Research project : Generative Modeling via Mean-Field Control in discrete time

January 20, 2026

1 Research orientation : A backward approach

Let $T > 0$ be a finite time horizon. We now fix $N \in \mathbb{N}^\star$ which will be the number of time-steps discretization of the interval $[0,T]$ and for any $k \in [\![0,N]\!]$, we denote $t_k := k\frac{T}{N}$ and $\Delta t = t_{k+1} - t_k$. Now, given a control process $\alpha = (\alpha_{t_k})_{k\in[\![0,N]\!]}$ and a family of independent gaussian random variables $W := (W_{t_k})_{k\in[\![0,N-1]\!]}$, we look at the following proces $X = (X_{t_k})_{k\in[\![0,N]\!]}$ with dynamics given by

$$\begin{cases} X_{t_{k+1}} &= b\big(k, X_{t_k}, \alpha_{t_k}\big) + \sigma\big(k, X_{t_k}, \alpha_{t_k}\big) W_{t_k}, \quad k \in [\![0, N-1]\!], \\ X_0 &\sim \mu, \end{cases} \tag{1.1}$$

where μ is a given initial probability measure and b_k, σ_k are measurable maps defined over $\mathbb{R}^d \times A$ into $\mathbb{R}^d, \mathbb{R}^{d\times n}$ and we are looking to minimize over $\alpha = (\alpha_{t_k})_{k\in[\![0,N]\!]}$ the quantity

$$\inf_{\alpha\in\mathcal{A}} J^l(\alpha) = \mathbb{E}\Big[\sum_{k=0}^{N-1} f\big(k, X_{t_k}, \alpha_{t_k}, \mathbb{P}_{X_{t_k}}\big) + lg\big(X_{t_N}, \mathbb{P}_{X_{t_N}}\big)\Big], \tag{1.2}$$

where $\mathcal{A}$ denotes the set of admissible controls, where $l \in \mathbb{N}^*$, and where f and g are respectively maps defined on $\mathbb{R}^d \times A \times \mathcal{P}_2(\mathbb{R}^d)$ and $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ into $\mathbb{R}$.

Remark 1.1

- *The maps L and g should be understood as penalty terms for the process X, in the sense that f and g will be maps penalizing if for any $k \in [\![0,N]\!]$, $\mathbb{P}_{X_{t_k}}$ is large from ρ_{t_k}.*
- *$(\rho_{t_k})_{k\in[\![0,N]\!]}$ is going to be a given family of probability distribution that we are aiming to "reproduce". On the numerical aspect, $\rho_T = \rho_{t_N}$ will correspond to the terminal distribution that we aim to be "close enough" in a sense to be determined and $[\![0,N]\!] \ni k \mapsto \rho_{t_k} \in \mathcal{P}_2(\mathbb{R}^d)$ denotes trajectory on the space of probability measures that we would like to be close enough.*
- *The parameter l is going to represent the diversity in the samples that we aim to reproduce. In fact, as $l \nearrow \infty$, the penalty term $g(X_{t_N}, \mathbb{P}_{X_{t_N}}; \rho_T)$ should go towards 0 and g should be constructed such that $Im(g) \subset \mathbb{R}^+$ with $g = 0$ i.i.f $\mathbb{P}_{X_{t_N}} = \rho_T$. Similarly, we can penalize the term f if necessary.*

The goal is to solve the optimal control problem (1.1)-(1.2) by finding an explicit construction of

$$\hat{\alpha}^l = \arg\min_{\alpha\in\mathcal{A}} J^l(\alpha),$$

We will discuss a way to solve the control problem (1.1)-(1.2) relying on the Pontryagin's maximum principle which states that under some regularity conditions on the maps (b, σ, f, g), one can construct an optimal control $\alpha \in \mathcal{A}$ by means of a Discrete-Time Forward–Backward Equations . For this, we introduce the Hamiltonian map defined on $[\![0, N]\!] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d\times n} \times A$ into $\mathbb{R}$ defined by

$$H(k, x, \mu, y, z, a) = b(k, u, x, \mu, a) \cdot y + \sigma(k, u, x, \mu, a) : z + f(k, u, x, \mu, a), \tag{1.3}$$

where $\cdot$ represent the scalar product on $\mathbb{R}^d$ and $:$ denotes the scalar product on matrices, i.e $A : B = \mathrm{Tr}(A^\top B)$. We also use the notation for derivatives with respect to measure ($\frac{\delta}{\delta m} l$ which is called linear functional derivative of the map $l : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ and is a map defined on $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ into $\mathbb{R}$. In the following, we denote by $\mathcal{F}_k = \sigma((W_{t_i})_{i\leq k}, U)$, i.e the filtration generated by the random process W and U where U is a uniform random variable independant of W used for randomization of the initial condition of X, i.e allowing X_0 to be a non-trivial random variable. We then define $\mathbb{F} = (\mathcal{F}_k)_{k\in[\![0,N]\!]}$. We refer to [2] (Chapter 6) for an introduction to this notion of differentiability over the space of measures and to [1] for an overview of its use in the context of optimal control via Pontryagin's maximum principle in **discrete time**.

$$\begin{cases} X_{t_k} = b\big(k-1, X_{t_{k-1}}, \mathbb{P}_{X_{t_{k-1}}}, \alpha_{t_{k-1}}\big)\Delta t + \sigma\big(k-1, X_{t_{k-1}}, \mathbb{P}_{X_{t_{k-1}}}, \alpha_{t_{k-1}}\big) W_{t_{k-1}} \ , \\ Y_{t_k} = \mathbb{E}\Big[\partial_x H(k+1, X_{t_{k+1}}, \mathbb{P}_{X_{t_{k+1}}}, Y_{t_{k+1}}, Z_{t_{k+1}}, \alpha_{t_{k+1}}) \\ \qquad\qquad + \tilde{\mathbb{E}}\Big[\partial_{\tilde{x}} \frac{\delta}{\delta m} H(k+1, \tilde{X}_{t_{k+1}}, \mathbb{P}_{X_{t_{k+1}}}, \tilde{Y}_{t_{k+1}}, \tilde{Z}_{t_{k+1}}, \tilde{\alpha}_{t_{k+1}})(X_{t_{k+1}})\Big] \Big| \mathcal{F}_k\Big] \ , \\ Z_k = \mathbb{E}\Big[\Big(\partial_x H(k+1, X_{t_{k+1}}, \mathbb{P}_{X_{t_{k+1}}}, Y_{t_{k+1}}, Z_{t_{k+1}}, \alpha_{t_{k+1}}) \\ \qquad\qquad + \tilde{\mathbb{E}}\Big[\partial_{\tilde{x}} \frac{\delta}{\delta m} H(k+1, \tilde{X}_{t_{k+1}}, \mathbb{P}_{X_{t_{k+1}}}, \tilde{Y}_{t_{k+1}}, \tilde{Z}_{t_{k+1}}, \tilde{\alpha}_{t_{k+1}})(X_{t_{k+1}})\Big) W_{t_k}^\top \Big| \mathcal{F}_k\Big] \ , \\ X_0 \sim \mu \ , \\ Y_{t_{N-1}} = \mathbb{E}\Big[\partial_x g\big(X_{t_N}, \mathbb{P}_{X_{t_N}}\big) + \tilde{\mathbb{E}}\Big[\partial_{\tilde{x}} \frac{\delta}{\delta m} g\big(\tilde{X}_{t_N}, \mathbb{P}_{X_{t_N}}\big)(X_{t_N})\Big] \Big| \mathcal{F}_{N-1}\Big] \ , \\ Z_{t_{N-1}} = \mathbb{E}\Big[\Big(\partial_x g\big(X_{t_N}, \mathbb{P}_{X_{t_N}}\big) + \tilde{\mathbb{E}}\Big[\partial_{\tilde{x}} \frac{\delta}{\delta m} g\big(\tilde{X}_{t_N}, \mathbb{P}_{X_{t_N}}\big)(X_{t_N})\Big) W_{t_{N-1}}^\top \Big| \mathcal{F}_{N-1}\Big] \end{cases} \tag{1.4}$$

where $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{\alpha})$ is an independent copy of (X, Y, Z, α) defined over the space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$.

The main result can be summarized as below

Proposition 1.2 *Under some regularity assumptions on (b, σ, f, g), denote by $\hat{\alpha} = (\hat{\alpha}_t)_{0\leq t\leq T}$ the process defined as*

$$\hat{\alpha}_{t_k} = \underset{a\in A}{arg\ min}\ H(k, X_{t_k}, \mathbb{P}_{X_{t_k}}, Y_{t_k}, Z_{t_k}, a) = \hat{a}(k, X_{t_k}, \mathbb{P}_{X_{t_k}}, Y_{t_k}, Z_{t_k}), \quad \mathbb{P} - a.s, \quad t \in [0, T],$$

where H has been defined in (1.3) and for $\hat{a}$ measurable map defined over $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R}^{d\times n}$ then the unique solution to the equation (1.4) when replacing $\alpha = (\alpha_k)_{k\in[\![0,N]\!]}$ by $\hat{\alpha} = (\hat{\alpha}_k)_{k\in[\![0,N]\!]}$, exists and yields an optimal control given by $\hat{\alpha}$.

Remark 1.3 (A simple model). *We consider the model (approximation of the SDE model $\mathrm{d}X_t = \alpha_t \mathrm{d}t + \sigma \mathrm{d}W_t$) given by*

$$\begin{cases} X_{t_{k+1}} & = X_{t_k} + \alpha_{t_k}\Delta t + \sigma\sqrt{\Delta t} W_{t_k}, \\ X_0 & \sim \mu \end{cases}$$

with cost functional given by

$$J^l(\alpha) = \mathbb{E}\Big[\sum_{k=0}^{N-1} \frac{1}{2}|\alpha_{t_k}|^2 \Delta t + l g(X_{t_N}, \mathbb{P}_{X_{t_N}})\Big],$$

where g can for instance represent the Kullback-Leibler divergence between 2 absolutely continuous measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$. According to Proposition 1.2, we have in this case

$$H(k, x, \mu, y, z, a) = \big(x + a\Delta t\big) \cdot y + \sigma \Delta t : z + \frac{1}{2}|a|^2,$$

and $\hat{\alpha}_{t_k} = -Y_{t_k}\Delta t$ such that we have

$$\begin{cases} X_{t_{k+1}} & = X_{t_k} - Y_{t_k}\Delta t + \sigma\sqrt{\Delta t} W_{t_k}, \quad k \in [\![0, N-1]\!] \\ Y_{t_k} & = \mathbb{E}\Big[Y_{t_{k+1}}|\mathcal{F}_k\Big], \quad k \in [\![0, N-2]\!], \\ Z_{t_k} & = \mathbb{E}\Big[Y_{t_{k+1}} W_{t_k}^\top |\mathcal{F}_k\Big], \quad k \in [\![0, N-2]\!], \\ Y_{t_{N-1}} & = \mathbb{E}\Bigg[\tilde{\mathbb{E}}\Big[\partial_{\tilde{x}}\frac{\delta}{\delta m} g(\tilde{X}_{t_N}, \mathbb{P}_{X_{t_N}})(X_{t_N})\Big]\Big|\mathcal{F}_{N-1}\Bigg], \\ Z_{t_{N-1}} & = \tilde{\mathbb{E}}\Bigg[\partial_{\tilde{x}}\frac{\delta}{\delta m} g(\tilde{X}_{t_N}, \mathbb{P}_{X_{t_N}})(X_{t_N}) W_{t_{N-1}}^\top |\mathcal{F}_{N-1}\Bigg] \end{cases} \tag{1.5}$$

In this case, we have

$$X_{t_N} = \sum_{k=0}^{n-1} X_{t_k} - \Delta t \sum_{k=0} Y_{t_k} + \sigma\sqrt{\Delta t}\sum_{k=0}^{N-1} W_{t_{k+1}}.$$

The goal is therefore to learn conditional expectations at every time t_k for $k \in [\![0, N-2]\!]$ and then to sample X_{t_N} such that $\mathbb{P}_{X_{t_N}}$ should be close enough in order to get new samples of μ_T (or likely samples). Therefore, we see that we need to learn efficiently $(X_{t_{k_1}}, Y_{t_{k_2}}, Z_{t_{k_2}})_{k_1 \in [\![1,N]\!], k_2 \in [\![0,N-1]\!]}$. We precise here the notion of $\frac{\delta}{\delta m} g(\mu|\nu)(\cdot)$ in this setting, we can show that

$$\partial_x \frac{\delta}{\delta m} g(\mu, \nu)(x) = \nabla_x log\ \mu(x) - \nabla_x\ log\ \nu(x). \tag{1.6}$$

In (1.6), we see the apparition of the score function $x \mapsto \nabla_x log\big(\mu(x)\big)$, for which we can rely to the known results on score matching ?

Here the process $(Z_{t_k})_{k \in [\![0,N-1]\!]}$ as we do not control the volatility of the process but it could be included in the model !.

1.1 Theoretical questions

On the theoretical aspects, the natural questions to be addressed are the following.

- Let $(X^l, Y^l, Z^l) = (X^l_{t_k}, Y^l_{t_k}, Z^l_{t_k})_{k_1 \in [\![1,N]\!], k_2 \in [\![0,N-1]\!]}$ the processes solution to the system (1.5). Can we find a positive sequence $\epsilon(l) \underset{l\to\infty}{\to} 0$ such that under a suitable metric on $\mathcal{P}_2(\mathbb{R}^d)$ denoted by d (for instance KL or $\mathcal{W}_2$ among others), we have

$$\mathrm{d}(\mathbb{P}_{X^l_{t_N}}, \mu_T) \leq \epsilon(l). \tag{1.7}$$

  In fact, Equation (1.7) should be understood as a suitable way to allow diversity in the samples we are going to generate via the parameter $k \in \mathbb{N}^*$.

- A discussion as we should maybe formulate (1.4) in the latent space in the case of images. This comes as a second part for me since we should focus for now on simulation on low dimensional spaces where we can focus on the direct simulation of $(X^l_{t_{k_1}}, Y^l_{t_{k_2}}, Z^l_{t_{k_2}})_{k_1 \in [\![1,N]\!], k_2 \in [\![0,N-1]\!]}$ to obtain likely samples μ_T by $\mathbb{P}_{X^l_{t_N}}$.

1.2 Numerical questions

We will now discuss some methods to numerically solve the control problem (1.1)-(**??**).

On the numerical aspects, the following questions should be addressed

- The key role of the map g. In fact, we should ensure some convexity assumptions on g and f to ensure existence and uniqueness for the system (1.5) which should be verified under some standard functions over the space of measures, i.e KL among others.
- We need to understand how to learn (X^l, Y^l, Z^l) efficiently but this will be done by means of approximating conditional expectations.
- We can start by trying to learn some standard distributions, e.g **Gaussian Distributions**, **Student distribution**, and even in higher dimensions. We need to test numerically the impact of the parameter $l \in \mathbb{N}^*$, showing that l big should lead to a quite good accuracy in the learning of μ_T while l low provides almost nothing.

References

[1] Tianyang Nie Bozhang Dong and Zhen Wu. Maximum principle for discrete-time stochastic control problem of mean-field type.

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