are both characterized by system like (4.17), though we need to replace  $t_0$  with  $t_1$  for  $\frac{\partial}{\partial w} \tilde{U}(t_1, z, \tilde{p}(t_1))$ . Let  $(\tilde{v}, \tilde{\zeta})$  satisfy (4.17). Then we know

$$\tilde{v}(t_0) = \frac{\partial}{\partial w} \tilde{U}(t_0, z, \tilde{p}(t_0)), \quad \tilde{v}(t_1) = \frac{\partial}{\partial \tilde{\zeta}(t_1)} \tilde{U}(t_1, z, \tilde{p}(t_1)).$$

$\frac{\partial}{\partial \tilde{\zeta}(t_1)} \tilde{U}(t_1, z, \tilde{p}(t_1))$  is also characterized by (4.17), except that  $t_0$  and initial value need to be replaced by  $t_1$  and  $\tilde{\zeta}(t_1)$ . It means  $\frac{\partial}{\partial \tilde{\zeta}(t_1)} \tilde{U}(t_1, z, \tilde{p}(t_1)) - \frac{\partial}{\partial w} \tilde{U}(t_1, z, \tilde{p}(t_1))$  can also be characterized by (4.17) except that  $t_0$  and initial value need to be replaced by  $t_1$  and  $\tilde{\zeta}(t_1) - w$ . From Proposition 4.5, we have there exists constant  $C$  such that

$$\left| \frac{\partial}{\partial \tilde{\zeta}(t_1)} \tilde{U}(t_1, z, \tilde{p}(t_1)) - \frac{\partial}{\partial w} \tilde{U}(t_1, z, \tilde{p}(t_1)) \right| \leq C|\tilde{\zeta}(t_1) - w|.$$

As  $\lambda^*$ ,  $D_\mu \lambda^*$  and the directional derivative of  $G$  are Lipschitz continuous and uniform bounded, as well as that both  $\tilde{v}$  and  $\tilde{\zeta}$  are uniformly bounded, we know hence both  $\frac{d\tilde{v}_z(t)}{dt}$  and  $\frac{d\tilde{\zeta}_z(t)}{dt}$  are also uniformly bounded by some constant  $C$ . We have

$$\begin{aligned} \|\tilde{\zeta}(t_1) - w\| &= \|\tilde{\zeta}(t_1) - \tilde{\zeta}(t_0)\| \leq C|t_1 - t_0|, \\ \left| \frac{\partial}{\partial w} \tilde{U}(t_0, z, \tilde{p}(t_0)) - \frac{\partial}{\partial \tilde{\zeta}(t_1)} \tilde{U}(t_1, z, \tilde{p}(t_1)) \right| &= |\tilde{v}_z(t_0) - \tilde{v}_z(t_1)| \leq C|t_1 - t_0|. \end{aligned}$$

Combine above, we know there exists constant  $C$  such that

$$\left| \frac{\partial}{\partial w} \tilde{U}(t_0, z, \tilde{p}(t_0)) - \frac{\partial}{\partial w} \tilde{U}(t_1, z, \tilde{p}(t_1)) \right| \leq C|t_1 - t_0|.$$

Then by the continuity of  $\frac{\partial}{\partial w} \tilde{U}$  w.r.t its third argument, as well as the continuity of  $\tilde{p}$ , we can also conclude that  $\frac{\partial}{\partial w} \tilde{U}$  is continuous w.r.t  $t$ , its first argument. □

From Proposition 4.6 and Theorem 4.7,  $\tilde{U}$  is  $C^1$  on compact set  $\bar{B}(P(\Sigma), \frac{1}{N_0})$ . Hence, both  $D\tilde{U}$  and the directional derivative of  $\tilde{U}$  along any direction are well defined, bounded, and

Lipschitz continuous, uniformly for  $t \in [0, T]$ . Theorem 4.7 also suggests that the directional derivative of  $\tilde{U}$  along any direction is continuous w.r.t  $t$ . Thanks to these properties, we can use similar idea of the proof for existence of solution to the master equation in [9, Section 5.3.1], to show that  $\tilde{U}$  also satisfies the master equation with some extra error terms.

**Theorem 4.8** *Let  $(\tilde{\theta}, \tilde{p})$  be the solution to ODE system (2.13). Define  $\tilde{U}$  as (4.10). There exist positive constants  $N_0$  and  $C$ , such that if we have condition (4.3) in Theorem 2.6, then  $\tilde{U}$  satisfies the following master equation along the path  $(t, \tilde{p}(t))$  on  $[t_0, T]$ , as long as  $\tilde{p}(t) \in B(P(\Sigma), \frac{1}{N_0})$ .*

$$\begin{aligned} & \frac{\partial \tilde{U}(t, z, \tilde{p}(t))}{\partial t} + H(z, \Delta^z \tilde{U}) + \sum_{y \in \Sigma} \tilde{p}_y(t) \lambda^*(y, \Delta^y \tilde{U}) \cdot D\tilde{U}(t, z, \tilde{p}(t)) \\ & = \epsilon(t, z) \tilde{U}(T, z, \tilde{p}(T)) = G(z, \tilde{p}(T)) + \epsilon_3(z), \end{aligned} \tag{4.18}$$

where  $\Delta^z \tilde{U} := (\tilde{U}(t, 1, \tilde{p}(t)) - \tilde{U}(t, z, \tilde{p}(t)), \dots, \tilde{U}(t, K, \tilde{p}(t)) - \tilde{U}(t, z, \tilde{p}(t)))$  and  $\|\epsilon\| < \frac{C+1}{N}$ , where  $N > N_0$  and  $C$  comes from the uniform bound coefficient in Proposition 4.2.

**Proof** From condition in Theorem 2.6,  $\tilde{p}(t) \in B(P(\Sigma), \frac{1}{N_0})$  for every  $t \in [t_0, T]$  where  $B(P(\Sigma), \frac{1}{N_0})$  being the open neighborhood of  $P(\Sigma)$ . Hence, from Proposition 4.1, 4.2 and Theorem 4.7,  $\tilde{U}$ ,  $D\tilde{U}$  and  $\frac{\partial}{\partial \delta_1} \tilde{U}$  are well defined on  $(t, \tilde{p}(t))$ . Take  $t$  as initial time and  $\tilde{p}(t)$  as initial value, there exists an unique solution to (2.13), and we can always choose  $h$  small enough such that this solution taking value on  $t + h$ , i.e.,  $\tilde{p}(t + h) \in B(P(\Sigma), \frac{1}{N_0})$ . Note that as  $\sum_{z \in \Sigma} \epsilon_2(t, z) = 0$  for all  $t \in [t_0, T]$ , we have

$$\sum_{z \in \Sigma} \tilde{p}_z(t) = \sum_{z \in \Sigma} \tilde{p}_z(t + h).$$

Let us first compute limit of the following when  $h$  tends to 0.

$$\begin{aligned} & \frac{\tilde{U}(t + h, z, \tilde{p}(t)) - \tilde{U}(t, z, \tilde{p}(t))}{h} \\ & = \frac{\tilde{U}(t + h, z, \tilde{p}(t)) - \tilde{U}(t + h, z, \tilde{p}(t + h))}{h} \\ & \quad + \frac{\tilde{U}(t + h, z, \tilde{p}(t + h)) - \tilde{U}(t, z, \tilde{p}(t))}{h} \end{aligned} \tag{4.19}$$

For the first term in (4.19), we first define

$$W(s) := \tilde{U}(t + h, z, \tilde{p}(t) + s(\tilde{p}(t + h) - \tilde{p}(t))).$$

By definition, we derive the derivative of  $W$  as

$$W'(s) = \frac{\partial}{\partial (\tilde{p}(t + h) - \tilde{p}(t))} \tilde{U}(t + h, z, \tilde{p}(t) + s(\tilde{p}(t + h) - \tilde{p}(t))).$$

Then the first term in (4.19) can be reformulated as

$$\frac{\tilde{U}(t + h, z, \tilde{p}(t)) - \tilde{U}(t + h, z, \tilde{p}(t + h))}{h} = \frac{W(0) - W(1)}{h} = -\frac{1}{h} \int_0^1 W'(s) ds.$$

From Lemma 4.4 and  $c(h) = \sum_{z=1}^K (\tilde{p}_z(t + h) - \tilde{p}_z(t)) = 0$ , we know

$$W'(s) = D\tilde{U}(t + h, z, \tilde{p}(t) + s(\tilde{p}(t + h) - \tilde{p}(t))) \cdot (\tilde{p}(t + h) - \tilde{p}(t))$$

Substitute above to the first term in (4.19), we get

$$\begin{aligned} & \frac{\tilde{U}(t+h, z, \tilde{p}(t)) - \tilde{U}(t+h, z, \tilde{p}(t+h))}{h} \\ &= -\frac{1}{h} \int_0^1 D\tilde{U}(t+h, z, \tilde{p}(t) + s(\tilde{p}(t+h) - \tilde{p}(t))) \cdot (\tilde{p}(t+h) - \tilde{p}(t)) ds \\ &= -\frac{1}{h} \int_0^1 \tilde{U}(t+h, z, \tilde{p}(t) + s(\tilde{p}(t+h) - \tilde{p}(t))) ds \\ & \quad \cdot \int_t^{t+h} \left( \sum_y \tilde{p}_y(u) \lambda^*(y, \Delta^y \tilde{\theta}(u)) + \epsilon_2(u) \right) du, \end{aligned}$$

where  $\epsilon_2(t) := (\epsilon_2(t, 1), \dots, \epsilon_2(t, z))$ . From Theorem 4.7, we know for any  $y \in \Sigma$ ,

$$\lim_{h \rightarrow 0} [D\tilde{U}(t+h, z, \tilde{p}(t) + s(\tilde{p}(t+h) - \tilde{p}(t)))]_y = [D\tilde{U}(t, z, \tilde{p}(t))]_y.$$

As  $D\tilde{U}$  is uniform bounded, we have the following with dominated convergence theorem:

$$\lim_{h \rightarrow 0} \int_0^1 D\tilde{U}(t+h, z, \tilde{p}(t) + s(\tilde{p}(t+h) - \tilde{p}(t))) ds = D\tilde{U}(t, z, \tilde{p}(t)).$$

On the other hand, dividing  $h$  and letting  $h \rightarrow 0$ , we have the following:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\int_t^{t+h} (\sum_y \tilde{p}_y(u) \lambda^*(y, \Delta^y \tilde{\theta}(u)) + \epsilon_2(u)) du}{h} &= \sum_y \tilde{p}_y(t) \lambda^*(y, \Delta^y \tilde{\theta}(t)) + \epsilon_2(t) \\ &= \sum_y \tilde{p}_y(t) \lambda^*(y, \Delta^y \tilde{U}) + \epsilon_2(t). \end{aligned}$$

The last equation comes from Definition of  $\tilde{U}$ , which suggests  $\Delta^y \tilde{U} = \Delta^y \tilde{\theta}(t)$ .

For the second term in (4.19), from definition of  $\tilde{U}$ , we know

$$\tilde{U}(t+h, z, \tilde{p}(t+h)) - \tilde{U}(t, z, \tilde{p}) = \frac{d\tilde{\theta}_z(t)}{dt} h + o(h),$$

and hence,

$$\lim_{h \rightarrow 0} \frac{\tilde{U}(t+h, z, \tilde{p}(t+h)) - \tilde{U}(t, z, \tilde{p}(t))}{h} = \frac{d\tilde{\theta}_z(t)}{dt} = -H(z, \Delta^z \tilde{U}) + \epsilon_1(t, z).$$

Combining both the results from first and second term in (4.19), taking  $h \rightarrow 0$ , we have

$$\begin{aligned} \frac{\partial \tilde{U}(t, z, \tilde{p}(t))}{\partial t} &= -H(z, \Delta^z \tilde{U}) - D\tilde{U}(t, z, \tilde{p}(t)) \\ & \quad \cdot \left( \sum_{y \in \Sigma} \tilde{p}_y(t) \lambda^*(y, \Delta^y \tilde{U}) + \epsilon_2(t) \right) + \epsilon_1(t, z), \end{aligned}$$

As  $\|D\tilde{U}(t, z, \tilde{p}(t))\| \leq C$  uniformly and  $\|\epsilon_2(t)\| \leq \frac{1}{N}$ , we know

$$|D\tilde{U}(t, z, \tilde{p}(t)) \cdot \epsilon_2(t)| \leq \frac{C}{N}.$$

Hence, defining  $\epsilon(t, z) := \epsilon_1(t, z) - D\tilde{U}(t, z, \tilde{p}(t)) \cdot \epsilon_2(t)$  concludes the proof. □

Then the DNN approximation  $(\tilde{\theta}, \tilde{p})$  is characterized by (4.18), while the true solution  $(\theta, p)$  of the MFG is characterized by similar one, except  $\epsilon$  and  $\epsilon_3$  are 0. Although the two master equations are now backward PDE, it is still difficult to directly compare their solutions. Hence, we would like to approximate the two PDEs by two ODE systems on some discrete grids of  $P(\Sigma)$ .

Define  $P^N(\Sigma) = \{(\frac{n_1}{N}, \dots, \frac{n_K}{N}), \sum_{z=0}^K n_z = N, n_z \in \mathbb{Z}^+\}$ . Then  $P^N(\Sigma)$  is a discrete grid of  $P(\Sigma)$ . For any  $p^N \in P^N(\Sigma)$ , define operators:

$$\begin{aligned} \alpha^{N,i,j}(p^N) &:= \begin{cases} p^N + \frac{1}{N}(\delta_j - \delta_i) & p_i^N > 0, p_j^N < 1 \\ p^N & \text{else} \end{cases} \\ \Delta^{N,y}\tilde{U}(t, z, p^N) &:= (\tilde{U}(t, z, \alpha^{N,y,1}(p^N)) - \tilde{U}(t, z, p^N), \dots, \tilde{U}(t, z, \alpha^{N,y,K}(p^N)) - \tilde{U}(t, z, p^N)) \\ \Delta^{N,z,z}\tilde{U}(t, z, p^N) &:= (\tilde{U}(t, 1, \alpha^{N,z,1}(p^N)) - \tilde{U}(t, z, p^N), \dots, \tilde{U}(t, K, \alpha^{N,z,K}(p^N)) - \tilde{U}(t, z, p^N)). \end{aligned} \tag{4.20}$$

With the discrete grid and discrete operators defined above, we next show in Proposition 4.9 that the master equation can be approximate by a backward ODE system.

**Proposition 4.9** *There exists  $N_0$  such that for  $N > N_0$ , every  $p^N \in P^N(\Sigma)$  and  $z \in \Sigma$ ,  $\tilde{U}$  solves*

$$\begin{aligned} \frac{\partial \tilde{U}}{\partial t}(t, z, p^N) &= \tilde{\epsilon}^N(t, z, p^N) - H(z, \Delta^{N,z,z}\tilde{U}(t, z, p^N)) \\ &\quad - \sum_{y \in \Sigma} \left( p_y^N - \frac{\mathbb{1}_{y=z}}{N} \right) \lambda^*(y, \Delta^{N,y,y}\tilde{U}(t, y, p^N)) \cdot \Delta^{N,y}\tilde{U}(t, z, p^N) \end{aligned} \tag{4.21}$$

$$\tilde{U}(T, z, p^N) = G(z, p^N) + \epsilon_3(z),$$

where  $\tilde{\epsilon}^N \in C^0([0, T] \times \Sigma \times P^N(\Sigma))$ ,  $\|\tilde{\epsilon}^N\| \leq \frac{C}{N}$ .

**Proof** From Theorem 4.8, there exists constant  $N_0$ , such that when  $N > N_0$  and (4.3),  $\tilde{U}$  satisfies (4.18) when taking value on point  $(t, z, p^N)$ .

$$\frac{\partial \tilde{U}(t, z, p^N)}{\partial t} = -H(z, \Delta^z \tilde{U}) - \sum_{y \in \Sigma} p_y^N \lambda^*(y, \Delta^y \tilde{U}) \cdot D \tilde{U}(t, z, p^N) + \epsilon(t, z).$$

It looks similar to (4.21), except for the discrete operator  $\Delta^{N,y}$  and the differential operator  $D^y$ . Hence, we next compare the two operators similar to [9, Proposition 3]. We first discuss the first component  $\delta_1 - \delta_y$  of  $\Delta^{N,y}\tilde{U}(t, z, p^N)$  defined in (4.20),

$$\begin{aligned} \tilde{U}(t, z, p^N + \frac{1}{N}(\delta_1 - \delta_y)) - \tilde{U}(t, z, p^N) &= \int_0^{\frac{1}{N}} [D^y \tilde{U}(t, z, p^N + s(\delta_1 - \delta_y))]_1 ds \\ &= [D^y \tilde{U}(t, z, p^N)]_1 + \int_0^{\frac{1}{N}} ([D^y \tilde{U}(t, z, p^N + s(\delta_1 - \delta_y))]_1 - [D^y \tilde{U}(t, z, p^N)]_1) ds \\ &= [D^y \tilde{U}(t, z, p^N)]_1 + O\left(\frac{1}{N^2}\right), \end{aligned}$$

where the last equality is derived by the Lipschitz continuity in  $p^N \in P(\Sigma)$  of  $D^y \tilde{U}$ . As above can be applied to every component in  $\Delta^{N,y}\tilde{U}(t, z, p^N)$ , we conclude that there exists  $N_0$  such that for  $N > N_0$ ,

$$\Delta^{N,y}\tilde{U}(t, z, p^N) = D^y \tilde{U}(t, z, p^N) + \epsilon^{N,y}(t, z, p^N),$$



where  $\epsilon^{N,y} \in C^0([0, T] \times \Sigma \times P^N(\Sigma); \mathbb{R}^K)$ ,  $\|\epsilon^{N,y}\| \leq \frac{C}{N^2}$ .

Hence, we have

$$\begin{aligned} \frac{\partial \tilde{U}}{\partial t}(t, z, p^N) &= - \sum_{y \in \Sigma} \left( p_y^N - \frac{\mathbb{1}_{y=z}}{N} \right) \lambda^*(y, \Delta^{N,y,y} \tilde{U}(t, y, p^N)) \cdot \Delta^{N,y} \tilde{U}(t, z, p^N) \\ &\quad - H(z, \Delta^{N,z,z} \tilde{U}(t, z, p^N)) + \sum_{i=1}^4 e_i(t, z), \end{aligned}$$

where

$$\begin{aligned} e_1(t, z) &:= H(z, \Delta^{N,z,z} \tilde{U}(t, z, p^N)) - H(z, \Delta^z \tilde{U}) \\ e_2(t, z) &:= \sum_{y \in \Sigma} p_y^N \Delta^{N,y} \tilde{U}(t, z, p^N) \cdot (\lambda^*(y, \Delta^{N,y,y} \tilde{U}(t, y, p^N)) - \lambda^*(y, \Delta^y \tilde{U})) \\ e_3(t, z) &:= \sum_{y \in \Sigma} p_y^N (\Delta^{N,y} \tilde{U}(t, z, p^N) - D \tilde{U}(t, z, p^N)) \cdot \lambda^*(y, \Delta^y \tilde{U}) \\ e_4(t, z) &:= - \frac{\mathbb{1}_{y=z}}{N} \lambda^*(y, \Delta^{N,y,y} \tilde{U}(t, y, p^N)) \cdot \Delta^{N,y} \tilde{U}(t, z, p^N) + \epsilon(t, z). \end{aligned}$$

From the Lipschitz continuity of  $H$  and  $\lambda^*$ , as well as that  $\tilde{U}$  is bounded, there exists constant  $C$  such that

$$\begin{aligned} |e_1(t, z)| &\leq C \|\Delta^{N,z,z} \tilde{U}(t, z, p^N) - \Delta^z \tilde{U}\| \\ |e_2(t, z)| &\leq C \max_{z \in \Sigma} \|\Delta^{N,z,z} \tilde{U}(t, z, p^N) - \Delta^z \tilde{U}\|. \end{aligned}$$

From Proposition 4.2, we know there exists constant  $C$  such that

$$|e_1(t, z)| + |e_2(t, z)| \leq \frac{C}{2N}.$$

From Lemma 4.4 and  $\sum_{z \in \Sigma} \lambda_z^*(y, \Delta^y \tilde{U}) = 0$  for every  $y \in \Sigma$ , we have

$$D^y \tilde{U}(t, z, p^N) \cdot \lambda^*(y, \Delta^y \tilde{U}) = D \tilde{U}(t, z, p^N) \cdot \lambda^*(y, \Delta^y \tilde{U}).$$

It follows that

$$e_3(t, z) = \sum_{y \in \Sigma} p_y^N (\Delta^{N,y} \tilde{U}(t, z, p^N) - D^y \tilde{U}(t, z, p^N)) \cdot \lambda^*(y, \Delta^y \tilde{U}).$$

From the boundedness of  $\lambda^*$  and  $\epsilon$ , there is constant  $C$  such that

$$\begin{aligned} |e_3(t, z)| &\leq \sum_{y \in \Sigma} p_y^N \left( \sum_{i \in \Sigma} \frac{C}{N^2} \right) \leq \frac{C}{4N} \\ |e_4(t, z)| &\leq \frac{C}{4N}. \end{aligned}$$

We can conclude the proof by defining

$$\tilde{\epsilon}^N(t, z, p^N) := \sum_{i=1}^4 e_i(t, z) < \frac{C}{N}.$$

□

Finally we can proceed to the proof of our main result. The main idea of the proof is to characterize both the DNN approximation  $(\tilde{\theta}, \tilde{p})$  and the true solution  $(\theta, p)$  by their corresponding master equations, which are further approximated by two backward ODE systems on certain discrete grid points. Then the error of the two can be directly estimated on these grid points using Gronwall inequality. As both  $(\tilde{\theta}, \tilde{p})$  and  $(\theta, p)$  are uniformly Lipschitz continuous with respect to their initial conditions, the error between the grid points can also be estimated.

**Completion of proof of Theorem 2.6** Since ODE system (2.2) admits a solution to any initial value  $p_0 \in P(\Sigma)$ , we can define

$$U(t, z, p) := \theta(t, z).$$

From Cecchin and Pelino [9],  $U$  satisfy the master equation for any  $p \in P(\Sigma)$ :

$$\frac{\partial U(t, z, p)}{\partial t} + H(z, \Delta^z U) + \sum_{y \in \Sigma} p_y D\tilde{U}(t, z, p) \cdot \lambda^*(y, \Delta^y U) = 0,$$

$$U(T, z, p) = G(z, p).$$

Similar to the proof of Proposition 4.9, we know  $U(t, z, p^N)$  satisfy ODE:

$$\begin{aligned} \frac{\partial U}{\partial t}(t, z, p^N) &= \epsilon^N(t, z, p^N) - H(z, \Delta^{N,z,z} U(t, z, p^N)) \\ &\quad - \sum_{y \in \Sigma} \left( p_y^N - \frac{\mathbb{1}_{y=z}}{N} \right) \lambda^*(y, \Delta^{N,y,y} U(t, y, p^N)) \cdot \Delta^{N,y} U(t, z, p^N), \\ U(T, z, p^N) &= G(z, p^N), \end{aligned} \tag{4.22}$$

where  $\epsilon^N \in \mathcal{C}^0([0, T] \times \Sigma \times P^N(\Sigma))$ ,  $\|\epsilon^N\| \leq \frac{C}{N}$ . From (4.21) and (4.22), there exists  $N_0$  such that when  $N > N_0$  and (4.3) holds, we have

$$\begin{aligned} \tilde{U}(t, z, p^N) - U(t, z, p^N) &= \epsilon_3(z) + e + A + \sum_{y \in \Sigma} \left( p_y^N - \frac{\mathbb{1}_{y=z}}{N} \right) (B_y + C_y), \\ e &:= \int_t^T (\tilde{\epsilon}^N(s, z, p^N) - \epsilon^N(s, z, p^N)) ds, \\ A &:= \int_t^T (H(z, \Delta^{N,z,z} \tilde{U}(s, z, p^N)) - H(z, \Delta^{N,z,z} U(s, z, p^N))) ds, \\ B_y &:= \int_t^T [\lambda^*(y, \Delta^{N,y,y} \tilde{U}(s, y, p^N)) \\ &\quad - \lambda^*(y, \Delta^{N,y,y} U(s, y, p^N))] \cdot \Delta^{N,y} \tilde{U}(s, z, p^N) ds, \\ C_y &:= \int_t^T \lambda^*(y, \Delta^{N,y,y} U(s, y, p^N)) \\ &\quad \cdot [\Delta^{N,y} \tilde{U}(s, z, p^N) - \Delta^{N,y} U(s, z, p^N)] ds. \end{aligned}$$

From Proposition 4.1, both  $U$  and  $\tilde{U}$  are bounded. Hence,  $H$  and  $\lambda^*$  are Lipschitz continuous w.r.t their second variable. Define

$$d(t) := \max_{z \in \Sigma, p^N \in P^N(\Sigma)} |\tilde{U}(t, z, p^N) - U(t, z, p^N)|.$$

There exists a constant  $C$  such that

$$|A| + |B_y| + |C_y| \leq C \int_t^T d(s)ds.$$

As  $p_N \in P^N(\Sigma)$ , there exists constant  $C$  such that

$$d(t) \leq \max_{z \in \Sigma, p^N \in P^N(\Sigma)} \left\{ \int_t^T |\tilde{\epsilon}^N(s, z, p^N) - \epsilon^N(s, z, p^N)|ds + \epsilon_3(z) \right\} + C \int_t^T d(s)ds \leq \frac{C}{N} + C \int_t^T d(s)ds.$$

By applying Gronwall inequality, there is constant  $C$  such that for every  $t \in [0, T]$ ,  $z \in \Sigma$  and  $p^N \in P^N(\Sigma)$  we have

$$|\tilde{U}(t, z, p^N) - U(t, z, p^N)| \leq \frac{C}{N}. \tag{4.23}$$

For  $N > 2N_0$  where  $N_0$  is defined in Proposition 4.9 above, if  $\tilde{p} \in B(P(\Sigma), \frac{1}{N})$ , there is  $p \in P(\Sigma)$  such that  $\tilde{p} = p + \epsilon_4$  and  $\epsilon_4 < \frac{1}{N}$ . And there exists  $p^N \in P^N(\Sigma)$  such that

$$\begin{aligned} \|p - p^N\| &< \frac{1}{N} \\ \|\tilde{p} - p^N\| &\leq \|\tilde{p} - p\| + \|p - p^N\| < \frac{2}{N} < \frac{1}{N_0}. \end{aligned}$$

From Proposition 4.1,  $\tilde{U}(t, z, \tilde{p})$  is well defined, and from Proposition 4.2, there exists constant  $C$  independent of  $N$  and  $p$ , such that for every  $t \in [0, T]$  and  $z \in \Sigma$ ,

$$|U(t, z, p) - U(t, z, p^N)| \leq \frac{C}{N}, \quad |\tilde{U}(t, z, \tilde{p}) - \tilde{U}(t, z, p^N)| \leq \frac{2C}{N}.$$

Combining the above inequalities with (4.23), we have  $|\tilde{U}(t, z, \tilde{p}) - U(t, z, p)| \leq \frac{C}{N}$  for some constant  $C$  independent to  $N$  and  $p$ , which is equivalent to

$$\|\tilde{\theta} - \theta\| \leq \frac{C}{N}.$$

By using the uniform boundedness and Lipschitz continuity of  $\lambda^*$ , we can prove  $p$  and  $\tilde{p}$  are Lipschitz continuous w.r.t  $\theta$  and  $\tilde{\theta}$ , respectively, with the help of Gronwall inequality and technique similar to the proof of Proposition 4.2. Note also that the Lipschitz coefficient only depends on the uniform bound and Lipschitz continuous coefficient of  $\lambda^*$ , which again only depend on the preliminary  $M$  given in Proposition 4.1. Hence, we know there exists a uniform constant  $C$  independent on  $N$  such that

$$\|\tilde{p} - p\| \leq \frac{C}{N}.$$

This concludes the proof. □

### 4.4 Proof of Proposition 3.1

**Proof** The proof is divided to several steps to prove the conditions for  $H$  and  $G$ , respectively.

Step 1: proof of  $\lambda^*$  for Assumption 2.1.

Let us first write out the Hamilton operator  $H$  for  $G_{c,R}$ . Define  $\mathcal{A}$  as the admissible control set for all  $\lambda$  that satisfy (3.4). Define  $\delta^a := \Lambda^{-1}(\lambda_{\beta^a(z)}(t, z))$  and  $\delta^b := \Lambda^{-1}(\lambda_{\beta^b(z)}(t, z))$ , then we have

$$\begin{aligned}
 H(z, \mu) &= \sup_{\lambda \in \mathcal{A}} \left\{ g(\Lambda^{-1}(\lambda_{\beta^a(z)}(t, z)), \mu_{\beta^a(z)}) + g(\Lambda^{-1}(\lambda_{\beta^b(z)}(t, z)), \mu_{\beta^b(z)}) - \frac{1}{2} \gamma \sigma^2 Z_1^{-1}(z)^2 \right\} \\
 &= \sup_{\delta^a \in \mathbb{R}} \{g(\delta^a, \mu_{\beta^a(z)})\} + \sup_{\delta^b \in \mathbb{R}} \{g(\delta^b, \mu_{\beta^b(z)})\} - \frac{1}{2} \gamma \sigma^2 Z_1^{-1}(z)^2,
 \end{aligned}$$

where  $g(\delta, \mu) := \Lambda(\delta)(\delta - c + \mu)$ . From (3.1) and according to the proof of Lemma 3.1 in Guéant [15],  $\zeta(\mu) := \sup_{\delta} \{g(\delta, \mu)\}$  is increasing w.r.t  $\mu$ . Moreover, the optimal  $\delta^*$  exists and is unique, which is a continuously differentiable function of  $\mu$ .

Step 2: proof of  $H$  satisfying Assumption 2.1.

We only need to prove that the second-order derivative  $\zeta''(\mu)$  is positive. From the proof of Lemma 3.1 in Guéant [15],  $\zeta$  is  $\mathcal{C}^2$ ,  $\zeta'(\mu) = \Lambda(\delta^*)$ , and  $\delta^*$  is strictly decreasing w.r.t  $\mu$ . Hence,  $\Lambda(\delta^*)$  is strictly increasing w.r.t  $\mu$ , which implies  $\zeta''(\mu) > 0$ . Then there exists constant  $C$  such that  $\zeta''(\mu) > C$  when  $\mu$  is bounded.

Step 3: proof of  $G$  satisfying Assumption 2.1.

From (3.5), the differentiability and (2.5) of  $G$  are trivial. We then only need to prove (2.6). Note that

$$\begin{aligned}
 &\sum_{z \in \Sigma} (G(z, p_z) - G(z, \bar{p}_z))(p_z - \bar{p}_z) \\
 &= \sum_{v=0}^{v_{\max}} \sum_{q=-Q}^Q \sum_{i=v}^{v_{\max}} (\bar{p}(T, i) - p(T, i))(p(T, q, v) - \bar{p}(T, q, v))R \\
 &= \sum_{v=0}^{v_{\max}} \sum_{i=v}^{v_{\max}} (\bar{p}(T, i) - p(T, i))(p(T, v) - \bar{p}(T, v))R \\
 &= -\frac{R}{2} \left( \left( \sum_{v=0}^{v_{\max}} (p(T, v) - \bar{p}(T, v)) \right)^2 + \sum_{v=0}^{v_{\max}} (p(T, v) - \bar{p}(T, v))^2 \right),
 \end{aligned}$$

which is non-positive, this concludes the proof. □

### 5 Conclusions

In this paper, we have solved the finite state mean field games problem by the deep neural network method. By transforming the fully coupled FBODE system to the master equation, we have proved that the error between the true solution and the approximate solution is linear to the square root of DNN loss function. We have also applied the DNN method to solve the optimal market making problem with terminal rank-based trading volume reward which is shown to perform better in liquidity provision and trading cost reduction than the linear trading volume reward. There remain many open questions such as general heterogeneous interaction structure and infinite state MFG. We leave these and other questions for future research.

**Acknowledgements** The authors are grateful to the anonymous reviewers whose constructive comments and suggestions have helped improve the paper of the previous version.

**Funding** This study was supported in part by EPSRC (UK) Grant (EP/V008331/1).

**Availability of data and material** Not applicable.

**Code Availability** Not available

## Declarations

**Conflict of interest** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Achdou Y, Camilli F, Capuzzo-Dolcetta I (2012) Mean field games: numerical methods for the planning problem. *SIAM J Control Optim* 50:77–109
2. Achdou Y, Capuzzo-Dolcetta I (2010) Mean field games: numerical methods. *SIAM J Numer Anal* 48:1136–1162
3. Cardaliaguet P, Delarue F, Lasry JM, Lions PL (2015) The master equation and the convergence problem in mean field games. [arXiv:1509.02505](https://arxiv.org/abs/1509.02505)
4. Carmona R, Delarue F (2013) Probabilistic analysis of mean-field games. *SIAM J Control Optim* 51:2705–2734
5. Carmona R, Laurière M (2021) Deep learning for mean field games and mean field control with applications to finance. [arXiv:2107.04568](https://arxiv.org/abs/2107.04568)
6. Carmona R, Wang P (2016) Finite state mean field games with major and minor players. [arXiv:1610.05408](https://arxiv.org/abs/1610.05408)
7. Carmona R, Wang P (2018) A probabilistic approach to extended finite state mean field games. [arXiv:1808.07635](https://arxiv.org/abs/1808.07635)
8. Cecchin A, Fischer M (2020) Probabilistic approach to finite state mean field games. *Appl Math Optim* 81:253–300
9. Cecchin A, Pelino G (2019) Convergence, fluctuations and large deviations for finite state mean field games via the master equation. *Stochast Process Appl* 129:4510–4555
10. El Euch O, Mastrolia T, Rosenbaum M, Touzi N (2018) Optimal make-take fees for market making regulation. SSRN 3174933
11. Fouque JP, Zhang Z (2020) Deep learning methods for mean field control problems with delay. *Front Appl Math Stat* 6:11
12. Gomes D, Mohr J, Souza R (2013) Continuous time finite state mean field games. *Appl Math Optim* 68:99–143
13. Gomes D, Saude J (2017) Monotone numerical methods for finite-state mean-field games. [arXiv:1705.00174](https://arxiv.org/abs/1705.00174)
14. Guéant O (2009) A reference case for mean field games models. *J mathématiques pures et appliquées* 92:276–294
15. Guéant O (2017) Optimal market making. *Appl Math Finance* 24:112–154
16. Guéant O, Lasry JM, Lions PL (2011) Mean field games and applications. In: Paris-Princeton lectures on mathematical finance 2010. Springer, pp 205–266
17. Han J, Jentzen A (2018) Solving high-dimensional partial differential equations using deep learning. *Proc Natl Acad Sci* 115:8505–8510
18. Huang M, Malhamé R, Caines P (2006) Large population stochastic dynamic games: closed-loop Mckean-Vlasov systems and the nash certainty equivalence principle. *Commun Inf Syst* 6:221–252
19. Lagaris I, Likas A, Fotiadis D (1998) Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans Neural Netw* 9:987–1000

20. Lagaris I, Likas A, Papageorgiou D (2000) Neural-network methods for boundary value problems with irregular boundaries. *IEEE Trans Neural Netw* 11:1041–1049
21. Lasry JM, Lions PL (2007) Mean field games. *Japan J Math* 2:229–260
22. Lasry JM, Lions PL, Guéant O (2008) Application of mean field games to growth theory. hal:00348376
23. Lauriere M (2021) Numerical methods for mean field games and mean field type control. [arXiv:2106.06231](https://arxiv.org/abs/2106.06231)
24. Lee H, Kang I (1990) Neural algorithm for solving differential equations. *J Comput Phys* 91:110–131
25. Li J, Yue J, Zhang W, Duan W (2020) The deep learning Galerkin method for the general stokes equations. [arXiv:2009.11701](https://arxiv.org/abs/2009.11701)
26. Li J, Zhang W, Yue J (2021) A deep learning Galerkin method for the second-order linear elliptic equations. *Int J Numer Anal Model* 18:427–441
27. Malek A, Beidokhti R (2006) Numerical solution for high order differential equations using a hybrid neural network-optimization method. *Appl Math Comput* 183:260–271
28. Mishra S, Molinaro R (2021) Estimates on the generalization error of physics-informed neural networks for approximating pdes. [arXiv:2006.16144](https://arxiv.org/abs/2006.16144)
29. Ruthotto L, Osher SJ, Li W, Nurbekyan L, Fung SW (2020) A machine learning framework for solving high-dimensional mean field game and mean field control problems. *Proc Natl Acad Sci* 117:9183–9193
30. Sirignano J, Spiliopoulos K (2018) DGM: a deep learning algorithm for solving partial differential equations. *J Comput Phys* 375:1339–1364

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.