A. Preliminaries on Rough Path Theory and Signatures

In this appendix, we shall follow Lyons & Qian (2002); Lyons et al. (2007); Friz & Victoir (2010) and briefly introduce rough path theory and signatures. We will also give the proof of Lemma 4.1 using the factorial decay property of signatures. Denote by Δ_T the simplex $\{(s,t) \in [0,T]^2 : 0 \leq s \leq t \leq T\}$, and by $T^n(\mathbb{R}^d) = \bigoplus_{k=0}^n (\mathbb{R}^d)^{\bigotimes k}$ the truncated tensor algebra.

Definition A.1 (Multiplicative Functional). *Let $\mathbb{X} : \Delta_T \to T^n(\mathbb{R}^d)$, with $n \geq 1$ as an integer. For each $(s,t) \in \Delta_T$, $\mathbb{X}_{s,t}$ denotes the image of (s,t) under the mapping $\mathbb{X}$, and we write*

$$\mathbb{X}_{s,t} = (\mathbb{X}^0_{s,t}, \mathbb{X}^1_{s,t}, \ldots, \mathbb{X}^n_{s,t}) \in T^n(\mathbb{R}^d).$$

The function $\mathbb{X}$ is called a multiplicative functional of degree n in $\mathbb{R}^d$ if $\mathbb{X}^0_{s,t} = 1$ for all $(s,t) \in \Delta_T$ and

$$\mathbb{X}_{s,u} \otimes \mathbb{X}_{u,t} = \mathbb{X}_{s,t}, \ \forall s, u, t \in [0,T], \ s \leq u \leq t, \tag{A.1}$$

which is called Chen's identity.

Rough paths will be defined as a multiplicative functional with extra regularization conditions.

Definition A.2 (Control). *A control function on $[0,T]$ is a continuous non-negative function ω on the simplex Δ_T which is supper-additive in the sense that*

$$\omega(s,u) + \omega(u,t) \leq \omega(s,t) \ \ \forall s \leq u \leq t \in [0,T].$$

It is easy to see that $\omega(t,t) = 0$ for any control ω. In the following, we use the notation $x! = \Gamma(x+1)$, where $\Gamma(\cdot)$ is the Gamma function and x is a positive real number.

Definition A.3. *Let $p \geq 1$ be a real number and $n \geq 1$ be an integer. Denote $\omega : \Delta_T \to [0, +\infty)$ as a control and $\mathbb{X} : \Delta_T \to T^n(\mathbb{R}^d)$ as a multiplicative functional. Then we say that $\mathbb{X}$ has finite p-variation on Δ_T controlled by ω if*

$$\|\mathbb{X}^i_{s,t}\| \leq \frac{\omega(s,t)^{\frac{i}{p}}}{\beta(\frac{i}{p})!} \ \ \forall i = 1, \ldots, n, \ \ \forall (s,t) \in \Delta_T, \tag{A.2}$$

where $\|\cdot\|$ is the tensor norm induced by the norm on $\mathbb{R}^d$. We will call that $\mathbb{X}$ has finite p-variation in short if there exists a control ω such that (A.2) is satisfied.

Note that in (A.2), β is a constant depending only on p. We are now ready to define the rough paths.

Definition A.4 (Rough Path). *Let $p \geq 1$ be a real number. A p-rough path in $\mathbb{R}^d$ is a multiplicative functional of degree $\lfloor p \rfloor$ with finite p-variation. The space of p-rough paths is denoted by $\Omega_p(\mathbb{R}^d)$.*

Given a continuous path $X : [0,T] \to \mathbb{R}^d$ with bounded p-variation, one can construct a $\lfloor p \rfloor$-rough path $\mathbb{X}$ with $\mathbb{X}^1_{s,t} = X_t - X_s$ for any $s \leq t$. In particular, truncated siganture $S^{\lfloor p \rfloor}(X) \in T^{\lfloor p \rfloor}(\mathbb{R}^d)$ is a p-rough path. The following fundamental theorem of rough paths allows us to make extension of a p-rough path,

Theorem A.1 (Extension Theorem, Lyons & Qian (2002)). *Let $p \geq 1$ be a real number and $n \geq 1$ an integer. Denote $\mathbb{X} : \Delta_T \to T^n(\mathbb{R}^d)$ as a multiplicative functional with finite p-variation controlled be a control ω. Assume that $n \geq \lfloor p \rfloor$, then there exists a unique extension of $\mathbb{X}$ to a multiplicative functional $\Delta_T \to T((\mathbb{R}^d))$ which possesses finite p-variation.*

More precisely, for every $m \geq \lfloor p \rfloor + 1$, there exists a unique continuous function $\mathbb{X}^m : \Delta_T \to (\mathbb{R}^d)^{\bigotimes m}$ such that

$$(s,t) \to \mathbb{X}_{s,t} = \left(1, \mathbb{X}^1_{s,t}, \ldots, \mathbb{X}^{\lfloor p \rfloor}_{s,t}, \ldots, \mathbb{X}^m_{s,t}, \ldots\right) \in T((\mathbb{R}^d))$$

is a multiplicative functional with finite p-variation controlled by ω. By this we mean that

$$\|\mathbb{X}^i_{s,t}\| \leq \frac{\omega(s,t)^{\frac{i}{p}}}{\beta(\frac{i}{p})!} \ \ \forall i \geq 1, \ \ \forall (s,t) \in \Delta_T. \tag{A.3}$$

Signature can be seen as an extension of rough path, and its factorial decay property follows by (A.3). The control function is related to p-variation of path. Given that $x \in \mathcal{V}^p([0,T],\mathbb{R}^d)$, $S^{\lfloor p \rfloor}(x)$ is a p-rough path and one candidate for its control function is

$$\omega(s,t) = \sum_{i=1}^{\lfloor p \rfloor} \sup_{D\subset[s,t]} \sum_k \|x^i_{t_{k+1}} - x^i_{t_k}\|^{p/i}, \tag{A.4}$$

where the norm is the tensor norm induced by Euclidean norm in $\mathbb{R}^d$.

Let $S^{\lfloor p \rfloor}(\Omega_1) = \{S^{\lfloor p \rfloor}(x) : x \in \Omega_1(\mathbb{R}^d)\}$, and $\mathbb{Y}$ be a p-rough path. We call $\mathbb{Y}$ a p-geometric rough path if $\mathbb{Y}$ is in the closure of $S^{\lfloor p \rfloor}(\Omega_1)$ under p-variation metric, where p-variation metric is given by

$$d_{p\text{-var}}(\mathbb{X},\mathbb{Y}) := \left(\sup_D \sum_{t_i\in D} \|\mathbb{X}_{t_i,t_{i+1}} - \mathbb{Y}_{t_i,t_{i+1}}\|^p\right)^{1/p}, \qquad \mathbb{X},\mathbb{Y} \in \Omega_p(\mathbb{R}^d). \tag{A.5}$$

Proof of Lemma 4.1. By constructing the iterated integral in Stratonovich sense, $S(\hat{B}_{0:T})$ is the signature of a p-geometric rough path $\forall p \in (2,3)$ (Friz & Victoir, 2010), and thus it characterizes $B_{0:T}$ uniquely. Therefore, conditional distribution $\mu_t = \mathbb{E}[\iota(X_t)|\mathcal{F}^B_t]$ can be written as $\mu_t := \mu(t, B_{0,t}) = \mu(\hat{B}_{0,t})$.

By Theorem 3.1, for any $\epsilon > 0$ there exits l such that

$$\sup_{\hat{B}\in K} |\mu(\hat{B}_{0:T}) - \langle l, S(\hat{B}_{0:T})\rangle| < \frac{\epsilon}{2}. \tag{A.6}$$

Since $|\langle l, S(\hat{B}_{0:T}) - S^M(\hat{B}_{0:T})\rangle| \le \|l\| \cdot \|S(\hat{B}_{0:T}) - S^M(\hat{B}_{0:T})\|$ where the first norm is functional norm and second is tensor norm and $\|S(\hat{B}_{0:T}) - S^M(\hat{B}_{0:T})\| = \sum_{i\ge M+1} \|\hat{B}^i_{0:T}\|$. By the compactness of K, and (A.3), (A.4), $\sum_{i\ge M+1} \|\hat{B}^i_{0:T}\|$ admits a convergent uniform norm over $\hat{B} \in K$ and goes to 0 as $M \to \infty$. Then for M large enough,

$$\sup_{\hat{B}\in K} |\mu(\hat{B}_{0:T}) - \langle l, S^M(\hat{B}_{0:T})\rangle| < \frac{\epsilon}{2} + \sup_{\hat{B}\in K} |\langle l, S(\hat{B}_{0:T}) - S^M(\hat{B}_{0:T})\rangle| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \tag{A.7}$$

For $t < T$, we extend path $\hat{B}_{0:t}$ to space $\mathcal{V}^p([0,T],\mathbb{R}^d)$ by defining

$$\tilde{B}^t_s := \begin{cases} \hat{B}_s, & 0 \le s \le t \\ \hat{B}_t, & t < s \le T. \end{cases}$$

Then $\tilde{B}^t_{0:T} \in \mathcal{V}^p([0,T],\mathbb{R}^d)$, $S(\tilde{B}^t_{0:T}) = S(\hat{B}_{0:t})$ by Chen's identity (A.1), and $\mu(\hat{B}_{0:t}) = \mu(\tilde{B}^t_{0,T})$. Denote $\tilde{K} = \{\tilde{B}^t_{0:T}, \forall t \in [0,T] : \tilde{B}^t_{0:T}$ is constructed by $\hat{B}_{0:t}$ and $\hat{B} \in K\}$. Thus $\tilde{K}$ is also compact.

$$\begin{aligned} \sup_{t\in[0,T]} \sup_{\hat{B}\in K} |\mu(\hat{B}_{0:t}) - \langle l, S^M(\hat{B}_{0:t})\rangle| &= \sup_{t\in[0,T]} \sup_{\hat{B}\in K} |\mu(\tilde{B}^t_{0:T}) - \langle l, S^M(\tilde{B}^t_{0:T})\rangle| \\ &= \sup_{\tilde{B}\in\tilde{K}} |\mu(\tilde{B}_{0:T}) - \langle l, S^M(\tilde{B}_{0:T})\rangle| < \epsilon, \end{aligned} \tag{A.8}$$

where the second equality is due to the construction of $\tilde{B}^t_{0:T}$ and the last inequality is by (A.7). □

B. Details of Implementing the Sig-DFP Algorithm

The simulation of $X^{i,(n)}$ and $J_B(\varphi, \hat{\mu}^{(n-1)})$ follows

$$J_B(\varphi, \hat{\mu}^{(n-1)}) = \frac{1}{B}\sum_{i=1}^B \left(\sum_{k=0}^{L-1} f(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i), \alpha_\varphi(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i))\Delta_k + g(X_L, \hat{\mu}^{(n-1)}_L(\omega^i))\right), \tag{B.1}$$

$$\begin{aligned} X^{i,(n)}_{k+1} = X^{i,(n)}_k &+ b(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i), \alpha_\varphi(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i))\Delta_k \\ &+ \sigma(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i), \alpha_\varphi(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i))\Delta W^i_k \\ &+ \sigma^0(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i), \alpha_\varphi(t_k, X^{i,(n)}_k, \hat{\mu}^{(n-1)}_k(\omega^i))\Delta B^i_k, \quad X^{i,(n)}_0 = X^i_0 \sim \mu_0, \end{aligned} \tag{B.2}$$

where $\hat{\mu}_k^{(n-1)}(\omega^i)$ is computed by $\hat{\mu}_k^{(n-1)}(\omega^i) = \langle \bar{l}^{(n-1)}, S^M(\hat{B}^i_{0:t_k})\rangle$ with $\bar{l}^{(n-1)}$ obtained from the previous round of fictitious play. Then $l^{(n)}$ is calculated by regressing $\{\iota(X_0^{i,(n)}), \iota(X_{L/2}^{i,(n)}), \iota(X_L^{i,(n)})\}_{i=1}^N$ on $\{S^M(\hat{B}_{0,0}), S^M(\hat{B}_{0,t_{L/2}}), S^M(\hat{B}_{0,t_L})\}_{i=1}^N$, and we update $\bar{l}^{(n)} = \frac{n-1}{n}\bar{l}^{(n-1)} + \frac{1}{n}l^{(n)}$ for $n \geq 1$. The algorithm starts with a random initialization $\bar{l}^{(0)}$ to produce $\hat{\mu}^{(0)}$.

Linear-Quadratic MFGs. We set α_φ to be a feed-forward NN with two hidden layers of width 64. The signature depth is chosen at $M = 2$. This model is trained for $N_{round} = 500$ iterations of fictitious play. Note that fictitious play has a slow convergence speed since our initial guess $m^{(0)}$ is far from the truth. Therefore, we only apply averaging over distributions (or linear functions) during the second half iteration. We set the learning rate as 0.1 for the first half iterations and 0.01 for the second half. The minibatch size is $B = 2^{10}$, and hence $N_{batch} = 2^5$.

Mean-field Portfolio Game. We consider signature depth $M = 2$ and use a fully connected neural network π_φ with four hidden layers to estimate π_t. Since different players are characterized by their type vectors ζ, π_φ takes (ζ, t, X_t, m_t) as inputs. Hidden neurons in each layer are (64, 32, 32, 16). We train our model with $N_{round} = 500$ rounds fictitious play. The learning rate starts at 0.1 and is reduced by a factor of 5 after every 200 rounds. The minibatch size is $B = 2^{10}$, and hence $N_{batch} = 2^5$.

Mean-field Game of Optimal Consumption and Investment. In this example, signature depth is $M = 4$. The optimal controls $(\pi_t, c_t)_{0\leq t\leq 1}$ are estimated by two neural networks π_φ and c_φ, each with three hidden layers. Due the nature of heterogeneous extended MFG, both α_φ and c_φ take $(\zeta_t, t, X_t, m_t, \Gamma_t)$ as the inputs. Hidden layers in each network have width (64, 64, 64). We will propagate two conditional distribution flows, *i.e.*, two linear functionals $\bar{l}^{(n)}, \bar{l}_c^{(n)}$ during each round fictitious play. Instead of estimating m_t, Γ_t directly, we estimate $\mathbb{E}[\log X_t^* | \mathcal{F}_t^B], \mathbb{E}[\log c_t^* | \mathcal{F}_t^B]$ by $\langle \bar{l}^{(n)}, S^4(\hat{B}_{0:t})\rangle$, $\langle \bar{l}_c^{(n)}, S^4(\hat{B}_{0:t})\rangle$, and then take the exponential to get m_t, Γ_t. To ensure the non-negativity condition, we evolve $\log X_t$ according to (D.4), use c_φ to predicted $\log c_t$, and then take exponential to get c_t, X_t. We use $N_{round} = 600$ rounds fictitious play training, learning rate 0.1 decaying by a factor of 5 for every 200 rounds, the minibatch size $B = 2^{11}$, and hence $N_{batch} = 2^4$.

The training time for all three experiments with sample size $N = 2^{13}, 2^{14}, 2^{15}$ is given in Table 7.

Table 7. Training time in minutes. Here LQ-MFG = Linear-Quadratic mean-field games, MF Portfolio = Mean-field Portfolio Game, and MFG with Consump. = Mean-field Game of Optimal Consumption and Investment.

	$N = 2^{13}$	$N = 2^{14}$	$N = 2^{15}$
LQ-MFG	12.4	23.7	46.7
MF Portfolio	12.3	23.3	45.5
MFG with Consump.	23.4	40.9	80.1

C. Proof of Theorems 4.1 and 4.2

We first list all main assumptions on $(b, \sigma, \sigma^0, f, g)$ that will be used to prove Theorem 4.1. Let $\|\cdot\|$ be the Euclidean norm and K be the same constant for all assumptions below.

Assumption C.1. *We make assumptions* **A1-A3** *and* **B1-B3** *as follows.*

A1. *(Lipschitz) $\partial_x f, \partial_\alpha f, \partial_x g$ exist and are K-Lipschitz continuous in (x, α) uniformly in (t, μ), i.e., for any $t \in [0, T]$, $x, x' \in \mathbb{R}^d, \alpha, \alpha' \in \mathbb{R}^m, \mu \in \mathcal{P}^2(\mathbb{R}^d)$,*

$$\|\partial_x g(x, \mu) - \partial_x g(x', \mu)\| \leq K\|x - x'\|,$$
$$\|\partial_x f(t, x, \mu, \alpha) - \partial_x f(t, x', \mu, \alpha')\| \leq K(\|x - x'\| + \|\alpha - \alpha'\|),$$
$$\|\partial_\alpha f(t, x, \mu, \alpha) - \partial_\alpha f(t, x', \mu, \alpha')\| \leq K(\|x - x'\| + \|\alpha - \alpha'\|).$$

The drift coefficient $b(t, x, \mu, \alpha)$ in (2.3) takes the form

$$b(t, x, \mu, \alpha) = b_0(t, \mu) + b_1(t)x + b_2(t)\alpha,$$

where $b_0 \in \mathbb{R}^d$, $b_1 \in \mathbb{R}^{d\times d}$ and $b_2 \in \mathbb{R}^{d\times m}$ are measurable functions and bounded by K. The diffusion coefficients $\sigma(t, x, \mu)$ and $\sigma^0(t, x, \mu)$ are uncontrolled and K-Lipschitz in x uniformly in (t, μ):

$$\|\sigma(t, x, \mu)\| \leq K\|x - x'\|, \quad \|\sigma^0(t, x, \mu)\| \leq K\|x - x'\|.$$

A2. *(Growth)* $\partial_x f, \partial_\alpha f, \partial_x g$ *satisfy a linear growth condition, i.e., for any* $t \in [0,T]$, $x \in \mathbb{R}^d, \alpha \in \mathbb{R}^m, \mu \in \mathcal{P}^2(\mathbb{R}^d)$,

$$\|\partial_x g(x,\mu)\| \leq K\bigg(1 + \|x\| + \Big(\int_{\mathbb{R}^d} \|y\|^2 \,\mathrm{d}\mu(y)\Big)^{\frac{1}{2}}\bigg),$$
$$\|\partial_x f(t,x,\mu,\alpha)\| \leq K\bigg(1 + \|x\| + \|\alpha\| + \Big(\int_{\mathbb{R}^d} \|y\|^2 \,\mathrm{d}\mu(y)\Big)^{\frac{1}{2}}\bigg),$$
$$\|\partial_\alpha f(t,x,\mu,\alpha)\| \leq K\bigg(1 + \|x\| + \|\alpha\| + \Big(\int_{\mathbb{R}^d} \|y\|^2 \,\mathrm{d}\mu(y)\Big)^{\frac{1}{2}}\bigg).$$

In addition f, g *satisfy a quadratic growth condition in* μ*:*

$$|g(0,\mu)| \leq K\bigg(1 + \int_{\mathbb{R}^d} \|y\|^2 \,\mathrm{d}\mu(y)\bigg),$$
$$|f(t,0,\mu,0)| \leq K\bigg(1 + \int_{\mathbb{R}^d} \|y\|^2 \,\mathrm{d}\mu(y)\bigg).$$

A3. *(Convexity)* g *is convex in* x *and* f *is convex jointly in* (x,α) *with strict convexity in* α*, i.e., for any* $x, x' \in \mathbb{R}^d, \mu \in \mathcal{P}^2(\mathbb{R}^d)$,

$$(\partial_x g(x,\mu) - \partial_x g(x',\mu))^T (x - x') \geq 0,$$

and there exist a constant $c_f > 0$ *such that for any* $t \in [0,T]$, $x, x' \in \mathbb{R}^d, \alpha, \alpha' \in \mathbb{R}^m$, $\mu \in \mathcal{P}^2(\mathbb{R}^d)$,

$$f(t,x',\alpha',\mu) \geq f(t,x,\alpha,\mu) + \partial_x f(t,x,\alpha,\mu)^T (x'-x) + \partial_\alpha f(t,x,\alpha,\mu)^T(\alpha'-\alpha) + c_f\|\alpha'-\alpha\|^2.$$

B1. *(Lipschitz in* μ*)* $\partial_x g, \partial_x f, \partial_\alpha f, b_0, \sigma, \sigma^0$ *are Lipschitz continuous in* μ *uniformly in* (t,x)*, i.e., there exists a constant* K *such that*

$$\begin{aligned}
\|\partial_x g(x,\mu) - \partial_x g(x,\mu')\| &\leq K\mathcal{W}_2(\mu,\mu'),\\
\|\partial_x f(t,x,\mu,\alpha) - \partial_x f(t,x,\mu',\alpha)\| &\leq K\mathcal{W}_2(\mu,\mu')\\
\|\partial_\alpha f(t,x,\mu,\alpha) - \partial_\alpha f(t,x,\mu',\alpha)\| &\leq K\mathcal{W}_2(\mu,\mu')\\
\|b_0(t,\mu) - b_0(t,\mu')\| &\leq K\mathcal{W}_2(\mu,\mu'),\\
\|\sigma(t,x,\mu) - \sigma(t,x,\mu')\| &\leq K\mathcal{W}_2(\mu,\mu'),\\
\|\sigma^0(t,x,\mu) - \sigma^0(t,x,\mu')\| &\leq K\mathcal{W}_2(\mu,\mu'),
\end{aligned}$$

for all $t \in [0,T], x \in \mathbb{R}^d, \alpha \in \mathbb{R}^m$, $\mu, \mu' \in \mathcal{P}^2(\mathbb{R}^d)$*, where* $\mathcal{W}_2$ *is the 2-Wasserstein distance.*

B2. *(Separable in* α, μ*)* f *is of the form*

$$f(t,x,\mu,\alpha) = f^0(t,x,\alpha) + f^1(t,x,\mu),$$

where f^0 *is assumed to be convex in* (x,α) *and strictly convex in* α*, and* f^1 *is assumed to be convex in* x*.*

B3. *(Weak monotonicity) For all* $t \in [0,T]$, $\mu, \mu' \in \mathcal{P}^2(\mathbb{R}^d)$ *and* $\gamma \in \mathcal{P}^2(\mathbb{R}^d \times \mathbb{R}^d)$ *with marginals* μ, μ' *respectively,*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \big[(\partial_x g(x,\mu) - \partial_x g(y,\mu'))^T (x-y)\big]\gamma(\,\mathrm{d}x,\,\mathrm{d}y) \geq 0,$$
$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \big[(\partial_x f(t,x,\mu,\alpha) - \partial_x g(t,y,\mu',\alpha))^T (x-y)\big]\gamma(\,\mathrm{d}x,\,\mathrm{d}y) \geq 0.$$

Note that Assumption C.1 extends conditions **A** and **B** in Ahuja (2015) by considering general drift coefficient $b(t,x,\mu,\alpha)$ and non-constant diffusion coefficients $\sigma(t,x,\mu)$ and $\sigma^0(t,x,\mu)$.

Our proof of Theorem 4.1 uses the probabilistic approach. To this end, we define the Hamiltonian by

$$H(t,x,y,\mu,\alpha) = b(t,x,\mu,\alpha)\cdot y + f(t,x,\mu,\alpha).$$

Denote by $\hat{\alpha}$ the minimizer of the Hamiltonian which is unique due to Assumptions **A1** and **A3**:

$$\hat{\alpha}(t,x,y,\mu) = \arg\min_{\alpha\in\mathbb{R}^m} H(t,x,y,\mu,\alpha). \tag{C.1}$$

By the Lipschitz property of $\partial_\alpha f$ in (t,μ,α) and the boundedness of $b_2(t)$, $\hat{\alpha}$ is Lipschitz in (x,y,μ). Let $\hat{H}$ be the Hamiltonian, with $\hat{\alpha}$ obtained in (C.1),

$$\hat{H}(t,x,y,\mu) = H(t,x,y,\mu,\hat{\alpha}(t,x,y,\mu)). \tag{C.2}$$

Under Assumptions **A1-A3**, with the stochastic maximum principle, the problem (2.2)-(2.3) is equivalent to solve the following FBSDE, given $\mu \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$,

$$\begin{aligned} \mathrm{d}X_t &= b(t,X_t,\mu_t,\hat{\alpha}(t,X_t,Y_t,\mu_t))\,\mathrm{d}t + \sigma(t,X_t,\mu_t)\,\mathrm{d}W_t + \sigma^0(t,X_t,\mu_t)\,\mathrm{d}B_t, \quad X_0 = x_0 \sim \mu_0, \\ \mathrm{d}Y_t &= -\partial_x \hat{H}(t,X_t,Y_t,\mu_t)\,\mathrm{d}t + Z_t\,\mathrm{d}W_t + Z_t^0\,\mathrm{d}B_t, \quad Y_T = \partial_x g(X_T,\mu_T). \end{aligned} \tag{C.3}$$

Moreover, the optimal control is given by

$$\hat{\alpha}_t = \hat{\alpha}(t,X_t,Y_t,\mu_t), \tag{C.4}$$

for any solution (X_t,Y_t,Z_t,Z_t^0) to FBSDE (C.3).

The next theorem describes the McKean-Vlasov FBSDE for finding the mean-field equilibrium (*cf.* Definition 2.1).

Theorem C.1 (Theorem 2.2.8, Ahuja (2015))**.** *Under Assumptions* **A1-A3***, the mean-field equilibrium of* (2.2)-(2.3) *exists if and only if the following McKean-Vlasov FBSDE is solvable:*

$$\begin{aligned} \mathrm{d}X_t &= b(t,X_t,\mathcal{L}(X_t|\mathcal{F}_t^B),\hat{\alpha}(t,X_t,Y_t,\mu_t))\,\mathrm{d}t + \sigma(t,X_t,\mathcal{L}(X_t|\mathcal{F}_t^B))\,\mathrm{d}W_t + \sigma^0(t,X_t,\mathcal{L}(X_t|\mathcal{F}_t^B))\,\mathrm{d}B_t, \\ \mathrm{d}Y_t &= -\partial_x \hat{H}(t,X_t,Y_t,\mathcal{L}(X_t|\mathcal{F}_t^B))\,\mathrm{d}t + Z_t\,\mathrm{d}W_t + Z_t^0\,\mathrm{d}B_t. \end{aligned} \tag{C.5}$$

Moreover, the mean-field control-distribution flow pair is given by

$$\alpha_t^* = \hat{\alpha}(t,X_t,Y_t,\mathcal{L}(X_t|\mathcal{F}_t^B)), \quad \mu_t^* = \mathcal{L}(X_t|\mathcal{F}_t^B), \quad \forall t\in[0,T]. \tag{C.6}$$

Theorem C.2. *Under Assumption C.1, the FBSDE systems* (C.3) *and* (C.5) *have unique solutions. Moreover, let* $\mu_t^1,\mu_t^2 \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$ *be different given flow of measures, and denote by* $(X_t^i,Y_t^i,Z_t^i,Z_t^{0,i})$ *the unique solution to FBSDE* (C.3) *given* μ_t^i*, then*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|\Delta X_t\|^2 + \sup_{t\in[0,T]}\|\Delta Y_t\|^2 + \int_0^T \|\Delta Z_t\|^2 + \|\Delta Z_t^0\|^2\,\mathrm{d}t\right] \le C_{K,T}\mathbb{E}\left[\int_0^T (\Delta\mu_t)^2\,\mathrm{d}t\right], \tag{C.7}$$

where $\Delta X_t = X_t^1 - X_t^2$*,* $\Delta Y_t, \Delta Z_t, \Delta Z_t^0$ *are defined similarly, and* $\Delta\mu_t = \mathcal{W}_2(\mu_t^1,\mu_t^2)$*.*

Proof. The results generalize Theorem 3.1.3, Proposition 3.1.4 and Theorem 3.1.6 in Ahuja (2015) to the multi-dimensional case and with Lipschitz SDE coefficients b,σ,σ^0. The original proofs rely on Theorem 3.1.1 and Theorem 3.1.2 under Assumption **H** in Ahuja (2015). With the additional conditions on (b,σ,σ^0) in our setting, Assumption **H** of Ahuja (2015) still holds. We omit the details because they essentially parallel the corresponding derivations in Ahuja (2015). □

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. The proof uses the estimate (C.7) repeatedly. We first observe that, for $\mu_t = \mathcal{L}(X_t|\mathcal{F}_t^B)$ and $\mu_t' = \mathcal{L}(X_t'|\mathcal{F}_t^B)$, one has

$$\mathbb{E}[\mathcal{W}_2^2(\mu_t,\mu_t')] \le \mathbb{E}[\|X_t - X_t'\|^2], \quad \forall t\in[0,T]. \tag{C.8}$$

Then we define a map Φ by

$$\mu = \{\mu_t\}_{0\le t\le T} \to \Phi(\mu) := \{\mathcal{L}(X_t^\mu|\mathcal{F}_t^B)\}_{0\le t\le T}, \tag{C.9}$$

where X_t^{μ} is the optimal controlled process in FBSDE (C.3) given $\mu \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$. Combining (C.8) and (C.7) gives

$$\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\Phi(\mu_t),\Phi(\mu_t'))] \le \sup_{t\in[0,T]} \mathbb{E}[\|X_t^{\mu} - X_t^{\mu'}\|^2]$$
$$\le C_{K,T}\mathbb{E}\left[\int_0^T \mathcal{W}_2^2(\mu_t,\mu_t')\,\mathrm{d}t\right] \le C_{K,T}T \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\mu_t,\mu_t')]. \quad \text{(C.10)}$$

Thus, for sufficiently small T, Φ is a contraction map. By definition, μ_t^* defined in (C.6) is a fixed point of Φ: $\Phi(\mu^*) = \mu^*$. Let $\mu^{(0)}$ be the initial guess of μ^*, and $\mu^{(n)}$ be the resulted flow of measures of X_t given $\tilde{\mu}^{(n-1)}$ which is the approximation of $\mu^{(n-1)}$ by truncated signatures. So the measure flows are generated by

$$\mu^{(0)} \to \mu^{(1)} \rightsquigarrow \tilde{\mu}^{(1)} \to \mu^{(2)} \rightsquigarrow \tilde{\mu}^{(2)} \cdots \to \mu^{(n-1)} \rightsquigarrow \tilde{\mu}^{(n-1)} \to \mu^{(n)} \rightsquigarrow \tilde{\mu}^{(n)} \quad \text{(C.11)}$$

where $\to$ corresponds to the map Φ, and $\rightsquigarrow$ corresponds to the truncated signature approximation. Therefore, with (C.10) and the assumption $\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n)},\mu_t^{(n)})] \le \epsilon$ in Theorem 4.1, and denoting by $2C_{K,T}T = q$, we deduce that

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n)},\mu_t^*)] &\le 2\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n)},\mu_t^{(n)})] + 2\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\mu_t^{(n)},\mu_t^*)] \\
&\le 2\epsilon + 2C_{K,T}T \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n-1)},\mu_t^*)] = 2\epsilon + q\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n-1)},\mu_t^*)] \\
&\le 2\epsilon + q(2\epsilon + q\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n-2)},\mu_t^*)]) \\
&\le \cdots \\
&\le 2\epsilon(1+q+q^2+\dots q^{n-1}) + q^n \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\mu_t^{(0)},\mu_t^*)] \\
&= \frac{2-2q^n}{1-q}\epsilon + q^n \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\mu_t^{(0)},\mu_t^*)].
\end{aligned}$$

With sufficiently small T, one has $0<q<1$. To estimate $\int_0^T \mathbb{E}|\alpha_t^{(n)} - \alpha_t^*|^2\,\mathrm{d}t$, we observe that

$$\alpha_t^{(n)} - \alpha_t^* = \hat{\alpha}(t, X_t^{\tilde{\mu}^{(n-1)}}, Y_t^{\tilde{\mu}^{(n-1)}}, \tilde{\mu}_t^{(n-1)}) - \hat{\alpha}(t, X_t^*, Y_t^*, \mu_t^*), \quad \text{(C.12)}$$

where $(X_t^{\tilde{\mu}^{(n-1)}}, Y_t^{\tilde{\mu}^{(n-1)}})$ is the solution to FBSDE (C.3) given $\tilde{\mu}^{(n-1)}$, and (X_t^*, Y_t^*) can be viewed as the solution to FBSDE (C.3) given μ^*. Then using the Lipschitz property of $\hat{\alpha}$ in (t,x,μ) and (C.7) again produces

$$\begin{aligned}
\int_0^T \mathbb{E}|\alpha_t^{(n)} - \alpha_t^*|^2\,\mathrm{d}t &\le C_{K,T}\mathbb{E}\left[\int_0^T \|X_t^{\tilde{\mu}^{(n-1)}} - X_t^*\|^2 + \|Y_t^{\tilde{\mu}^{(n-1)}} - Y_t^*\|^2 + \mathcal{W}_2^2(\tilde{\mu}_t^{(n-1)},\mu_t^*)\,\mathrm{d}t\right] \\
&\le C_{K,T}T \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n-1)},\mu_t^*)].
\end{aligned}$$

Therefore, we obtain the desired result. □

Next we give the proof to Theorem 4.2.

Proof of Theorem 4.2. Consider a partition of $[0,T] : 0 = t_0 < \cdots < t_L = T$, and define $\pi(t) = t_k$ for $t \in [t_k, t_{k+1})$ with $\|\pi\| = \max_{1\le k<L} |t_k - t_{k-1}|$, then by following the line of the proof to Theorem 4.1, one only needs an additional estimate on $\mathbb{E}|X_t^{\mu} - X_{t_k}^{(n)}|^2$ to complete the proof. Noticing that X_t solves (2.3) with μ^* and $X_{t_k}^{(n)}$ satisfies (4.5) with $\tilde{\mu}^{(n-1)}$, one can obtain the estimate by following Lemma 14 in Carmona & Laurière (2019) with $N=1$. □

D. Benchmark Solutions

This appendix summarizes the analytical solutions to the three examples in Section 5, which are used to benchmark our algorithm's performance.

Linear-Quadratic MFGs. The analytical solution is provided in Carmona et al. (2015):

$$m_t := \mathbb{E}[X_t|\mathcal{F}_t^B] = \mathbb{E}[X_0] + \rho\sigma B_t, \quad t \in [0,T], \tag{D.1}$$

$$\alpha_t = (q+\eta_t)(m_t - X_t), \quad t \in [0,T], \tag{D.2}$$

where η_t is a deterministic function solving the Riccati equation:

$$\dot{\eta}_t = 2(a+q)\eta_t + \eta_t^2 - (\epsilon - q^2), \quad \eta_T = c,$$

with the solution given by

$$\eta_t = \frac{-(\epsilon - q^2)(e^{(\delta^+ - \delta^-)(T-t)} - 1) - c(\delta^+ e^{(\delta^+ - \delta^-)(T-t)} - \delta^-)}{(\delta^- e^{(\delta^+ - \delta^-)(T-t)} - \delta^+) - c(e^{(\delta^+ - \delta^-)(T-t)} - 1)}.$$

Here $\delta^{\pm} = -(a+q) \pm \sqrt{R}$, $R = (a+q)^2 + (\epsilon - q^2) > 0$, and the minimized expected cost is $V(0, x_0 - \mathbb{E}[x_0])$ with

$$V(t,x) = \frac{\eta_t}{2}x^2 + \mu_t, \quad \mu_t = \frac{1}{2}\sigma^2(1-\rho^2)\int_t^T \eta_s \,\mathrm{d}s.$$

The benchmark trajectories in Figure 2 are simulated according to (5.3) with m_t and α_t in (D.1) and (D.2).

Mean-field Portfolio Game Given the type vector $\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma)$, the analytical solution provided in Lacker & Zariphopoulou (2019) is summarized below

$$\pi_t^* = \delta\frac{\mu}{\sigma^2+\nu^2} + \theta\frac{\sigma}{\sigma^2+\nu^2}\frac{\phi}{1-\psi},$$
$$m_t = \mathbb{E}[\xi] + \mathbb{E}[\mu\pi^*]t + \mathbb{E}[\sigma\pi^*]B_t,$$

where $\phi = \mathbb{E}[\delta\frac{\mu\sigma}{\sigma^2+\nu^2}]$ and $\psi = \mathbb{E}[\theta\frac{\sigma^2}{\sigma^2+\nu^2}]$. Note that, since the type vector ζ is random representing the heterogenuity of agents in this mean-field game, π^* is a random strategy. The maximized expected utility of this game is given by $\mathbb{E}[v(0, \xi - \theta\mathbb{E}[\xi])]$, with

$$v(t,x) = -e^{-x/\delta}e^{-\rho(T-t)}, \quad \rho = \frac{1}{2(\sigma^2+\nu^2)}\left(\mu + \frac{\theta}{\delta}\frac{\phi}{1-\psi}\sigma\right)^2 - \frac{\theta}{\delta}\left(\tilde{\psi} + \frac{\tilde{\phi}\phi}{1-\psi}\right) - \frac{1}{2}\left(\frac{\theta}{\delta}\frac{\phi}{1-\psi}\right)^2,$$

$$\tilde{\psi} = \mathbb{E}\left[\delta\frac{\mu^2}{\sigma^2+\nu^2}\right], \quad \tilde{\phi} = \mathbb{E}\left[\theta\frac{\mu\sigma}{\sigma^2+\nu^2}\right].$$

Note that Figure 3(c) plots the absolute value of $\mathbb{E}[v(0, \xi - \theta\mathbb{E}[\xi])]$.

Mean-field Game of Optimal Consumption and Investment Following Lacker & Soret (2020), the analytical solution is given by

$$\pi_t^* \equiv \pi^* = \frac{\delta\mu}{\sigma^2+\nu^2} - \frac{\theta(\delta-1)\sigma}{\sigma^2+\nu^2}\frac{\phi}{1+\psi}, \quad c_t^* = \left(\frac{1}{\beta} + (\frac{1}{\lambda} - \frac{1}{\beta})e^{-\beta(T-t)}\right)^{-1}, \tag{D.3}$$

where

$$\phi = \mathbb{E}\left[\frac{\delta\mu\sigma}{\sigma^2+\nu^2}\right], \quad \psi = \mathbb{E}\left[\frac{\theta(\delta-1)\sigma^2}{\sigma^2+\nu^2}\right], \quad \lambda = \epsilon^{-\delta}\left(e^{\mathbb{E}\left[\log(\epsilon^{-\delta})\right]}\right)^{-\frac{\theta(\delta-1)}{1+\mathbb{E}[\theta(\delta-1)]}},$$
$$\beta = \theta(\delta-1)\frac{\mathbb{E}\left[\delta\rho\right]}{1+\mathbb{E}\left[\theta(\delta-1)\right]} - \delta\rho,$$

and

$$\rho = \left(1 - \frac{1}{\delta}\right)\left\{\frac{\delta}{2(\sigma^2+\nu^2)}\left(\mu - \sigma\frac{\phi}{1+\psi}\theta(1-\frac{1}{\delta})\right)^2 + \frac{1}{2}\left(\frac{\phi}{1+\psi}\right)^2\theta^2\left(1-\frac{1}{\delta}\right)\right.$$
$$\left. - \theta\mathbb{E}\left[\frac{\delta\mu^2 - \theta(\delta-1)\sigma\mu\frac{\phi}{1+\psi}}{\sigma^2+\nu^2}\right] + \frac{\theta}{2}\mathbb{E}\left[\frac{(\delta\mu - \theta(\delta-1)\sigma\frac{\phi}{1+\psi})^2}{\sigma^2+\nu^2}\right]\right\}.$$

Note that the expression of m_t, Γ_t and the maximized expected utility are not given in Lacker & Soret (2020). For completeness, we give their derivations below. Since c_t^* in (D.3) doesn't depend on the common noise B, $\Gamma_t := \exp \mathbb{E}[\log c_t^* | \mathcal{F}_t^B]$ admits a unique formula for all agents

$$\Gamma_t = \exp \mathbb{E}[\log c_t^*].$$

To obtain the formula for $m_t := \exp \mathbb{E}[\log X_t^* | \mathcal{F}_t^B]$, we first deduce by Itô's formula that

$$\mathrm{d} \log X_t^* = \pi_t^* (\mu \,\mathrm{d}t + \nu \,\mathrm{d}W_t + \sigma \,\mathrm{d}B_t) - \frac{1}{2}(2c_t^* + (\pi_t^*)^2 \sigma^2 + (\pi_t^*)^2 \nu^2) \,\mathrm{d}t, \tag{D.4}$$

from which we easily get

$$\mathbb{E}[\log X_t^* | \mathcal{F}_t^B] = \mathbb{E}[\log \xi] + \mathbb{E}[\pi^* \mu - \frac{1}{2}(\pi^*)^2(\sigma^2 + \nu^2)]t - \int_0^t \mathbb{E}[c_s^*] \,\mathrm{d}t + \pi^* \sigma B_t,$$

and $m_t = \exp \mathbb{E}[\log X_t^* | \mathcal{F}_t^B]$. The maximized expected utility of this game is given by $\mathbb{E}[v(0, \xi, \mathbb{E}[\xi])]$, with

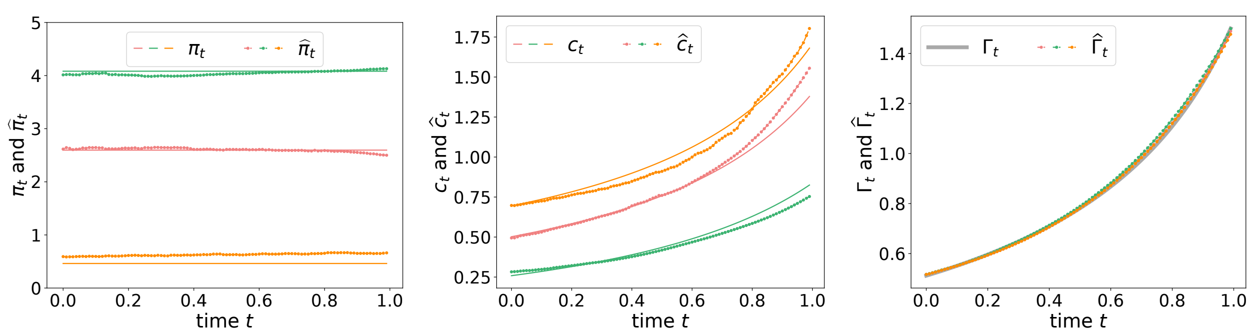
$$v(t, x, y) = \epsilon \left(1 - \frac{1}{\delta}\right)^{-1} x^{1 - \frac{1}{\delta}} y^{-\theta(1 - \frac{1}{\delta})} f(t),$$

and $f(t)$ is defined by

$$f(t) = \exp \left\{ \int_t^T \left(\rho + \frac{1}{\delta} c_s^* + \mathbb{E}[c_s^*] \left(1 - \frac{1}{\delta}\right) \theta \right) \mathrm{d}s \right\}.$$

Note that, to ensure the positiveness of X_t required by using the power utility, the trajectories of X_t are obtained by simulating $\log X_t$ via (D.4) then taking the exponential.

E. Plots of π_t, c_t, $\Gamma_t = \exp \mathbb{E}(\log c_t | \mathcal{F}_t^B)$ for Mean-Field Game of Optimal Consumption and Investment



(a) π_t (b) c_t (c) $\Gamma_t = \exp \mathbb{E}(\log c_t | \mathcal{F}_t^B)$

Figure 5. Plots on test data for three different $(X_0^i, W^i, B^i, \zeta^i)$. Solid line is the benchmark solution and dashed line is the numerical approximation using the Sig-DFP algorithm. Each panel presents three trajectories of π_t, c_t, and $\Gamma_t = \exp \mathbb{E}(\log c_t | \mathcal{F}_t^B)$ and their approximations. Parameter choices are: $\delta \sim U(2, 2.5), \mu \sim U(0.25, 0.35), \nu \sim U(0.2, 0.4), \theta, \xi \sim U(0, 1), \sigma \sim U(0.2, 0.4)$, $\epsilon \sim U(0.5, 1)$.

F. Experiment setup for the high-dimensional case $n_0 = 5$

To test the performance of Sig-DFP in high dimensions, we implement a toy experiment on the mean-field game of optimal consumption and investment with the common noise of dimension $n_0 = 5$. Specifically, we modify the $\sigma \,\mathrm{d}B_t$ term in (5.7) to be in high dimensions, *i.e.*, X_t now follows

$$\mathrm{d}X_t = \pi_t X_t (\mu \,\mathrm{d}t + \nu \,\mathrm{d}W_t + \boldsymbol{\sigma}^{\mathrm{T}} \,\mathrm{d}\boldsymbol{B}_t) - c_t X_t \,\mathrm{d}t,$$

where $\boldsymbol{\sigma} := (\sigma_1, \dots, \sigma_5)^{\mathrm{T}}$, $\boldsymbol{B}_t$ is a 5-dimensional Brownian motion, and $X_0 = \xi$. We use the same hyperparameters for training and provide the running time in Table 6.