Signatured Deep Fictitious Play for Mean Field Games with Common Noise

Ming Min [*1] **Ruimeng Hu** [*1 2]

Abstract

Existing deep learning methods for solving mean-field games (MFGs) with common noise fix the sampling common noise paths and then solve the corresponding MFGs. This leads to a nested-loop structure with millions of simulations of common noise paths in order to produce accurate solutions, which results in prohibitive computational cost and limits the applications to a large extent. In this paper, based on the rough path theory, we propose a novel single-loop algorithm, named signatured deep fictitious play, by which we can work with the unfixed common noise setup to avoid the nested-loop structure and reduce the computational complexity significantly. The proposed algorithm can accurately capture the effect of common uncertainty changes on mean-field equilibria without further training of neural networks, as previously needed in the existing machine learning algorithms. The efficiency is supported by three applications, including linear-quadratic MFGs, mean-field portfolio game, and mean-field game of optimal consumption and investment. Overall, we provide a new point of view from the rough path theory to solve MFGs with common noise with significantly improved efficiency and an extensive range of applications. In addition, we report the first deep learning work to deal with extended MFGs (a mean-field interaction via both the states and controls) with common noise.

1. Introduction

Stochastic differential games study the strategic interaction of rational decision-makers in an uncertain dynamical system, and have been widely applied to many areas, including social science, system science, and computer science. For realistic models, the problem usually lacks tractability and needs numerical methods. With a large number of players resulting in high-dimensional problems, conventional algorithms soon lose efficiency and one may resort to recently developed machine learning tools (Hu, 2021; Han & Hu, 2020; Han et al., 2020). On the other hand, one could utilize its limiting mean-field version, mean-field games (MFGs), to approximate the n-player game for large n (*e.g.*, Han et al. (2021)). Introduced independently in Huang et al. (2006); Lasry & Lions (2007), MFGs study the decision making problem of a continuum of agents, aiming to provide asymptotic analysis of the finite player model in which players interact through their empirical distribution. In an MFG, each agent is infinitesimal, whose decision can not affect the population law. Therefore, the problem can be solved by focusing on the optimal decision of a representative agent in response to the average behavior of the entire population and a fixed-point problem (*cf.* equation (2.5)). The MFG model has inspired tremendous applications, not only in finance and economics, such as system risk (Carmona et al., 2015), high-frequency trading (Lachapelle et al., 2016) and crowd trading (Cardaliaguet & Lehalle, 2018), but also to population dynamics (Achdou et al., 2017; Djehiche et al., 2017; Achdou & Lasry, 2019) and sanitary vaccination (Hubert & Turinici, 2018; Elie et al., 2020a), to list a few. For a systematical introduction of MFGs, see Caines et al. (2017); Carmona & Delarue (2018a;b).

In MFGs, the random shocks to the dynamical system can be from two sources: idiosyncratic to the individual players and common to all players, *i.e.*, decision-makers face correlated randomness. While MFGs were initially introduced with only idiosyncratic noise as seen in most of the literature, games with common noise, referred to as *MFGs with common noise*, have attracted significant attention recently (Lacker & Webster, 2015; Carmona et al., 2016; Ahuja, 2016; Graber, 2016). The inclusion of common noise is natural in many contexts, such as multi-agent trading in a common stock market, or systemic risk induced through inter-bank lending/borrowing. In reality, players make decisions in a common environment (*e.g.*, trade in the same stock market). Therefore, their states are subject to correlated random shocks, which can be modeled by individual noises and a common noise. In this modeling, observing the state dynamics will be sufficient, and one does not need

[*]Equal contribution [1]Department of Statistics and Applied Probability, University of California, Santa Barbara, CA 93106-3110, USA [2]Department of Mathematics, University of California, Santa Barbara, CA 93106-3080, USA. Correspondence to: Ruimeng Hu <rhu@ucsb.edu>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

to observe the noises. These applications make it crucial to develop efficient and accurate algorithms for computing MFGs with common noise.

Theoretically, MFGs with common noise can be formulated as an infinite-dimensional master equation, which is the type of second-order nonlinear Hamilton-Jacobi-Bellman equation involving derivatives with respect to a probability measure. Therefore, direct simulation is infeasible due to the difficulty of discretizing the probability space. An alternative way of solving MFGs with common noise is to formulate it into a stochastic Fokker-Planck/Hamilton-Jacobi-Bellman system, which has a complicated form with common noise, forward-backward coupling, and second-order differential operators. The third kind of approaches turns it into forward backward stochastic differential equations (FBSDE) of McKean-Vlasov type (*cf.* Carmona & Delarue (2018b, Chapter 2)), which in general requires convexity of the Hamiltonian. For all three approaches, the common assumption is the monotonicity condition that ensures uniqueness. Regarding simulation, existing deep learning methods fix the sampling common noise paths and then solve the corresponding MFGs, which leads to a nested-loop structure with millions of simulations of common noise paths to produce accurate predictions for unseen common shock realizations. Then the computational cost becomes prohibitive and limits the applications to a large extent.

In this paper, we solve MFGs with common noise by directly parameterizing the optimal control using deep neural networks in spirit of (Han & E, 2016), and conducting a global optimization. We integrate the signature from rough path theory, and fictitious play from game theory for efficiency and accuracy, and term the algorithm *Signatured Deep Fictitious Play* (Sig-DFP). The proposed algorithm avoids solving the three aforementioned complicated equations (master equation, Stochastic FP/HJB, FBSDE) and does not have uniqueness issues.

Contribution. We design a novel efficient single-loop deep learning algorithm, Sig-DFP, for solving MFGs with common noise by integrating fictitious play (Brown, 1949) and Signature (Lyons et al., 2007) from rough path theory. To our best knowledge, this is the first work focusing on the common noise setting, which can address heterogeneous MFGs and heterogeneous extended MFGs, both with common noise.

We prove that the Sig-DFP algorithm can reach mean-field equilibria as both the depth M of the truncated signature and the stage n of the fictitious play approaching infinity, subject to the universal approximation of neural networks. We demonstrate its convergence superiority on three benchmark examples, including homogeneous MFGs, heterogeneous MFGs, and heterogeneous extended MFGs, all with common noise, and with assumptions even beyond the technical requirements in the theorems. Moreover, the algorithm has the following advantages:

1. Temporal and spacial complexity are $\mathcal{O}(NLp+Np^2)$ and $\mathcal{O}(NLp)$, compared to $\mathcal{O}(N^2L)$ (for both time and space) in existing machine learning algorithms, with N as the sample size, L as the time discretization size, $p = \mathcal{O}(n_0^M)$, n_0 as the dimension of common noise.

2. Easy to apply the fictitious play strategy: only need to average over linear functionals with $\mathcal{O}(1)$ complexity.

Related Literature. After MFGs firstly introduced by Huang et al. (2006) and Lasry & Lions (2007) under the setting of a continuum of homogeneous players but without common noise, it has been extended to many applicable settings, *e.g.*, heterogeneous players games (Lacker & Zariphopoulou, 2019; Lacker & Soret, 2020) and major-minor players games (Huang, 2010; Nourian & Caines, 2013; Carmona & Zhu, 2016). A recent line of work studies MFGs with common noise (Carmona et al., 2015; Bensoussan et al., 2015; Ahuja, 2016; Cardaliaguet et al., 2019). Despite its theoretical progress and importance for applications, efficient numerical algorithms focusing on common noise settings are still missing. Our work will fill this gap by integrating machine learning tools with learning procedures from game theory and signature from rough path theory.

Fictitious play was firstly proposed in Brown (1949; 1951) for normal-form games, as a learning procedure for finding Nash equilibria. It has been widely used in the Economic literature, and adapted to MFGs (Cardaliaguet & Hadikhanloo, 2017; Briani & Cardaliaguet, 2018) and finite-player stochastic differential games (Hu, 2021; Han & Hu, 2020; Han et al., 2020; Xuan et al., 2021).

Using machine learning to solve MFGs has also been considered, for both model-based setting (Carmona & Laurière, 2019; Ruthotto et al., 2020; Lin et al., 2020) and model-free reinforcement learning setting (Guo et al., 2019; Tiwari et al., 2019; Angiuli et al., 2020; Elie et al., 2020b), most of which did not consider common noise. Existing machine learning methods for MFGs with common noise were studied in Perrin et al. (2020), which have a nested-loop structure and require millions of simulations of common noise paths to produce accurate predictions for unseen common shock realizations.

The signature in rough path theory has been recently applied to machine learning as a feature map for sequential data. For example, Király & Oberhauser (2019); Bonnier et al. (2019); Toth & Oberhauser (2020); Min & Ichiba (2020) have used signatures in natural language processing, time series, and handwriting recognition, and Chevyrev & Oberhauser (2018); Ni et al. (2020) studied the relation between signatures and distributions of sequential data. We refer to Lyons & Qian (2002); Lyons et al. (2007) for a more

detailed introduction of the signature and rough path theory.

2. Mean Field Games with Common Noise

We first introduce the following notations to precisely define MFGs with common noise. For a fixed time horizon T, let $(W_t)_{0\le t\le T}$ and $(B_t)_{0\le t\le T}$ be independent n- and n_0-dimensional Brownian motions defined on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$. We shall refer W as the *idiosyncratic noise* and B as the *common noise* of the system. Let $\mathcal{F}_t^B$ be the filtration generated by $(B_t)_{0\le t\le T}$, and $\mathcal{P}^p(\mathbb{R}^d)$ be the collection of probability measures on $\mathbb{R}^d$ with finite p^{th} moment, *i.e.*, $\mu \in \mathcal{P}^p(\mathbb{R}^d)$ if

$$\left(\int_{\mathbb{R}^d} \|x\|^p \,\mathrm{d}\mu(x)\right)^{1/p} < \infty. \tag{2.1}$$

We denote by $\mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$ the space of continuous $\mathcal{F}^B$-adapted stochastic flow of probability measures with the finite second moment, and by $\mathcal{H}^2([0,T];\mathbb{R}^m)$ the set of all $\mathcal{F}$-progressively measurable $\mathbb{R}^m$-valued square-integrable processes.

Next, we introduce the concept of MFGs with common noise. Given an initial distribution $\mu_0 \in \mathcal{P}^2(\mathbb{R}^d)$, and a stochastic flow of probability measures $\mu = (\mu_t)_{0\le t\le T} \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$, we consider the stochastic control

$$\inf_{(\alpha_t)_{0\le t\le T}} \mathbb{E}[\int_0^T f(t, X_t, \mu_t, \alpha_t)\,\mathrm{d}t + g(X_T, \mu_T)], \tag{2.2}$$

$$\text{where } \mathrm{d}X_t = b(t, X_t, \mu_t, \alpha_t)\,\mathrm{d}t + \sigma(t, X_t, \mu_t, \alpha_t)\,\mathrm{d}W_t + \sigma^0(t, X_t, \mu_t, \alpha_t)\,\mathrm{d}B_t, \tag{2.3}$$

with $X_0 \sim \mu_0$. Here the representative agent controls his dynamics X_t through a $\mathbb{R}^m$-dimensional control process α_t, and the drift coefficient b, diffusion coefficients σ and σ^0, running cost f and terminal cost g are all measurable functions, with $(b, \sigma, \sigma^0, f) : [0,T]\times\mathbb{R}^d\times\mathcal{P}^2(\mathbb{R}^d)\times\mathbb{R}^m \to \mathbb{R}^d \times \mathbb{R}^{d\times n} \times \mathbb{R}^{d\times n_0} \times \mathbb{R}$, and $g : \mathbb{R}^d \times \mathcal{P}^2(\mathbb{R}^d) \to \mathbb{R}$.

Note that since μ is stochastic, (2.2)–(2.3) is a control problem with random coefficients.

Definition 2.1 (Mean-field equilibrium)**.** *The control-distribution flow pair $\alpha^* = (\alpha_t^*)_{0\le t\le T} \in \mathcal{H}^2([0,T];\mathbb{R}^m)$, $\mu^* \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$ is a mean-field equilibrium to the MFG with common noise, if α^* solves (2.2) given the stochastic measure flow μ^*, and the conditional marginal distribution of the optimal path $X_t^{\alpha^*}$ given the common noise B coincides with the measure flow μ^*:*

$$\mu_t^* = \mathcal{L}(X_t^{\alpha^*}|\mathcal{F}_t^B), \tag{2.4}$$

where $\mathcal{L}(\cdot|\mathcal{F})$ is the conditional law given a filtration $\mathcal{F}$.

We remark that, with a continuum of agents, the measure μ^* is not affected by a single agent's choice, and the MFG is a standard control problem plus an additional fixed-point problem. More precisely, denote by $\hat{\alpha}^\mu$ the optimal control of (2.2)–(2.3) given the stochastic measure flow $\mu \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$, then μ^* is a fixed point of

$$\mu_t = \mathcal{L}(X_t^{\hat{\alpha}^\mu}|\mathcal{F}_t^B). \tag{2.5}$$

MFGs without common noise: Note that with $\sigma^0 \equiv 0$, (2.2)–(2.3) is a MFG without common noise, and the flow of measures μ_t becomes deterministic.

Extended MFGs: In extended mean field games, the interactions between the representative agent and the population happen via both the states and controls, thus the functions $(b, \sigma, \sigma^0, f, g)$ can also depend on $\mathcal{L}(\alpha_t|\mathcal{F}_t^B)$.

3. Fictitious Play and Signatures

The Signatured Deep Fictitious Play (Sig-DFP) algorithm is built on fictitious play, and propagates conditional distributions $\mu = \{\mu_t\}_{0\le t\le T} \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$ by signatures. This section briefly introduces these two ingredients.

In the learning procedure of *fictitious play*, players myopically choose their best responses against the empirical distribution of others' actions at every subsequent stage after arbitrary initial moves. When Cardaliaguet & Hadikhanloo (2017); Cardaliaguet & Lehalle (2018) extended it to mean-field settings, the empirical distribution of actions is naturally replaced by the average of distribution flows. More precisely, let $\bar{\mu}^{(0)} \in \mathcal{M}([0,T];\mathcal{P}^2(\mathbb{R}^d))$ be the initial guess of μ^* in (2.4), and consider the following iterative algorithm: (1) take $\bar{\mu}^{(n-1)} \in \mathcal{P}^2(\mathbb{R}^d)$ as the given flow of measures in (2.2)–(2.3) for the n-th iteration, and solve the optimal control in (2.2) denoted by $\alpha^{(n)}$; (2) solve the controlled stochastic differential equation (SDE) (2.3) for $X^{\alpha^{(n)}}$ and then infer the conditional distribution flow $\mu^{(n)} = \mathcal{L}(X^{\alpha^{(n)}}|\mathcal{F}_t^B)$; (3) average distributions $\bar{\mu}^{(n)} = \frac{n-1}{n}\bar{\mu}^{(n-1)} + \frac{1}{n}\mu^{(n)}$ and pass $\bar{\mu}^{(n)}$ to the next iteration. If $\mu^{(n)}$ converges and the strategy corresponding to the limiting measure flow is admissible, then by construction, it is a fixed-point of (2.5) and thus a mean-field equilibrium.

Signatures of Paths. Let $T((\mathbb{R}^d)) := \bigoplus_{k=0}^\infty (\mathbb{R}^d)^{\otimes k}$ be the tensor algebra, and denote by $\mathcal{V}^p([0,T],\mathbb{R}^d)$ the space of continuous mappings from $[0,T]$ to $\mathbb{R}^d$ with finite p-variation. For a path $x : [0,T] \to \mathbb{R}^d$, define the p-variation

$$\|x\|_p := \left(\sup_{D\subset[0,T]} \sum_{i=0}^{r-1} \|x_{t_{i+1}} - x_{t_i}\|^p\right)^{1/p}, \tag{3.1}$$

where $D \subset [0,T]$ denotes a partition $0 \le t_0 < t_1 < \ldots < t_r \le T$. We equip the space $\mathcal{V}^p([0,T],\mathbb{R}^d)$ with the norm $\|\cdot\|_{\mathcal{V}^p} := \|\cdot\|_\infty + \|\cdot\|_p$.

Definition 3.1 (Signature)**.** *Let $X \in \mathcal{V}^p([0,T],\mathbb{R}^d)$ such that the following integral makes sense. The signature of*

X, denoted by $S(X)$, is an element of $T((\mathbb{R}^d))$ defined by $S(X) = (1, X^1, \cdots, X^k \cdots)$ with

$$X^k = \int_{0<t_1<t_2<\cdots<t_k<T} \mathrm{d}X_{t_1} \otimes \cdots \otimes \mathrm{d}X_{t_k}. \quad (3.2)$$

We denote by $S^M(X)$ the truncated signature of X of depth M, i.e., $S^M(X) = (1, X^1, \cdots, X^M)$ and has the dimension $\frac{d^{M+1}-1}{d-1}$.

Note that when X is a semi-martingale (the case of our problems), equation (3.2) is understood in the Stratonovich sense. The following properties of the signature make it an ideal choice for our problem, with more details in Appendix A.

1. Signatures characterize paths uniquely up to the tree-like equivalence, and the equivalence is removed if at least one dimension of the path is strictly increasing (Boedihardjo et al., 2016). Therefore, we shall augment the original path with the time dimension in the algorithm, *i.e.*, working with $\hat{X}_t = (t, X_t)$ since $S(\hat{X})$ characterizes paths $\hat{X}$ uniquely.

2. Terms in the signature present a factorial decay property, which provides the accuracy of using a few terms in the signature (small M) to approximate a path.

3. As a feature map of sequential data, the signature has a universality detailed in the following theorem.

Theorem 3.1 (Universality, Bonnier et al. (2019))**.** *Let $p \geq 1$ and $f : \mathcal{V}^p([0,T], \mathbb{R}^d) \to \mathbb{R}$ be a continuous function in paths. For any compact set $K \subset \mathcal{V}^p([0,T], \mathbb{R}^d)$, if $S(x)$ is a geometric rough path for any $x \in K$, then for any $\epsilon > 0$ there exist $M > 0$ and a linear functional $l \in T((\mathbb{R}^d))^*$ such that*

$$\sup_{x \in K} |f(x) - \langle l, S(x) \rangle| < \epsilon. \quad (3.3)$$

4. The Sig-DFP Algorithm

We introduce two shorthand notations: if x is a path indexed by $t \in [0,T]$, then $x := (x_t)_{0 \leq t \leq T}$ denotes the whole path and $x_{s:t} := (x_u)_{s \leq u \leq t}$ denotes the path between s and t.

4.1. Propagation of Distribution with Signatures

With the presence of common noise, existing algorithms mostly consider a nested-loop structure, with the inner one for idiosyncratic noise W and the outer one for common noise B. More precisely, if one works with N idiosyncratic Brownian paths $\{W^k\}_{k=1}^N$ and N common Brownian paths $\{B^k\}_{k=1}^N$, then for each B^j, one needs to simulate N paths $\{X^{i,j}\}_{i=1}^N$ defined by (2.3) over all idiosyncratic Brownian paths and solve the problem (2.2) associated to B^j. This requires a total of N^2 simulations of (2.3). With a sufficiently large N, $\mu_t = \mathcal{L}(X_t|\mathcal{F}_t^B)$ is approximated well by $\frac{1}{N^2}\sum_{i,j=1}^N \delta_{X_t^{i,j}} \mathbb{1}_{\omega^{(0,j)}}$ with $\omega^{0,j} \in \Omega$ corresponding to the trajectory B^j. The double summation is of $\mathcal{O}(N^2)$ which is computationally expensive for large N.

We shall address the aforementioned numerical difficulties by signatures. The key idea is to approximate μ_t by

$$\mu_t \equiv \mathcal{L}(X_t|\mathcal{F}_t^B) = \mathcal{L}(X_t|S(\hat{B}_t)) \approx \mathcal{L}(X_t|S^M(\hat{B}_t)),$$
$$\text{with } \hat{B}_t = (t, B_t), \quad (4.1)$$

where the equal sign comes from the unique characterization of signatures $S(\hat{B})$ to the paths $B_{0:t}$, and the approximation is accurate for large M due to the factorial decay property of the signature. The last term is then computed by machine learning methods, *e.g.*, by Generative Adversarial Networks (GANs). In addition, if the agents interact via some population average subject to common noise: $\mu_t = \mathbb{E}[\iota(X_t)|\mathcal{F}_t^B]$, the approximation in (4.1) can be arbitrarily close to the true measure flow for sufficiently large M. The following lemma gives a precise statement.

Lemma 4.1. *Suppose $\mu_t = \mathbb{E}[\iota(X_t)|\mathcal{F}_t^B]$ where $\iota : \mathbb{R}^d \to \mathbb{R}$ is a measurable function. View μ_t as $\mu(t, B_{0:t})$ with $\mu : \mathcal{V}^p([0,T], \mathbb{R}^{n_0+1}) \to \mathbb{R}$ continuous for some $p \in (2,3)$, and let $K \subset \mathcal{V}^p([0,T], \mathbb{R}^{n_0+1})$ be a compact set, then for any $\epsilon > 0$, there exist a positive integer M and a linear functional $l \in T((\mathbb{R}^{n_0+1}))^*$, such that*

$$\sup_{t \in [0,T]} \sup_{\hat{B} \in K} |\mu_t - \langle l, S^M(\hat{B}_{0:t}) \rangle| < \epsilon. \quad (4.2)$$

Proof. See Appendix A for details due to the page limit. □

With all the above preparations, we now explain how the approximation to $\mu = \{\mu_t\}_{0 \leq t \leq T}$ using signatures is implemented. Given N pairs of idiosyncratic and common Brownian paths (W^i, B^i) and assume α_t in (2.3) is already obtained (which will be explained in Section 4.2), we first sample the optimized state processes $(X_t^i)_{0 \leq t \leq T}$, producing N samples $\{X^i\}_{i=1}^N$. Then the linear functional l in Lemma 4.1 is approximated by implementing linear regressions on $\{S^M(\hat{B}_{0:t}^i)\}_{i=1}^N$ with dependent variable $\{\iota(X_t^i)\}_{i=1}^N$ at several time stamps t, *i.e.*,

$$\hat{l} = \arg\min_{\boldsymbol{\beta}} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\|^2, \quad (4.3)$$
$$\boldsymbol{y} = \{\iota(X_t^i)\}_{i=1}^N, \ \boldsymbol{X} = \{S^M(\hat{B}_{0:t}^i)\}_{i=1}^N.$$

In all experiments in Section 5, we get decent approximations of μ on $[0,T]$ by considering only three time stamps $t = 0, \frac{T}{2}, T$. Note that such a framework can also deal with multi-dimensional ι, where the regression coefficients become a matrix.

The choice in (4.3) is mainly motivated by Lemma 4.1 stating l is a linear functional, and by the probability model underlying ordinary linear regression (OLS) which interprets that the least square minimization (4.3) gives the best

prediction of $E[\boldsymbol{y}|\boldsymbol{X}]$ restricting to linear relations. There are other benefits for choosing OLS: Once $\hat{l}$ is obtained in (4.3), the prediction for unseen common paths is efficient: $\mu_t(\tilde{\omega}) \approx \langle \hat{l}, S^M(\hat{B}_{0:t}(\tilde{\omega}))\rangle$ for any $\tilde{\omega}$ and t. Moreover, it is easy to integrate with fictitious play: averaging $\mu_t^{(n)}$ from different iterations, commonly needed in fictitious play, now means simply averaging $\hat{l}^{(n)}$ over n. Next, we analyze the temporal and spatial complexity of using signatures and linear regression as below.

Temporal Complexity: Suppose we discretize $[0, T]$ into L time stamps: $0 = t_0 \leq t_1 \leq \ldots \leq t_L = T$, and simulate N paths of W, B and X_t. The simulation cost is of $\mathcal{O}(NL)$. For computing the truncated signature $S^M(\hat{B})$ of depth M, we use the Python package Signatory (Kidger & Lyons, 2020), yielding a complexity of $\mathcal{O}(NLp)$ where $p = \frac{(n_0+1)^{M+1}-1}{n_0} = \mathcal{O}(n_0^M)$. Note that one can choose a large N and reuse all sampled common noise paths B for each iteration of fictitious play, thus the computation of $S^M(B)$ is done only once, and $S^M(\hat{B}_{0:t})$ is accessible in constant time for all t. The linear regression[1] (or Ridge regression) takes time $\mathcal{O}(Np^2)$. Thus, the total temporal complexity is of $\mathcal{O}(NLp + Np^2)$, which is linear in N given[2] $p \ll N$. Comparing to the nested-loop algorithm, where the cost of simulating SDEs is $\mathcal{O}(N^2L)$ and computing conditional distribution flows takes time $\mathcal{O}(N^2L)$, we claim that our algorithm reduced the temporal complexity by a factor of the sample size N by using signatures.

Spatial Complexity: In fictitious play, one may choose to average all past flow of measures $\mu^{(n)}$ as the given measures in (2.2)–(2.3) for the current iteration. Using signatures simplifies it to average $\hat{l}^{(n)}$. To update it between iterations, one needs to store the current average which costs $\mathcal{O}(p)$ of the memory. Combining $\mathcal{O}(NL)$ and $\mathcal{O}(NLp)$ for storing SDEs and truncated signatures, the overall spacial complexity is $\mathcal{O}(NLp)$. The complexity of the nested-loop case is again $\mathcal{O}(N^2L)$, which we reduce by a factor of N.

We conclude this section by the following remark: For the general case $\mu_t = \mathcal{L}(X_t|\mathcal{F}_t^B)$, though the linear regression is no longer available, the one-to-one mapping between μ and $S(\hat{B})$ persists. Therefore, one can train a Generative Adversarial Network (GAN, Goodfellow et al. (2014)) for generating samples following the distribution μ by taking truncated signatures as part of the network inputs.

4.2. Deep Learning Algorithm

Having explained the key idea on how to approximate μ efficiently, we describe the Sig-DFP algorithm in this subsection. The algorithm consists of repeatedly solving (2.2)–(2.3) for a given measure flow μ using deep learning in the spirit of Han & E (2016), and passing the yielded μ to the next iteration by using signatures. The flowchart of the idea is illustrated in Figure 1. Consider a partition π of $[0, T] : 0 = t_0 < \cdots < t_L = T$, denote by $\hat{\mu}^{(n-1)}$ the given flow of measures at stage n, the stochastic optimal control problem (2.2)–(2.3) is solved by

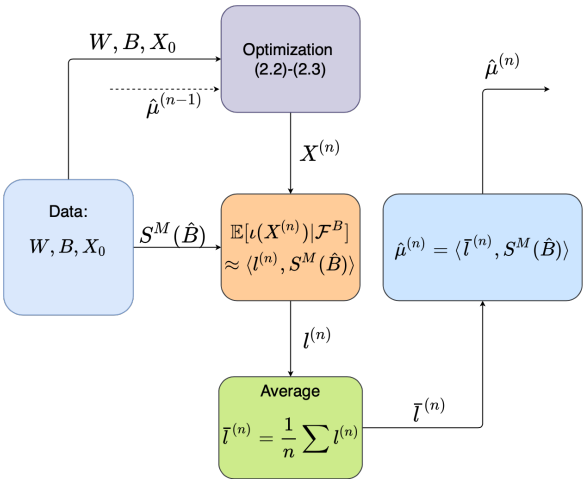


Figure 1. Flowchart of one iteration in the Sig-DFP Algorithm. Input: idiosyncratic noise W, common noise B, initial position X_0 and measure flow $\hat{\mu}^{(n-1)}$ from the last iteration. Output: measure flow $\hat{\mu}^{(n)}$ for the next iteration.

$$\inf_{\{\alpha_k\}_{k=0}^{N-1}} \frac{1}{N}\sum_{i=1}^{N}\Big(\sum_{k=0}^{L-1} f(t_k, X_k^i, \hat{\mu}_k^{(n-1)}(\omega^i), \alpha_k^i)\Delta_k + g(X_L^i, \hat{\mu}_L^{(n-1)}(\omega^i))\Big), \tag{4.4}$$

$$\begin{aligned} \text{where } X_{k+1}^i = X_k^i &+ b(t_k, X_k^i, \hat{\mu}_k^{(n-1)}(\omega^i), \alpha_k^i)\Delta_k \\ &+ \sigma(t_k, X_k^i, \hat{\mu}_k^{(n-1)}(\omega^i), \alpha_k^i)\Delta W_k^i \\ &+ \sigma^0(t_k, X_k^i, \hat{\mu}_k^{(n-1)}(\omega^i), \alpha_t^i)\Delta B_k^i, \end{aligned} \tag{4.5}$$

where we replace the subscript t_k by k to simplify notations, and let $\Delta_k = t_{k+1} - t_k$, $\Delta W_k^i = W_{t_{k+1}}^i - W_{t_k}^i$, $\Delta B_k^i = B_{t_{k+1}}^i - B_{t_k}^i$. Here, we use the superscript i to represent the i^{th} sample path and $\hat{\mu}_k^{(n-1)}(\omega^i)$ to emphasize the stochastic measure's dependence on the i^{th} sample path of B up to time t_k. The control α_k is then parameterized by neural networks (NNs) in the feedback form:

$$\alpha_k^i := \alpha_\varphi(t_k, X_k^i, \hat{\mu}_k^{(n-1)}(\omega^i); \varphi), \tag{4.6}$$

where α_φ denotes the NN map with parameters φ, and searching the infimum in (4.4) is translated into minimizing φ. The yielded optimizer φ^* gives $\alpha_k^{i,*}$, with which the optimized state process paths $\{X^{i,*}\}_{i=1}^N$ are simulated and

[1] We use the Python package scikit-learn (Pedregosa et al., 2011) to do the linear regression.

[2] M is usually small due to the factorial decay property of the signature. For n_0 not large, we have $p \ll N$.

Algorithm 1 The Sig-DFP Algorithm

Input: $b, \sigma, \sigma_0, f, g, \iota$ and $X_0^i, (W_{t_k}^i)_{k=0}^L, (B_{t_k}^i)_{k=0}^L$ for $i = 1, 2, \ldots, N$; N_{round}: rounds for FP;
B: minibatch size; N_{batch}: number of minibatches.
Compute the signatures of $\hat{B}^i_{0:t_k}$ for $i = 1, \ldots, N$, $k = 1, \ldots, L$;
Initialize $\hat{\mu}^{(0)}$, φ;
for $n = 1$ **to** N_{round} **do**
  for $r = 1$ **to** N_{batch} **do**
    Simulate the r^{th} minibatch of $X^{i,(n)}$ using $\hat{\mu}^{(n-1)}$ and compute $J_B(\varphi, \hat{\mu}^{(n-1)})$;
    Minimize $J_B(\varphi, \hat{\mu}^{(n-1)})$ over φ, then update α_φ;
  end for
  Simulate $X^{i,(n)}$ with the optimized α^*_φ, for $i = 1, \ldots, N$;
  Regress $\iota(X_0^{i,(n)}), \iota(X_{L/2}^{i,(n)}), \iota(X_L^{i,(n)})$ on $S^M(\hat{B}^i_{0:0})$, $S^M(\hat{B}^i_{0:t_{L/2}})$, $S^M(\hat{B}^i_{0:t_L})$ to get $l^{(n)}$;
  Update $\bar{l}^{(n)} = \frac{n-1}{n}\bar{l}^{(n-1)} + \frac{1}{n} l^{(n)}$;
  Compute $\hat{\mu}^{(n)}$ by $\hat{\mu}_k^{(n)}(\omega^i) = \langle \bar{l}^{(n)}, S^M(\hat{B}^i_{0:t_k})\rangle$, for $i = 1, 2, \ldots, N, k = 1, \ldots, L$;
end for
Output: the optimized α^*_φ and $\bar{l}^{(N_{\text{round}})}$.

its conditional law $\mathcal{L}(X^*|\mathcal{F}^B)$, denoted by $\mu^{(n)}$, is approximated using signatures as described in Section 4.1. This finishes one iteration of fictitious play. Denote by $\tilde{\mu}^{(n)}$ the approximation of $\mu^{(n)}$, we then pass $\tilde{\mu}^{(n)}$ to the next iteration via updating $\hat{\mu}^{(n)} = \frac{1}{n}\tilde{\mu}^{(n)} + \frac{n-1}{n}\hat{\mu}^{(n-1)}$ by averaging the coefficients in (4.3).

We summarize it in Algorithm 1, with implementation details deferred to Appendix B. Note that the simulation of $X^{i,(n)}$ and $J_B(\varphi, \bar{\mu}^{(n-1)})$ uses the equations (B.2) and (B.1) in Appendix B, respectively.

Theorem 4.1 (Convergence analysis)**.** *Let (α^*, μ^*) be the mean-field equilibrium in Definition 2.1, $\alpha^{(n)}$ be the optimal control, and $\mu^{(n)}$ be the measure flow of the optimized state process after the n^{th} iteration of fictitious play, and $\tilde{\mu}^{(n)}$ be the approximation by truncated signatures. Under Assumption C.1 and $\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n)}, \mu_t^{(n)})] \le \epsilon$, we have*

$$\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_t^{(n)}, \mu_t^*)] + \int_0^T \mathbb{E}|\alpha_t^{(n)} - \alpha_t^*|^2 \,\mathrm{d}t \le C(q^n \sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\mu_t^{(0)}, \mu_t^*)] + \epsilon),$$

for some constants $C > 0$ and $0 < q < 1$, where $\mathcal{W}_2$ denotes the 2-Wasserstein metric.

Moreover, if we 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

Theorem 4.2 (Convergence in discrete time)**.** *Let $\mu_{t_k}^{(n)}$ be the conditional law of the discretized optimal process $X_{t_k}^{(n)}$ after the n^{th} iteration of fictitious play (cf. (4.5)), and $\tilde{\mu}_{t_k}^{(n)}$ be the approximation by truncated signatures. Under Assumption C.1 and $\sup_{0\le k\le L} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_{t_k}^{(n)}, \mu_{t_k}^{(n)})] \le \epsilon$, one has*

$$\sup_{t\in[0,T]} \mathbb{E}[\mathcal{W}_2^2(\tilde{\mu}_{\pi(t)}^{(n)}, \mu_t^*)] + \int_0^T \mathbb{E}|\alpha_{\pi(t)}^{(n)} - \alpha_t^*|^2 \,\mathrm{d}t \le C(q^n \sup_{0\le k\le L} \mathbb{E}[\mathcal{W}_2^2(\mu_{t_k}^{(0)}, \mu_{t_k}^*)] + \epsilon + \|\pi\|),$$

for some constants $C > 0$ and $0 < q < 1$, where $\alpha_{t_k}^{(n)} = \hat{\alpha}(t_k, X_{t_k}, Y_{t_k}, \tilde{\mu}_{t_k}^{(n-1)})$, and (X_t, Y_t) solves (C.3) with μ replaced by $\tilde{\mu}_{t_k}^{(n-1)}$.

The proofs of Theorems 4.1 and 4.2 are given in Appendix C due to the page limit.

Remark that the Sig-DFP framework is flexible. We choose to solve (2.2)-(2.3) by direct parameterizing control policies α_t for the sake of easy implementation and the possible exploration of multiple mean-field equilibria. If the equilibrium is unique, with proper conditions on the coefficients b, σ, σ^0, f and g, one can reformulate (2.2)-(2.3) into McKean-Vlasov FBSDEs or stochastic FP/HJB equations, and solve them by fictitious play and propagating the common noise using signatures.

5. Experiments

In this section, we present the performance of Sig-DFP for three examples: homogeneous, heterogeneous, and heterogeneous extended MFGs. A relative L^2 metric will be used for performance measurement, defined for progressively measurable random processes as

$$L_R^2(x, \hat{x}) := \sqrt{\frac{\mathbb{E}[\int_0^T \|x_t - \hat{x}_t\|^2 \,\mathrm{d}t]}{\mathbb{E}[\int_0^T \|x_t\|^2 \,\mathrm{d}t]}}, \tag{5.1}$$

where x is a benchmark process and $\hat{x}$ is its prediction. We shall use stochastic gradient descent (SGD) optimizer for all three experiments. Training processes are done on a server with Intel Core i9-9820X (10 cores, 3.30 GHz) and RTX 2080 Ti GPU, and training time will be reported in Appendix B. Implementation codes are available at `https://github.com/mmin0/SigDFP`.

Data Preparation. For all three experiments, the size of both training and test data is $N = 2^{15}$, and the size of validation data is $N/2$. We fix $T = 1$ and discretize $[0, 1]$ by $t_k = \frac{k}{100}$, $k = 0, 1, \ldots, 100$. Initial states are generated independently by $X_0^i \sim \mu_0$, with $\mu_0 = U(0, 1)$ as the uniform distribution. The idiosyncratic Brownian motions W and

common noises B are generated by antithetic variates for variance reduction, *i.e.*, we generate the first half samples (W^i, B^i) and get the other half $(-W^i, -B^i)$ by flipping.

Benchmarks. The examples below are carefully chosen with analytical benchmark solutions. Due to the space limit, we provide the details in Appendix D.

Linear-Quadratic MFGs. We first consider a Linear-Quadratic MFG with common noise proposed in Carmona et al. (2015), formulated as below:

$$\inf_{\alpha} \mathbb{E}\bigg\{\int_0^T \left[\frac{\alpha_t^2}{2} - q\alpha_t(m_t - X_t) + \frac{\epsilon}{2}(m_t - X_t)^2\right] \mathrm{d}t + \frac{c}{2}(m_T - X_T)^2\bigg\}, \tag{5.2}$$

$$\text{where } \mathrm{d}X_t = [a(m_t - X_t) + \alpha_t]\,\mathrm{d}t + \sigma(\rho\,\mathrm{d}B_t + \sqrt{1-\rho^2}\,\mathrm{d}W_t). \tag{5.3}$$

Here $m_t = \mathbb{E}[X_t|\mathcal{F}_t^B]$ is the conditional population mean, $\rho \in [0,1]$ characterizes the noise correlation between agents, and $q, \epsilon, c, a, \sigma$ are positive constants. The agents have homogeneous preferences and aim to minimize their individual costs. We assume $q \le \epsilon^2$ so that the Hamiltonian is jointly convex in state and control variables, ensuring a unique mean-field equilibrium.

Training & Results. α_φ is a feedforward NN with two hidden layers of width 64. The truncated signature depth is chosen at $M = 2$. The model is trained for 500 iterations of fictitious play. The optimized state process $\hat{X}$ and its conditional mean $\hat{m}$ generated by test data are shown in Figures 2a and 2b. The minimized cost after each iteration computed using validation data is given in Figure 2c, where one can see a rapid convergence to the benchmark cost. During the experiments, we notice a slow convergence speed when using the average of $m^{(n)}$ in (5.3). This is because the initial guess $m^{(0)}$ is in general far from the truth. Therefore, for the first half of iterations, we simply use the previous-step result $m^{(n-1)}$. The learning rate is set as 0.1 for the first half and 0.01 for the second half of training. The relative L^2 errors for test data are listed in Table 1.

Table 1. Relative L^2 errors on test data for the LQ MFG.

	SDE X_t	CONTROL α_t	EQUILIBRIUM m_t
L^2_R	0.0031	0.0044	0.058

Mean-Field Portfolio Game. Our second experiment is performed on a heterogeneous MFG proposed by Lacker & Zariphopoulou (2019), where the agent's preference is different, characterized by a type vector ζ which is random and drawn at time 0. They all aim to maximize their exponential utility of terminal wealth compared to the population average:

$$\sup_{\pi} \mathbb{E}\left[-\exp\left(-\frac{1}{\delta}(X_T - \theta m_T)\right)\right], \tag{5.4}$$

where the dynamics are

$$\mathrm{d}X_t = \pi_t(\mu\,\mathrm{d}t + \nu\,\mathrm{d}W_t + \sigma\,\mathrm{d}B_t), \quad X_0 = \xi. \tag{5.5}$$

Here m represents the conditional mean $m_t := \mathbb{E}[X_t|\mathcal{F}_t^B]$, and $\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma)$ is random.

Training & Results. We use truncated signatures of depth $M = 2$ and a feedforward NN π_φ with 4 hidden layers[3] to approximate π. We train our model with 500 iterations of fictitious play. The learning rate starts at 0.1 and is reduced by a factor of 5 every 200 rounds. The relative L^2 errors evaluated under test data are listed in Table 2. Figure 3 compares X and m to their approximations, and plots the maximized utilities.

Table 2. Relative L^2 errors on test data for MF Portfolio Game.

	SDE X_t	INVEST π_t	EQUILIBRIUM m_t
L^2_R	0.068	0.035	0.085

Mean-Field Game of Optimal Consumption and Investment. Our last experiment considers an extended heterogeneous MFG proposed by Lacker & Soret (2020), where agents interact via both states and controls. The setup is similar to Lacker & Zariphopoulou (2019) except for including consumption and using power utilities. More precisely, each agent is characterized by a type vector $\zeta = (\xi, \delta, \theta, \mu, \nu, \sigma, \epsilon)$, and the optimization problem reads

$$\sup_{\pi, c} \mathbb{E}\left[\int_0^T U(c_t X_t (\Gamma_t m_t)^{-\theta}; \delta)\,\mathrm{d}t + \epsilon U(X_T m_T^{-\theta}; \delta)\right], \tag{5.6}$$

where $U(x;\delta) = \frac{1}{1-\frac{1}{\delta}} x^{1-\frac{1}{\delta}}$, $\delta \neq 1$, X_t follows

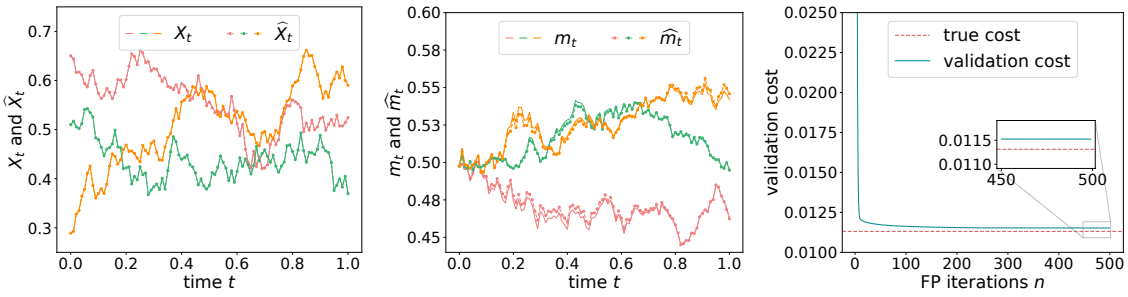
$$\mathrm{d}X_t = \pi_t X_t(\mu\,\mathrm{d}t + \nu\,\mathrm{d}W_t + \sigma\,\mathrm{d}B_t) - c_t X_t\,\mathrm{d}t, \tag{5.7}$$

and $X_0 = \xi$. Here $\Gamma_t = \exp\mathbb{E}[\log c_t|\mathcal{F}_t^B]$ and $m_t = \exp\mathbb{E}[\log X_t|\mathcal{F}_t^B]$ are the mean-field interactions from consumption and wealth.

Training & Results. For this experiment, we use truncated signatures of depth $M = 4$. The optimal controls $(\pi_t, c_t)_{0\le t\le 1}$ are parameterized by two neural networks π_φ and c_φ, each with three hidden layers.[4] Due to the extended mean-field interaction term Γ_t, we will propagate two conditional distribution flows, *i.e.*, two linear functionals $\bar{l}^{(n)}, \bar{l}_c^{(n)}$

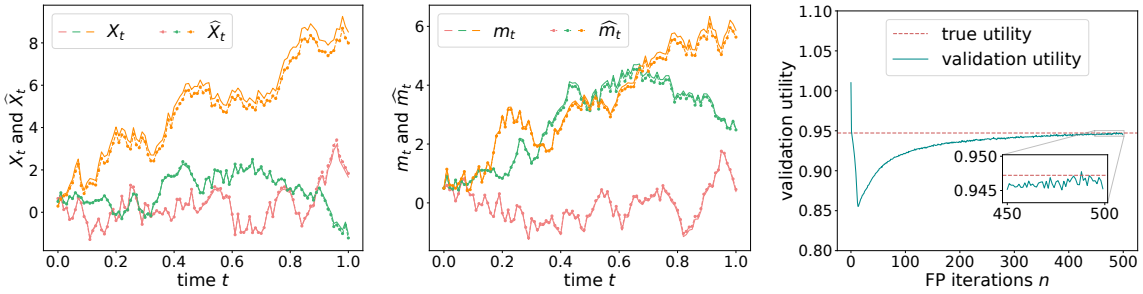
[3] Since agents are heterogeneous characterized by their type vectors ζ, π_φ takes (ζ, t, X_t, m_t) as inputs. Hidden neurons in each layer are (64, 32, 32, 16).

[4] Due to the nature of heterogeneous extended MFG, both α_φ and c_φ take $(\zeta_t, t, X_t, m_t, \Gamma_t)$ as inputs. Hidden neurons in each layer are (64, 64, 64).



(a) X_t (b) $m_t = \mathbb{E}(X_t|\mathcal{F}_t^B)$ (c) Minimized Cost

Figure 2. Panels (a) and (b) give three trajectories of X_t, $m_t = \mathbb{E}[X_t|\mathcal{F}_b^B]$ (solid lines) and their approximations (dashed lines) using different (X_0, W, B) from test data. Panel (c) shows the minimized cost computed using validation data over fictitious play iterations. Parameter choices are: $\sigma = 0.2, q = 1, a = 1, \epsilon = 1.5, \rho = 0.2, c = 1, x_0 \sim U(0, 1)$.



(a) X_t (b) $m_t = \mathbb{E}(X_t|\mathcal{F}_t^B)$ (c) Maximized Utility

Figure 3. Panels (a) and (b) give three trajectories of X_t, $m_t = \mathbb{E}[X_t|\mathcal{F}_b^B]$ (solid lines) and their approximations (dashed lines) using different (X_0, W, B) from test data. Panel (c) shows the maximized utility computed using validation data over fictitious play iterations. Parameter choices are: $\delta \sim U(5, 5.5), \mu \sim U(0.25, 0.35), \nu \sim U(0.2, 0.4), \theta \sim U(0, 1), \sigma \sim U(0.2, 0.4), \xi \sim U(0, 1)$.

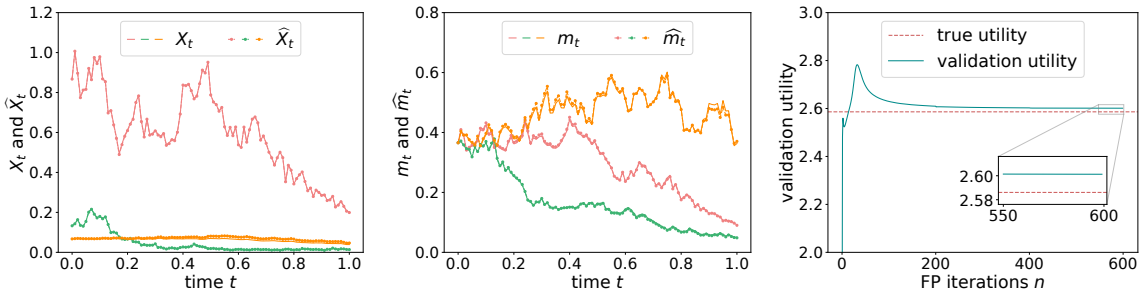
during each iteration of fictitious play. Instead of estimating m_t, Γ_t directly, we estimate $\mathbb{E}[\log X_t|\mathcal{F}_t^B], \mathbb{E}[\log c_t|\mathcal{F}_t^B]$ by $\langle \bar{l}^{(n)}, S^4(B_{0:t})\rangle$, $\langle \bar{l}_c^{(n)}, S^4(B_{0:t})\rangle$ and then take exponential to get m_t, Γ_t. To ensure the non-negativity condition of X_t, we evolve $\log X_t$ according to (D.4) and then take exponential to get X_t. For optimal consumption, c_φ is used to predicted $\log c_t$ and thus $\exp c_\varphi$ gives the predicted c_t. With 600 iterations of fictitious play and a learning rate of 0.1 decaying by a factor of 5 for every 200 iterations, the relative L^2 errors for test data are listed in Table 3. Figure 4 compares X and m to their approximations, and plots the maximized utilities. Plots of $\pi_t, c_t, \Gamma_t = \exp \mathbb{E}(\log c_t|\mathcal{F}_t^B)$ are provided in Appendix E.

Table 3. Relative L^2 errors on test data for Optimal Consumption and Investment MFG.

	INVEST π_t	CONSUMPTION c_t	m_t	Γ_t
L_R^2	0.1126	0.0614	0.0279	0.0121

Comparison with the nested algorithm. We run both Sig-DFP and the nested algorithm for the training data size of (INP, CNP)= $(2^4, 2^4)$, $(2^6, 2^6)$ $(2^8, 2^8)$, where INP means the number of individual noise paths and CNP means the number of common noise paths. From the comparisons of running time, memory, and relative L^2 errors in Tables 4 and 5, one can see that the accuracy is mainly affected by the size of (INP, CNP) used for training the neural network. Sig-DFP has the advantage of reducing memory request and running time, which allows it to use a larger size of data, *e.g.*, (INP, CNP)= $(2^{15}, 2^{15})$, to produce much better accuracy. The quadratic growth of memory in the nested algorithm, evidenced by the first three columns of data in Tables 4 (least squares growth rate ≈ 2), makes us unable to run the nested algorithm beyond $(2^8, 2^8)$ in our current computing environment due to its high demand for memory.

Comparisons of running time for different signature depth M and dimension n_0. We choose the data size (INP, CNP)= $(2^{15}, 2^{15})$ and compare the running time for different (n_0, M)'s in Table 6. Choosing $M = 1, 2, 3, 4$ yield the relative L^2 errors of controls (π, c) as $(15.9\%, 9.5\%)$, $(11.4\%, 6.3\%)$, $(11.4\%, 6.3\%)$ and $(11.3\%, 6.1\%)$ for $n_0 = 1$, respectively. Note that, compared to $M = 1$,



(a) X_t (b) $m_t = \exp \mathbb{E}(\log X_t | \mathcal{F}_t^B)$ (c) Maximized Utility

Figure 4. Panels (a) and (b) give three trajectories of X_t and $m_t = \exp \mathbb{E}(\log X_t | \mathcal{F}_t^B)$ (solid lines) and their approximation (dashed lines) using different (X_0, W, B) from test data. Panel (c) shows the maximized utility computed using validation data over fictitious play iterations. Parameter choices are: $\delta \sim U(2, 2.5), \mu \sim U(0.25, 0.35), \nu \sim U(0.2, 0.4), \theta, \xi \sim U(0, 1), \sigma \sim U(0.2, 0.4), \epsilon \sim U(0.5, 1)$.

Table 4. Running time (hours) and Memory (GBs) comparisons between Sig-DFP and the nested algorithm for different (INP, CNP)'s. INP = # of individual noise paths, CNP = # of common noise paths, and NA = Not Available due to high demand for memory.

(INP, CNP)	$(2^4, 2^4)$	$(2^6, 2^6)$	$(2^8, 2^8)$	$(2^{12}, 2^{12})$	$(2^{15}, 2^{15})$
NESTED ALGORITHM	$(0.09, 2.1)$	$(0.46, 4.1)$	$(4.3, 43.5)$	NA	NA
SIG-DFP	$(0.09, 1.9)$	$(0.1, 2.0)$	$(0.17, 2.3)$	$(0.33, 4.8)$	$(1.3, 27)$

Table 5. The comparisons of relative L^2 errors on (π, c) between Sig-DFP and the nested algorithm for different (INP, CNP)'s. INP = # of individual noise paths, CNP = # of common noise paths, and NA = Not Available due to high demand for memory.

(INP, CNP)	$(2^4, 2^4)$	$(2^6, 2^6)$	$(2^8, 2^8)$	$(2^{12}, 2^{12})$	$(2^{15}, 2^{15})$
NESTED ALGORITHM	$(53\%, 44\%)$	$(36\%, 41\%)$	$(79.4\%, 16.2\%)$	NA	NA
SIG-DFP	$(85.8\%, 48.1\%)$	$(43.3\%, 44.9\%)$	$(49\%, 43\%)$	$(18\%, 38\%)$	$(11\%, 6\%)$

Table 6. The comparisons of running time (hours) for different signature depth M and dimension n_0 using (INP, CNP)= $(2^{15}, 2^{15})$.

(n_0, DEPTH M)	$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$
RUNNING TIME (HOURS)	1.2	1.2	1.2	1.3	1.2	1.3	1.5	2.6

taking $M = 2$ improves the accuracy significantly but not $M = 3, 4$. This is because the curves of $\log(c_t)$ and $\log(X_t^*)$ are approximately either linear or quadratic in t, as shown in Figure 5 in Appendix E after taking a logarithm, which implies that the signatures of depth $M = 2$ will be sufficient to produce good accuracy. We remark that Sig-DFP has no difficulty computing high-dimensional problems, evidenced by the running time of $n_0 = 5$ cases in Table 6. We focus on one-dimensional problems since, to our best knowledge, the closed-form non-trivial solutions only exist in one-dimensional cases, which can serve as the benchmark solutions. More details about $n_0 = 5$ are given in Appendix F.

6. Conclusion

In this paper, we propose a novel single-loop algorithm, named signatured deep fictitious play, for solving mean-field games (MFGs) with common noise. We incorporate signature from rough path theory into the strategy of deep fictitious play (Hu, 2021; Han & Hu, 2020; Han et al., 2020), and avoid the nested-loop structure in existing machine learning methods, which reduces the computational cost significantly. Analysis of the complexity and convergence for the proposed algorithm is provided. The effectiveness of the algorithm is justified by three applications, and in particular, we report the first deep learning work to deal with extended MFGs with common noise. In the future, we shall study deep learning algorithms for MFGs with common noise in more general settings (Hu & Zariphopoulou, 2021).

Acknowledgement

R.H. was partially supported by NSF grant DMS-1953035, the Faculty Career Development Award and the Research Assistance Program Award, University of California, Santa Barbara. M.M. and R.H. are grateful to the reviewers for their valuable and constructive comments.

References

Achdou, Y. and Lasry, J.-M. Mean field games for modeling crowd motion. In *Contributions to partial differential equations and applications*, pp. 17–42. Springer, 2019.

Achdou, Y., Bardi, M., and Cirant, M. Mean field games models of segregation. *Mathematical Models and Methods in Applied Sciences*, 27(01):75–113, 2017.

Ahuja, S. *Mean Field Games with Common Noise*. PhD thesis, Stanford University, 2015.

Ahuja, S. Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM Journal on Control and Optimization*, 54(1):30–48, 2016.

Angiuli, A., Fouque, J.-P., and Laurière, M. Unified reinforcement Q-learning for mean field game and control problems. *arXiv preprint arXiv:2006.13912*, 2020.

Bensoussan, A., Frehse, J., and Yam, S. C. P. The master equation in mean field theory. *Journal de Mathématiques Pures et Appliquées*, 103(6):1441–1474, 2015.

Boedihardjo, H., Geng, X., Lyons, T., and Yang, D. The signature of a rough path: uniqueness. *Advances in Mathematics*, 293:720–737, 2016.

Bonnier, P., Kidger, P., Arribas, I. P., Salvi, C., and Lyons, T. Deep signature transforms. In *Advances in Neural Information Processing Systems 32 (NeurIPS)*, 2019.

Briani, A. and Cardaliaguet, P. Stable solutions in potential mean field game systems. *Nonlinear Differential Equations and Applications NoDEA*, 25(1):1–26, 2018.

Brown, G. W. Some notes on computation of games solutions. Technical report, RAND CORP SANTA MONICA CA, 1949.

Brown, G. W. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

Caines, P. E., Huang, M., and Malhamé, R. P. Mean field games. In *Handbook of Dynamic Game Theory*, pp. 1–28. Springer, 2017.

Cardaliaguet, P. and Hadikhanloo, S. Learning in mean field games: the fictitious play. *ESAIM: Control, Optimisation and Calculus of Variations*, 23(2):569–591, 2017.

Cardaliaguet, P. and Lehalle, C.-A. Mean field game of controls and an application to trade crowding. *Mathematics and Financial Economics*, 12(3):335–363, 2018.

Cardaliaguet, P., Delarue, F., Lasry, J.-M., and Lions, P.-L. *The Master Equation and the Convergence Problem in Mean Field Games:(AMS-201)*, volume 201. Princeton University Press, 2019.

Carmona, R. and Delarue, F. *Probabilistic Theory of Mean Field Games with Applications I*. Springer, 2018a.

Carmona, R. and Delarue, F. *Probabilistic Theory of Mean Field Games with Applications II*. Springer, 2018b.

Carmona, R. and Laurière, M. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II–the finite horizon case. *arXiv preprint arXiv:1908.01613*, 2019.

Carmona, R. and Zhu, X. A probabilistic approach to mean field games with major and minor players. *Annals of Applied Probability*, 26(3):1535–1580, 2016.

Carmona, R., Fouque, J.-P., and Sun, L.-H. Mean field games and systemic risk. *Communications in Mathematical Sciences*, 13(4):911–933, 2015.

Carmona, R., Delarue, F., and Lacker, D. Mean field games with common noise. *Annals of Probability*, 44(6):3740–3803, 2016.

Chevyrev, I. and Oberhauser, H. Signature moments to characterize laws of stochastic processes. *arXiv preprint arXiv:1810.10971*, 2018.

Djehiche, B., Tcheukam, A., and Tembine, H. A mean-field game of evacuation in multilevel building. *IEEE Transactions on Automatic Control*, 62(10):5154–5169, 2017.

Elie, R., Hubert, E., and Turinici, G. Contact rate epidemic control of COVID-19: an equilibrium view. *Mathematical Modelling of Natural Phenomena*, 15:35, 2020a.

Elie, R., Pérolat, J., Laurière, M., Geist, M., and Pietquin, O. On the convergence of model free learning in mean field games. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pp. 7143–7150, 2020b.

Friz, P. K. and Victoir, N. B. *Multidimensional stochastic processes as rough paths: theory and applications*, volume 120. Cambridge University Press, 2010.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27 (NIPS)*, 2014.

Graber, P. J. Linear quadratic mean field type control and mean field games with common noise, with application to production of an exhaustible resource. *Applied Mathematics & Optimization*, 74(3):459–486, 2016.

Guo, X., Hu, A., Xu, R., and Zhang, J. Learning mean-field games. In *Advances in Neural Information Processing Systems 32 (NeurIPS)*, 2019.

Han, J. and E, W. Deep learning approximation for stochastic control problems. *Deep Reinforcement Learning Workshop, NIPS*, 2016.

Han, J. and Hu, R. Deep fictitious play for finding Markovian Nash equilibrium in multi-agent games. In *Mathematical and Scientific Machine Learning (MSML)*, volume 107, pp. 221–245. PMLR, 2020.

Han, J., Hu, R., and Long, J. Convergence of deep fictitious play for stochastic differential games. *arXiv preprint arXiv:2008.05519*, 2020.

Han, J., Hu, R., and Long, J. A class of dimensionality-free metrics for the convergence of empirical measures. *arXiv preprint arXiv:2104.12036*, 2021.

Hu, R. Deep fictitious play for stochastic differential games. *Communications in Mathematical Sciences*, 19(2):325–353, 2021.

Hu, R. and Zariphopoulou, T. N-player and mean-field games in Itô-diffusion markets with competitive or homophilous interaction. *arXiv preprint arXiv:2106.00581*, 2021.

Huang, M. Large-population LQG games involving a major player: the Nash certainty equivalence principle. *SIAM Journal on Control and Optimization*, 48(5):3318–3353, 2010.

Huang, M., Malhamé, R. P., and Caines, P. E. Large population stochastic dynamic games: closed-loop mckean-vlasov systems and the nash certainty equivalence principle. *Communications in Information and Systems*, 6(3): 221–252, 2006.

Hubert, E. and Turinici, G. Nash-MFG equilibrium in a SIR model with time dependent newborn vaccination. *Ricerche di matematica*, 67(1):227–246, 2018.

Kidger, P. and Lyons, T. Signatory: differentiable computations of the signature and logsignature transforms, on both CPU and GPU. *arXiv preprint arXiv:2001.00706*, 2020.

Király, F. J. and Oberhauser, H. Kernels for sequentially ordered data. *Journal of Machine Learning Research*, 20 (31):1–45, 2019.

Lachapelle, A., Lasry, J.-M., Lehalle, C.-A., and Lions, P.-L. Efficiency of the price formation process in presence of high frequency participants: a mean field game analysis. *Mathematics and Financial Economics*, 10(3):223–262, 2016.

Lacker, D. and Soret, A. Many-player games of optimal consumption and investment under relative performance criteria. *Mathematics and Financial Economics*, 14(2): 263–281, 2020.

Lacker, D. and Webster, K. Translation invariant mean field games with common noise. *Electronic Communications in Probability*, 20, 2015.

Lacker, D. and Zariphopoulou, T. Mean field and n-agent games for optimal investment under relative performance criteria. *Mathematical Finance*, 29(4):1003–1038, 2019.

Lasry, J.-M. and Lions, P.-L. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

Lin, A. T., Fung, S. W., Li, W., Nurbekyan, L., and Osher, S. J. APAC-Net: Alternating the population and agent control via two neural networks to solve high-dimensional stochastic mean field games. *arXiv preprint arXiv:2002.10113*, 2020.

Lyons, T. and Qian, Z. *System control and rough paths*. Oxford University Press, 2002.

Lyons, T. J., Caruana, M., and Lévy, T. *Differential equations driven by rough paths*. Springer, 2007.

Min, M. and Ichiba, T. Convolutional signature for sequential data. *arXiv preprint arXiv:2009.06719*, 2020.

Ni, H., Szpruch, L., Wiese, M., Liao, S., and Xiao, B. Conditional Sig-Wasserstein GANs for time series generation. *arXiv preprint arXiv:2006.05421*, 2020.

Nourian, M. and Caines, P. E. ϵ-Nash mean field game theory for nonlinear stochastic dynamical systems with major and minor agents. *SIAM Journal on Control and Optimization*, 51(4):3302–3331, 2013.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

Perrin, S., Pérolat, J., Laurière, M., Geist, M., Elie, R., and Pietquin, O. Fictitious play for mean field games: Continuous time analysis and applications. In *Advances in Neural Information Processing Systems 33 (NeurIPS)*, 2020.

Ruthotto, L., Osher, S. J., Li, W., Nurbekyan, L., and Fung, S. W. A machine learning framework for solving high-dimensional mean field game and mean field control problems. *Proceedings of the National Academy of Sciences*, 117(17):9183–9193, 2020.

Tiwari, N., Ghosh, A., and Aggarwal, V. Reinforcement learning for mean field game. *arXiv preprint arXiv:1905.13357*, 2019.

Toth, C. and Oberhauser, H. Bayesian learning from sequential data using gaussian processes with signature covariances. In *Proceedings of the 37th International Conference on Machine Learning (ICML)*, volume 119, pp. 9548–9560. PMLR, 2020.

Xuan, Y., Balkin, R., Han, J., Hu, R., and Ceniceros, H. D. Optimal policies for a pandemic: A stochastic game approach and a deep learning algorithm. In *Mathematical and Scientific Machine Learning (MSML)*, 2021. Accepted. arXiv:2012.06745.