Mean-field Langevin System, Optimal Control and Deep Neural Networks

Kaitong Hu, Anna Kazeykina, Zhenjie Ren

Mean-field Langevin System, Optimal Control and Deep Neural Networks

Kaitong HU * Anna KAZEYKINA † Zhenjie REN‡

Abstract

In this paper, we study a regularised relaxed optimal control problem and, in particular, we are concerned with the case where the control variable is of large dimension. We introduce a system of mean-field Langevin equations, the invariant measure of which is shown to be the optimal control of the initial problem under mild conditions. Therefore, this system of processes can be viewed as a continuous-time numerical algorithm for computing the optimal control. As an application, this result endorses the solvability of the stochastic gradient descent algorithm for a wide class of deep neural networks.

Keywords: Mean-Field Langevin Dynamics, Gradient Flow, Neural Networks

MSC: 60H30, 37M25

1 Introduction

This paper revisits the classical optimal control problem, that is,

$$\inf_\alpha V^0(\alpha), \quad \text{where} \quad V^0(\alpha) := \int_0^T L(t, X_t^\alpha, \alpha_t)dt + G(X_T^\alpha) \quad \text{and} \quad X_t = x_0 + \int_0^t \phi(r, X_r^\alpha, \alpha_r)dr \tag{1.1}$$

In particular, we aim at providing a feasible algorithm for solving such problem (indeed, its regularized version) when the dimensions of the state X and of the control α are both large.

It has been more than half a century since the discovery of Pontryagin's maximum principle [1], which states that in order to be an optimal control to the problem (1.1), α^* needs to satisfy the forward-backward ODE system:

$$\begin{cases} \alpha_t^* = \operatorname{argmin}_a H(t, X_t^*, a, P_t^*), \quad \text{where} \quad H(t,x,a,p) := L(t,x,a) + p \cdot \phi(t,x,a), \\ X_t^* = x_0 + \int_0^t \phi(r, X_r^*, \alpha_r^*)dr, \\ P_t^* = \nabla_x G(X_T^*) + \int_0^t \nabla_x H(r, X_r^*, \alpha_r^*, P_r^*)dr. \end{cases} \tag{1.2}$$

It is worth mentioning that this necessary condition becomes sufficient if one imposes convexity condition on the coefficients. To solve the forward-backward system, the most naive way is to follow a fixed-point algorithm, that is, starting with an arbitrary control α, evaluate the forward equation and then the backward one, and eventually compute a new control $\tilde{\alpha}$ by solving the optimization problem on the top line. Under some mild conditions, one may show that this mapping $\alpha \mapsto \tilde{\alpha}$ is a contraction at least on short horizon (i.e. for small T), see e.g. [22] for a discussion on a more general setting where the dynamics of X and P are allowed to be SDE.

*CMAP, Ecole Polytechnique, IP Paris, 91128 Palaiseau Cedex, France, kaitong.hu@polytechnique.edu.

†Laboratoire de Mathématiques d'Orsay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France, anna.kazeykina@math.u-psud.fr.

‡Université Paris-Dauphine, PSL Research University, CNRS, UMR [7534], Ceremade, 75016 Paris, France, ren@ceremade.dauphine.fr.

However, this algorithm has a major drawback, that is, the optimization on the top line is hard to solve in high dimension (unless in some special cases when the optimizers have analytic forms). That is why after the discovery of Pontryagin's maximum principle, people have studied and widely applied the related gradient descent algorithm, see e.g. [2,25]. In such iterative algorithm at each step we update the value of the control variable along the direction opposite to that of $\nabla_a H$, that is,

$$\alpha^{i+1} := \alpha^i - \eta \nabla_a H(t, X_t^i, a, P_t^i),$$

where η is the learning rate and X^i, P^i are the forward and backward processes evaluated with α^i. In order to look into the convergence of such iteration, let us consider the continuous version of this gradient descent algorithm, governed by the following system of ODEs on the infinite horizon:

$$\begin{cases} \frac{d\alpha_t^s}{ds} = -\nabla_a H(t, X_t^s, a_t^s, P_t^s) \quad \text{on} \quad \{s \geq 0\} \quad \text{for all} \quad t \in [0,T], \\ \\ X_t^s = x_0 + \int_0^t \phi(r, X_r^s, \alpha_r^s) dr, \\ P_t^s = \nabla_x G(X_T^s) + \int_0^t \nabla_x H(r, X_r^s, \alpha_r^s, P_r^s) dr. \end{cases} \tag{1.3}$$

Curiously, after some careful calculus, one may verify that

$$\frac{dV^0(\alpha^s)}{ds} = -\int_0^T \left| \nabla_a H(r, X_r^s, a_r^s, P_r^s) \right|^2 dr.$$

Therefore V^0 is a natural Lyapunov function for the process (α^s), and in order for the equality $\frac{dV^0(\alpha)}{ds} = 0$ to be true, the control α must satisfy the forward-backward system (1.2). This analysis (though not completely rigorous) reflects why this algorithm would converge. However, like other gradient-descent type algorithms, it would converge to a local minimizer, since Pontryagin's maximum principle is only a necessary first-order condition. One may attempt to put a convexity condition on the coefficients in order to ensure the local minimizer to be the global one. However, this usually urges X to be linear in α (so the function ϕ needs to be linear in (x, a)), which largely limits the application of this method.

In order to go beyond the convex case for the optimal control problem, it is natural to recall how the Langevin equation helps to approximate the solution of the non-convex optimization on the real space. Given a function F not necessarily convex, we know that under some mild conditions the unique invariant measure of the following Langevin equation

$$d\Theta_s = -\dot{F}(\Theta_s) ds + \sigma dW_s \tag{1.4}$$

is the global minimizer of the regularized optimization:

$$\min_{\nu \in \mathcal{P}} \int_{\mathbb{R}^m} F(a) \nu(da) + \frac{\sigma^2}{2} \mathsf{Ent}(\nu), \tag{1.5}$$

where W is the Brownian motion, $\mathcal{P}$ is the space of probability measures and the regularizer Ent is the relative entropy with respect to the Lebesgue measure, see e.g. [17]. Moreover, the marginal law of the process (1.4) converges to its invariant measure. As analyzed in the recent paper [16], this result is basically due to the fact that the function $\nu \mapsto \int F(a)\nu(da)$ is convex (indeed linear). In the present paper we wish to apply a similar regularization to the optimal control problem. In order to do that we first recall the relaxed formulation of the control problem (1.1). Instead of controlling the process α, we will control the flow of laws $(\nu_t)_{t \in [0,T]}$. Then the controlled process reads

$$X_t = x_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r, a) \nu_r(da) dr,$$

and we aim at minimizing

$$\inf_\nu V(\nu), \quad \text{where} \quad V(\nu) := \int_0^T \int L(t, X_t, a)\nu_t(da)dt + G(X_T).$$

Comparing it to the original control problem (1.1), we obtain that $\inf_\nu V(\nu) \le \inf_\alpha V^0(\alpha)$. Indeed, due to the classical results in [9, 10], under some mild conditions the values of the minimums of the two formulations remain the same. Further we add the relative entropy as a regularizer, as in (1.5), and focus on the regularized optimization:

$$\inf_\nu V^\sigma(\nu), \quad \text{where} \quad V^\sigma(\nu) := V(\nu) + \frac{\sigma^2}{2}\int_0^T \mathsf{Ent}(\nu_t)dt. \tag{1.6}$$

We will show that the global minimizer of this regularized control problem is again characterized by the invariant measure of Langevin-type dynamics, however, not a single Langevin equation as in (1.4), but a system of mean-field Langevin equations in the spirit of (1.3), that is,

$$\begin{cases} d\Theta_t^s = -\nabla_a H(t, X_t^s, \Theta_t^s, P_t^s)ds + \sigma dW_s, & \text{for } s \in \mathbb{R}^+, \quad \text{for } t \in [0,T], \quad \text{where} \\ X_t^s = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s, a)\nu_r^s(da)dr, & \text{with} \quad \nu_r^s := \mathrm{Law}(\Theta_r^s), \\ P_t^s = \nabla_x G(X_T^s) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s, a, P_t^s)\nu_r^s(da)dr, & \end{cases} \tag{1.7}$$

The name 'mean-field' reflects the fact that the different equations in the system are coupled through (and only through) the marginal laws $(\nu_t^s)_{t\in[0,T],s\in\mathbb{R}^+}$. Moreover, we shall show that this characterization holds true not only when V is convex in ν (which is still a quite restrictive case), but also under a set of milder conditions on the coefficients. Also, we prove in both cases that the marginal laws of the system (1.7) converge to its unique invariant measure. In particular, in the latter case we may quantitively compute the convergence rate.

One concrete motivation of this work is to shed some light on the solvability of the gradient descent method for the deep neural networks. Our work can be viewed as a natural extension to the recent works [16, 23, 24] in which the authors endorse the solvability of the two-layer (i.e. with one hidden layer) neural networks using the mean-field Langevin equations. It has been proposed in the recent papers [3, 4, 6, 21] among others, as well as in the course of P.-L. Lions in Collège de France (indeed similar ideas can be dated back to [19, 27], see also the very recent review on this topic [20]), that one may use the continuous-time optimal control problem as a model to study the deep neural networks. However, to our knowledge, there is no existing literature which succeeds in explaining why the stochastic gradient descent algorithm may approach the global optimum of the deep neural network under mild conditions. Our system of mean-field Langevin equations (1.7) and its relation to the regularized optimization (1.6) show a clear clue to how numerically compute the optimal control. Meanwhile, it is curious to observe that the standard discretization scheme (explicite Euler scheme) for the dynamics (1.7) is equivalent to the (noised) stochastic gradient descent algorithm for a class of deep neural networks, such as residual networks, convolutional networks, recurrent networks and so on.

The rest of the paper is organised in the following way. In Section 2, we define the relaxed optimal control problem under study and the corresponding system of mean-field Langevin equations. In Section 3 we announce the main results of the paper, namely, the wellposedness of the system of mean-field Langevin equations and the convergence of the marginal laws of the system towards the optimal control, both in the convex case and in a contraction case. Before giving detailed proofs for the theoretical results, we introduce the application to deep neural networks in Section 4. Then in Sections 5, 6 and 7 we provide the proofs of the main results.

2 Preliminaries

2.1 Regularized Relaxed Optimal Control

In this paper we aim to solve the optimal control problems in large dimension (in particular, the control variable is of large dimension). We shall allow the player to apply a mixed strategy, namely, a probability measure ν of which the marginal law on the time dimension is the Lebesgue measure, i.e.

$$\nu \in \mathcal{V} := \Big\{\nu \in \mathcal{M}([0,T]\times\mathbb{R}^m): \quad \nu(dt,da) = \nu_t(da)dt, \quad \text{for some} \;\; \nu_t \in \mathcal{P}(\mathbb{R}^m)\Big\},$$

where we denote by $\mathcal{M}$ the space of measures. The controlled process X reads:

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r, a, Z)\nu_r(\mathrm{d}a)\mathrm{d}r, \quad \text{for} \quad t \in [0,T],$$

where Z is an exogenous random variable taking values in a set $\mathcal{Z}$. In particular, in the application to the neural networks, Z would represent the input data. Denote $\mathcal{G} := \sigma(Z)$ and assume that X_0 is a bounded $\mathcal{G}$-measurable random variable. We use the notation E as the expectation of random variables on $\mathcal{G}$. The relaxed control problem writes:

$$\inf_\nu V(\nu), \quad \text{where} \quad V(\nu) := \mathsf{E}\left[\int_0^T \int_{\mathbb{R}^m} L(t, X_t, a, Z)\nu_t(\mathrm{d}a)\mathrm{d}t + G(X_T, Z)\right]. \tag{2.1}$$

Further in this paper, instead of addressing the optimal control problem itself, we introduce the following regularized version:

$$\inf_\nu V^\sigma(\nu), \quad \text{where} \quad V^\sigma(\nu) := V(\nu) + \frac{\sigma^2}{2}\int_0^T \mathsf{Ent}(\nu_t)dt, \tag{2.2}$$

where Ent is the relative entropy with respect to the Lebesgue measure on $\mathbb{R}^m$. It is noteworthy that $\int_0^T \mathsf{Ent}(\nu_t)dt$ is equal to the relative entropy of ν with respect to the Lebesgue measure on $[0,T]\times\mathbb{R}^m$.

2.2 System of Mean-Field Langevin Equations

The following remark establishes a link between the control problem (2.2) and the mean-field Langevin equation.

Remark 2.1. *Let us consider a simple example of a control problem with the following coefficients: $X_0 \equiv 0$, $L(a) = \lambda|a|^2$ and $\phi(t,x,a) = \hat{\phi}(a)$, that is, we aim to minimize*

$$\inf_\nu \; \mathsf{E}\Big[G\Big(\int_0^T \int_{\mathbb{R}^m} \hat{\phi}(a,Z)\nu_t(da)dt\Big)\Big] + \int_0^T \int_{\mathbb{R}^m} \lambda|a|^2\nu_t(da)dt + \frac{\sigma^2}{2}\int_0^T \mathsf{Ent}(\nu_t)dt.$$

Clearly, $(\nu_t)_{t\in[0,T]}$ are exchangeable, so the optimal control ν^* must satisfy $\nu_0^* = \nu_t^*$ for any $t \in [0,T]$. Therefore it is equivalent to minimize

$$\inf_{\nu_0} \; \mathsf{E}\Big[G\Big(T\int_{\mathbb{R}^m} \hat{\phi}(a,Z)\nu_0(da)\Big)\Big] + \int_{\mathbb{R}^m} \lambda|a|^2\nu_0(da) + \frac{\sigma^2 T}{2}\mathsf{Ent}(\nu_0).$$

Given a convex function G, this minimization problem is studied in the recent paper [16], where the authors prove that the marginal laws of the corresponding mean-field Langevin equation converge to the global minimizer. In the present paper we are going to generalize this result.

For the general control problem (2.2), we assume that all the coefficients are smooth enough. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and W an m-dimensional Brownian motion on it. Introduce the following system of mean-field Langevin equations:

$$\begin{cases} d\Theta_t^s = -\mathsf{E}\big[\nabla_a H(t, X_t^s, \Theta_t^s, P_t^s, Z)\big]ds + \sigma dW_s, & \text{for } s \in \mathbb{R}^+, \quad \text{for } t \in [0,T], \\ \text{where} \quad X_t^s = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^s, a, Z)\nu_r^s(da)dr, & \text{with} \quad \nu_r^s := \mathrm{Law}(\Theta_r^s), \\ \qquad\qquad P_t^s = \nabla_x G(X_T^s, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^s, a, P_t^s, Z)\nu_r^s(da)dr, & \end{cases} \tag{2.3}$$

and H is the Hamiltonian function:

$$H(t,x,a,p,z) := L(t,x,a,z) + p \cdot \phi(t,x,a,z) \qquad \text{for} \quad (t,x,a,p) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^d.$$

For the readers familiar with the variational calculus of optimal control (Pontryagin's maximum principle), we note that the process $(P_t^s)_{t\in[0,T], s\in\mathbb{R}^+}$ has an obvious link to the adjoint process in the maximum principle. This connection will be made clear in the discussion of Section 6. We are going to prove that under reasonable assumptions the system of mean-field Langevin equations has a unique solution, and the marginal distribution $(\nu_t^s)_{t\in[0,T]}$ converges to the global minimizer of the control problem (2.2) as $s \to \infty$.

2.3 Notations

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Denote by $\mathbb{E}$ the expectation under the probability $\mathbb{P}$, or roughly speaking, the expectation of the randomness produced by the Brownian motion W. In particular, note the difference between the notations $\mathbb{E}$ and E.

In the present paper we will use several different metrics on the measure space. First, recall the p-Wasserstein distance ($p \geq 1$) on the probability space $\mathcal{P}(\mathbb{R}^m)$:

$$\mathcal{W}_p(\mu_0, \nu_0)^p := \inf \left\{ \int_{\mathbb{R}^m} |x-y|^p \pi(dx, dy) : \quad \text{where } \pi \text{ is a coupling of } \mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^m) \right\}.$$

Further, for $\mu, \nu \in \mathcal{V}$ we define the metric

$$\overline{\mathcal{W}}_p^T(\mu, \nu) := \Big(\int_0^T \mathcal{W}_p(\mu_t, \nu_t)^p dt \Big)^{1/p}.$$

In some part of the paper, in particular during the discussion of the convex case (see Section 3.2 and 6), we shall use the following generalized p-Wasserstein distance on $\mathcal{V}$ (by abuse of notation, we still denote it by $\mathcal{W}_p$):

$$\mathcal{W}_p(\mu, \nu) := T^{1/p} \mathcal{W}_p\Big(\frac{\mu}{T}, \frac{\nu}{T}\Big) \quad \text{for } \mu, \nu \in \mathcal{V}. \tag{2.4}$$

Comparing the above definitions, we clearly see that $\mathcal{W}_p(\mu,\nu) \leq \overline{\mathcal{W}}_p^T(\mu,\nu)$ for any $\mu, \nu \in \mathcal{V}$.

In the proofs the constant C can vary from line to line. Without further specification C is always positive.

3 Main Results

In this section we announce the main results. Their proofs are given in Sections 5, 6 and 7.

Throughout the paper we assume that the Hamiltonian function H and the terminal cost G are continuously differentiable in the variables (x, a), and the coefficients

$$\phi, \nabla_x G, \nabla_x L, \nabla_x \phi \quad \text{exist and are all bounded.}$$

Therefore (X_t^s, P_t^s) lives in a compact set $\mathcal{K}_x \times \mathcal{K}_p$. From now on we treat $(t,x,a,p,z) \mapsto H(t,x,a,p,z)$ as a function defined on $[0,T] \times \mathcal{K}_x \times \mathbb{R}^m \times \mathcal{K}_p \times \mathcal{Z}$. In particular, whenever we claim that H satisfies a property (e.g. Lipschitz continuity) globally, it is meant to be true on this set instead of the whole space.

3.1 Wellposedness of the System of Mean-Field Langevin Equation

Assumption 3.1. *Assume that the coefficients ϕ, L, G are continuously differentiable in the variables (x, a) and*

- $\nabla_a \phi, \nabla_a L, \nabla_x \phi, \nabla_x L, \nabla_x G, \phi$ *are uniformly Lipschitz continuous in the variables (x, a, p);*
- *the coefficients $H, \nabla_a H$ satisfy*

$$\sup_{t,z} |H(t,0,0,0,z)| < \infty, \quad \sup_{t,z} |\nabla_a H(t,0,0,0,z)| < \infty. \tag{3.1}$$

Define the space of the continuous measure flows on the horizon $[0, S]$:

$$C_p\big([0,S], \mathcal{V}\big) := \left\{ \mu = (\mu^s)_{s\in[0,S]} : \quad \mu^s \in \mathcal{V} \ \text{ and } \ \lim_{s'\to s} \overline{\mathcal{W}}_p^T(\mu^{s'}, \mu^s) = 0 \ \text{ for all } \ s \in [0,S] \right\}.$$

Theorem 3.2. *Let Assumption 3.1 hold true. Given $(\Theta_t^0)_{t\in[0,T]}$ such that*

$$\int_0^T \mathbb{E}[|\Theta_t^0|^p] dt < \infty, \qquad \textit{for some} \quad p \geq 1, \tag{3.2}$$

the system of SDE (2.3) has a unique solution. In particular, the law of the solution $(\nu^s) \in C_p\big(\mathbb{R}^+, \mathcal{V}\big)$.

One of our main contributions is to observe the decrease of energy along the flow of the solution to the system of mean-field Langevin equation (2.3).

Assumption 3.3. *We further assume that*

- *the coefficients ϕ, L are second-order continuously differentiable in a;*
- *there is $\varepsilon > 0$ such that*

$$a \cdot \nabla_a H(t,x,a,p,z) \geq \varepsilon |a|^2, \quad \textit{for } |a| \textit{ big enough;} \tag{3.3}$$

- *for fixed (t, x, p) the mapping $a \mapsto \mathsf{E}\big[\nabla_a H(t,x,a,p,Z)\big]$ belongs to C^∞.*

Theorem 3.4 (Gradient flow)**.** *Let Assumptions 3.1 and 3.3 hold true, and assume that*

$$\int_0^T \mathbb{E}\big[|\Theta_t^0|^p\big] dt < \infty \quad \textit{for some} \quad p \geq 2. \tag{3.4}$$

Recall the function V^σ defined in (2.2). Let (ν_t^s) be the marginal laws of the solution to the system of mean-field Langevin equations (2.3). Then, for each $s > 0$, $t \in [0,T]$ the law ν_t^s admits a density, and for $s' > s > 0$ we have

$$V^\sigma(\nu^{s'}) - V^\sigma(\nu^s) = -\int_s^{s'} \int_0^T \int_{\mathbb{R}^m} \Big| \mathsf{E}\big[\nabla_a H(t, X_t^r, a, P_t^r, Z)\big] + \frac{\sigma^2}{2} \nabla_a \ln \nu_t^r(a) \Big|^2 \nu_t^r(a) da dt dr. \tag{3.5}$$

3.2 Convex Case

We first consider the case where the objective function V, defined in (2.1), is convex in ν. More precisely, we assume the following.

Assumption 3.5. *Let the controlled process X be linear in ν, i.e.*

$$X_t = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(s, a, Z)\nu_s(da)ds, \tag{3.6}$$

and the dependence on the variables (x, a) of the function L is seperated, i.e.

$$L(t, x, a, z) = \ell(t, x, z) + c(t, a, z).$$

Further we assume that

- *for all $t \in [0, T]$ the functions ℓ, G are convex in x;*
- *the functions $\phi, \nabla_x H, \nabla_a H$ are globally Lipschitz continuous in t;*
- *the Hamiltonian H is continuously differentiable in t and $\partial_t H$ is globally Lipschitz continuous in (t, a).*

Remark 3.6. *In the present section concerning the convex case we add the regularity assumptions on the coefficients with respect to the variable t. That is due to the fact that in this part of the paper we will apply the metric $\mathcal{W}_p$ (defined in (2.4)) on the space $\mathcal{V}$ instead of the usual one $\overline{\mathcal{W}}_p^T$.*

Under the above assumptions, it is clear that there exists at least one global minimizer of V^σ. Moreover, the function V is convex in ν, and thus V^σ is strictly convex in ν for any $\sigma > 0$, so there is one unique global minimizer. By standard variational calculus, we shall show the following sufficient condition for being the unique global minimizer of the control problem.

Theorem 3.7 (Sufficient first order condition)**.** *Let Assumption 3.5 hold true. If $\nu^* \in \mathcal{V}$, equivalent to the Lebesgue measure, satisfies*

$$\mathsf{E}\big[\nabla_a H(t, X_t^*, \cdot, P_t^*, Z)\big] + \frac{\sigma^2}{2}\nabla_a \ln\left(\nu_t^*\right) = 0 \tag{3.7}$$

for Leb-a.s. t, where X^* is the controlled process with the control ν^* as in (3.6) and P is the following adjoint process

$$P_t^* := \nabla_x G(X_T^*, Z) + \int_t^T \nabla_x \ell(r, X_r^*, Z)dr, \tag{3.8}$$

then ν^* is an optimal control of the regularized control problem (2.2).

Combining the sufficient condition above and Theorem 3.4, we can prove the following main result in the convex case.

Theorem 3.8. *Assume that Assumptions 3.1, 3.3 and 3.5 hold true and $(\Theta_t^0)_{t\in[0,T]}$ satisfies (3.4) with $p > 2$. Further assume that V is $\mathcal{W}_2$-continuous and bounded from below. Denote $(\nu_t^s)_{t\in[0,T]}^{s\in\mathbb{R}^+}$ the flow of marginal laws of the solution to (2.3). Then there exists an invariant measure of (2.3) equal to $(\nu_t^*)_{t\in[0,T]} := \operatorname{argmin}_\nu V^\sigma(\nu)$, and $(\nu_t^s)_{t\in[0,T]}$ converges to $(\nu_t^*)_{t\in[0,T]}$.*

3.3 Contraction Case

Clearly, the previous convex case has restrictive requirements on the structure of the coefficients. In particular, these requirements cannot be all satisfied in the application to the deep neural networks. That drives us to look for another setting in which the system of mean-field Langevin equations leads us to the optimal control of (2.2).

Proposition 3.9 (Necessary first order condition)**.** *Let Assumptions 3.1, 3.3 hold true and assume that the function V is bounded from below. Let ν^* be an optimal control of V^σ such that $V^\sigma(\nu^*) < \infty$. Then ν^* is equivalent to the Lebesgue measure and satisfies*

$$\mathsf{E}\big[\nabla_a H(t, X_t^*, \cdot, P_t^*, Z)\big] + \frac{\sigma^2}{2}\nabla_a \ln\big(\nu_t^*\big) = 0, \quad \text{for Leb-a.s. } t \in [0,T], \tag{3.9}$$

where (X^*, P^*) is the solution of the following ODE

$$X_t^* = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X_r^*, a, Z)\nu_r^*(da)dr, \tag{3.10}$$

$$P_t^* = \nabla_x G(X_T^*, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H(r, X_r^*, a, P_t^*, Z)\nu_r^*(da)dr. \tag{3.11}$$

In particular, for a.s. t, the probability measure ν_t^* admits a continuous density such that

$$-C(1+|a|^2) \le \ln\big(\nu_t^*(a)\big) \le C - C'|a|^2, \quad \text{for some } C \ge C' > 0 \text{ independent of } t, \tag{3.12}$$

and thus ν^* has finite p-moment for all $p \ge 1$.

Corollary 3.10. *Let Assumptions 3.1 and 3.3 hold true, and assume that V is $\overline{\mathcal{W}}_2$-continuous, bounded from below. Then any optimal control ν^* of V^σ, such that $V^\sigma(\nu^*) < \infty$, is an invariant measure of the system* (2.3)*.*

Assume that the control problem (2.2) admits at least one optimal control. The corollary above implies that once we ensure the convergence of the marginal laws of the system (2.3) towards the unique invariant measure, then the limit measure is an (indeed the unique) optimal control of (2.2).

Next we find a sufficient condition for the existence of the unique invariant measure for the system of mean-field Langevin equations. In particular, the convergent rate towards the limit measure is computed explicitly.

Assumption 3.11. *Assume that there exists a continuous function $\kappa : (0, +\infty) \to \mathbb{R}$ such that $\int_0^1 r\kappa(r)dr < +\infty$, $\overline{\lim}_{r\to+\infty} \kappa(r) < 0$ and for any (t, x, p, z) we have*

$$(a-\tilde{a})\cdot\Big(-\nabla_a H(t,x,a,p,z)+\nabla_a H(t,x,\tilde{a},p,z)\Big) \le \kappa\left(|a-\tilde{a}|\right)|a-\tilde{a}|^2 \quad \text{for all } a, \tilde{a} \in \mathbb{R}^m,\ a \ne \tilde{a}.$$

Theorem 3.12 (Contraction towards the unique invariant measure)**.** *Assume that Assumptions 3.1, 3.3 and 3.11 hold true. Let $(\Theta_t^0)_{t\in[0,T]}$, resp. $(\tilde{\Theta}_t^0)_{t\in[0,T]}$, satisfy* (3.2) *and denote ν^s, resp. $\tilde{\nu}^s$, the marginal law of the solution of the mean-field Langevin system* (2.3)*. Then there exist constants $c, \gamma \in (0,\infty)$ such that for any $s \ge 0$, we have*

$$\overline{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s) \le e^{\left(\frac{2\gamma T}{\varphi(R_1)} - c\sigma^2\right)s} \frac{2}{\varphi(R_1)} \overline{\mathcal{W}}_1^T(\nu^0, \tilde{\nu}^0). \tag{3.13}$$

The constants are explicitly given by

$$\varphi(r) = \exp\left(-\frac{1}{2}\int_0^r \frac{u\kappa^+(u)}{\sigma^2}du\right), \quad \Phi(r) = \int_0^r \varphi(s)ds,$$

$$c^{-1} = \int_0^{R_2} \Phi(s)\varphi(s)^{-1} ds \quad \textit{and} \quad \gamma = K^2(1+K)\exp(2KT),$$

where K is a common Lipschitz coefficient of $\nabla_a H$, $\nabla_x H$, $\nabla_x G$ and ϕ, and

$$R_1 := \inf\{R \geq 0 : \kappa(r) \leq 0 \textit{ for all } r \geq R\}$$

$$R_2 := \inf\{R \geq R_1 : \kappa(r)R(R - R_1) \leq -4\sigma^2 \textit{ for all } r \geq R\}.$$

In particular, if $\frac{2\gamma T}{\varphi(R_1)} < c\sigma^2$, there is a unique invariant measure with finite 1*-moment.*

Remark 3.13. *The result and the proof of Theorem 3.12 reveal the importance of considering the relaxed formulation of the control problem instead of the strict one* (1.1)*. As discussed in the introduction, in the setting of the strict formulation, one may let the control $(\alpha_t)_{t\in[0,T]}$ evolve along the gradient as in* (1.3)*. In this case the limit $\lim_{s\to\infty}(a_t^s)_{t\in[0,T]}$, if exists, in general depends on the initial value $(\alpha_t^0)_{t\in[0,T]}$, so it is unlikely to be the optimal control. On the contrary, Theorem 3.12 ensures that the gradient flow* (2.3) *for the relaxed formulation converges to a unique invariant measure independent of the initialization ν^0.*

In order to link the unique invariant measure to the optimal control of (2.2), we need an additional assumption.

Assumption 3.14. *Assume that*

- *the functional V is $\overline{\mathcal{W}}_2$-continuous and bounded from below,*
- *V^σ has at least an optimal control ν^* such that $V^\sigma(\nu^*) < \infty$.*

In particular, we will see a sufficient condition for the existence of optimal control in Lemma 7.1. Finally, combining the results in Corollary 3.10 and Theorem 3.12, we may conclude:

Corollary 3.15. *Let Assumption 3.1, 3.3, 3.11 and 3.14 hold true. Recall the constants defined in Theorem 3.12. If $\frac{2\gamma T}{\varphi(R_1)} < c\sigma^2$, then the unique invariant measure is the optimal control of* (2.2)*.*

4 Application to Deep Neural Networks

In this section, we apply the previous theoretical results, in particular the results of Section 3.3, to show the solvability of the stochastic gradient descent algorithm for optimizing the weights of a deep neural network.

Let the data Z take value in a compact subset $\mathcal{Z}$ of $\mathbb{R}^D$, and denote by $Y = f(Z)$ the label of the data. The function f is unknown, and we want to approximate it using the parametrized function generated by a deep neural network. Here we shall model the deep neural network using a controlled dynamic.

More precisely, consider the following choice of coefficients of the control problem (2.2): for $x \in \mathbb{R}^d, \beta \in \mathbb{R}^d \times \mathbb{R}^d, A \in \mathbb{R}^d \times \mathbb{R}^d, A' \in \mathbb{R}^d \times \mathbb{R}^D, k \in \mathbb{R}^d, a = (\beta, A, A', k) \in \mathbb{R}^m, z \in \mathcal{Z}$

$$\phi(x,a,z) := \ell(\beta)\varphi\big(\ell(A)x + A'z + k\big), \quad L(a) := \lambda|a|^2, \quad G(x,z) := g\big(Tf(z) - x\big) \tag{4.1}$$

where $\lambda > 0$ is a constant, φ is a nonlinear activation function, ℓ is a bounded truncation function and g is a cost function.

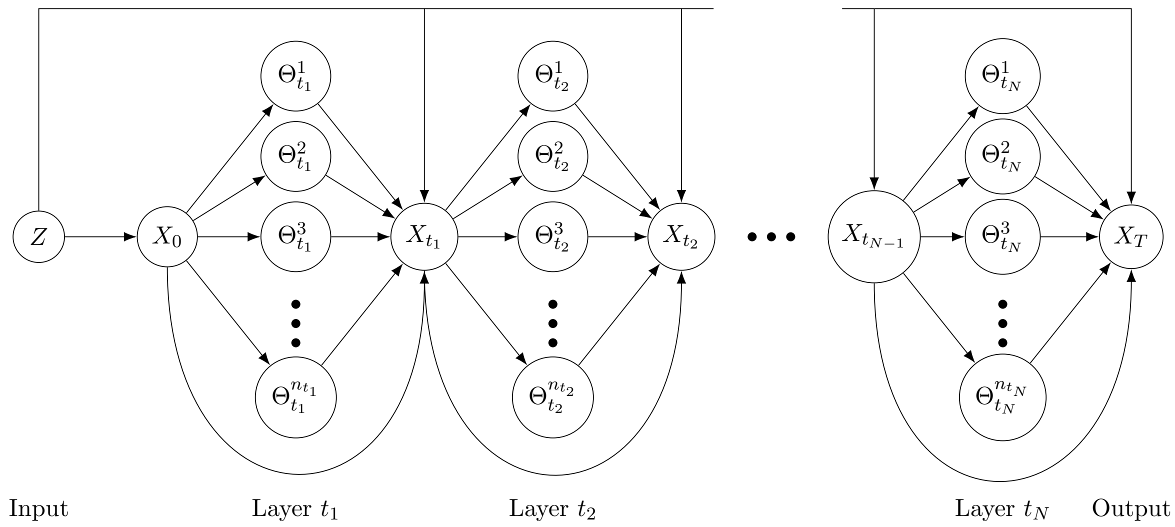
Assume that the coefficients satisfy all the assumptions needed for Corollary 3.15. Recall that the terminal value of the controlled process is equal to

$$X_T = X_0 + \int_0^T \int_{\mathbb{R}^m} \phi(X_t, a, Z)\nu_t(da)dt,$$

so it is a parametrized function of Z, where $(\nu_t)_{t\in[0,T]}$ is the parameter process. As we solve the regularized relaxed optimal control problem (2.2), we find the optimal parameter ν so as to minimize the (regularized) statistic error between the label Y and the output $\frac{1}{T}X_T$. Once we discretize the equation using the explicit Euler scheme and introduce random variables $(\Theta_t^j)_{j=1,\cdots,n_t}$, which are independent copies following the law ν_t, we find

$$X_{t_{i+1}} \approx X_{t_i} + \frac{\delta t}{n_{t_{i+1}}} \sum_{j=1}^{n_{t_{i+1}}} \phi(X_{t_i}, \Theta^j_{t_{i+1}}, Z), \quad \text{where } \delta t := t_{i+1} - t_i. \tag{4.2}$$

This discrete dynamics characterizes a type of structure of deep neural networks, and can be equivalently represented by the scheme of the figure below.



Neural network corresponding to the relaxed controlled process

This structure can describe a class of widely (and successfully) applied deep neural networks. Here are some examples:

- the process X can be viewed as the outputs of intermediate layers in a Residual Neural Network, see [12];
- (X_{t_i}) can also be interpreted as recurrent neurones in a Recurrent Neural Network or an LSTM Neural Network, see [14];
- once we take the function φ to be a convolutional-type activation function, the structure forms a Convolutional Neural Network with average pooling, see [18].

The most significant feature of this structure is the average pooling (averaging the outputs of the nonlinear activation φ as in (4.2)) on each layer, and it is due to the adoption of the relaxed formulation in our model of controlled process.

Given the structure of the neural network, or the scheme of the forward propagation (4.2), it is conventional to optimize the parameters $(\Theta^j_{t_i})$ using the stochastic gradient descent algorithm. The gradients of the parameters are easy to compute, due to the chain rule (or backward propagation):

$$\begin{cases} \Theta_t^{s_{j+1}} = \Theta_t^{s_j} - \delta s \mathsf{E}\big[\nabla_a H(t, X_t^{s_j}, \Theta_t^{s_j}, P_t^{s_j}, Z)\big] + \sigma \delta W_{s_j}, \text{ with } \delta s = s_{j+1} - s_j, \\ \text{where } P^s_{t_{i-1}} = P^s_{t_i} - \delta t \displaystyle\sum_{j=1}^{n_{t_{i+1}}} \nabla_x H\big(t_i, X^s_{t_i}, \Theta^{j,s}_{t_i}, P^s_{t_i}, Z\big), \quad P^s_T = \nabla_x G(X^s_T, Z), \end{cases} \tag{4.3}$$

where (δW_{s_j}) are independent copies of $\mathcal{N}(0, \delta s)$. In the conventional gradient descent algorithm σ is set to be 0, wheras we add a (small) positive volatility in our model of the regularized optimal control problem. It is important to observe that the continuous-time version of the noised gradient descent algorithm follows exactly the dynamics of the system of mean-field Langevin equations (2.3), where the horizon $s \in \mathbb{R}^+$ represents the iterations of gradient descents.

We remark that the evaluation of the parameters $\Theta_t^\cdot$ on a given layer t does not depend directly on the values of $\Theta_{t'}^\cdot$ on the other layers ($t' \neq t$), but only through the (empirical) law of $\Theta_{t'}^\cdot$. This 'mean-field' dependence between the parameters is due to the average pooling on each layer in this particular structure, and is the starting point of our theoretical investigation.

Recall that we showed in Section 3.3 that under a set of mild assumptions on the coefficients the marginal laws (ν^s) of (2.3) converge to the optimal control of (2.2). It approximately implies that the stochastic gradient descent algorithm converges to the global minimizer. One of the main insights provided by this theory is the quantitative convergence rate. In particular, the theory ensures the exponential convergence once the coefficients satisfy $\frac{\sigma^2}{T} > \frac{2\gamma}{c\varphi(R_1)}$. Hopefully, it could shed some light on how to tune the coefficients in practice.

Further it remains crucial to justify that the output $\frac{1}{T} X_T^*$ given by the optimal parameter ν^* is a good approximation to the label $Y = f(Z)$. In order for the contraction result to hold true we consider the horizon $T_\sigma := \frac{c\varphi(R_1)\sigma^2}{4\gamma}$. Assume that the cost function g is such that

$$|\zeta| \leq g(\zeta) \leq |\zeta| + \varepsilon\sigma^2$$

for some small constant $\varepsilon > 0$. Then we have

$$\mathsf{E}\Big|f(Z) - \frac{1}{T_\sigma} X_{T_\sigma}^*\Big| \leq \frac{1}{T_\sigma}\mathsf{E}\Big|T_\sigma f(Z) - X_{T_\sigma}^*\Big| \leq \frac{1}{T_\sigma} V^\sigma(\nu^*) = \frac{1}{T_\sigma}\inf_{\nu\in\mathcal{V}} V^\sigma(\nu). \tag{4.4}$$

Now consider the particular controls in the set $\mathcal{A} := \big\{\nu \in \mathcal{V} :\ \ell(A) = 0,\ \nu\text{-a.s.}\big\}$ where $\ell(A)$ is the coefficient in front of x in the activation function (see (4.1)). The optimization over the set $\mathcal{A}$ is equivalent to the optimization concerning the controlled process

$$d\tilde{X}_t = \int_{\mathbb{R}^m} \phi(0, a, Z)\nu_t(da)dt.$$

Together with the observation in Remark 2.1, we obtain

$$\begin{aligned}
\frac{1}{T_\sigma}\inf_{\nu\in\mathcal{V}} V^\sigma(\nu) &\leq \frac{1}{T_\sigma}\inf_{\nu\in\mathcal{A}} V^\sigma(\nu) \\
&\leq \frac{\varepsilon\sigma^2}{T_\sigma} + \frac{1}{T_\sigma}\inf_{\nu\in\mathcal{A}}\left(\mathsf{E}\Big|T_\sigma f(Z) - X_{T_\sigma}\Big| + \int_0^{T_\sigma}\int_{\mathbb{R}^m}\lambda|a|^2\nu_t(da)dt + \frac{\sigma^2}{2}\int_0^{T_\sigma}\mathsf{Ent}(\nu_t)dt\right) \\
&= \frac{4\gamma}{c\varphi(R_1)}\varepsilon + \inf_{\nu_0}\left(\mathsf{E}\Big|f(Z) - \int_{\mathbb{R}^m}\phi(0,a,Z)\nu_0(da)\Big| + \int_{\mathbb{R}^m}\lambda|a|^2\nu_0(da) + \frac{\sigma^2}{2}\mathsf{Ent}(\nu_0)\right) \\
&\longrightarrow \frac{4\gamma}{c\varphi(R_1)}\varepsilon + \inf_{\nu_0}\mathsf{E}\Big|f(Z) - \int_{\mathbb{R}^m}\phi(0,a,Z)\nu_0(da)\Big|, \qquad \text{as } \sigma, \lambda \to 0.
\end{aligned}$$

The last convergence is due to Proposition 2.3 of [16]. Together with (4.4) we have

$$\overline{\lim_{\sigma,\lambda\to 0}}\,\mathsf{E}\Big|f(Z) - \frac{1}{T_\sigma}X_{T_\sigma}^*\Big| \leq \frac{4\gamma}{c\varphi(R_1)}\varepsilon + \inf_{\nu_0}\mathsf{E}\Big|f(Z) - \int_{\mathbb{R}^m}\phi(0,a,Z)\nu_0(da)\Big|.$$

If one ignores the truncation function ℓ in (4.1), the universal representation theorem (see Theorem 1 of [15]) ensures that the value of the infimum on the right hand side is equal to 0. Therefore we have shown that $\frac{1}{T_\sigma}X_{T_\sigma}^*$ is an appropriate parametrized approximation for the label $f(Z)$.

5 Wellposedness of the system of mean-field Langevin equations

5.1 Wellposedness of the System

Proof of Theorem 3.2: Let $S > 0$. Given any process $(\mu^s)_{s\in[0,S]} \in C_p\big([0,S],\mathcal{V}\big)$, we can define for any $t \in [0,T]$ the process $(\Theta^s_t)_{s\in[0,S]}$ as the solution of the classical SDE:

$$\begin{cases} d\Theta^s_t = -\mathsf{E}\big[\nabla_a H(t, X^s_t(\mu), \Theta^s_t, P^s_t(\mu), Z)\big]ds + \sigma dW_s, & \text{for } s \in [0,S], \\ \\ \text{where} \quad X^s_t(\mu) = X_0 + \int_0^t \int_{\mathbb{R}^m} \phi(r, X^s_r(\mu), a, Z)\mu^s_r(da)dr, & \\ \qquad\qquad P^s_t(\mu) = \nabla_x G(X^s_T, Z) + \int_t^T \int_{\mathbb{R}^m} \nabla_x H\big(r, X^s_r(\mu), a, P^s_t(\mu), Z\big)\mu^s_r(da)dr, & \text{for } t \in [0,T], \end{cases} \tag{5.1}$$

We are going to show that the mapping $(\mu^s)_{s\in[0,S]} \mapsto \big(dt \times \mathrm{Law}(\Theta^s_t)\big)_{s\in[0,S]} =: (\nu^s)_{s\in[0,S]}$ is a contraction in the space $C_p\big([0,S],\mathcal{V}\big)$ for small enough S with the following metric:

$$d^{T,S}_p(\nu,\mu) := \sup_{s\le S} \overline{\mathcal{W}}^T_p(\nu^s,\mu^s),$$

and thus has a fixed point. Then the existence of a unique solution to the system (2.3) follows.

Step 1. First we will show the following property for the image of the mapping: $(\nu^s)_{s\in[0,S]} \in C_p\big([0,S],\mathcal{V}\big)$. It suffices to show that

$$\lim_{s'\to s} \int_0^T \mathbb{E}\big[|\Theta^{s'}_t - \Theta^s_t|^p\big]dt = 0 \quad \text{for all} \quad s \in [0,S]. \tag{5.2}$$

Since $\phi, \nabla_x H, \nabla_x G$ are all bounded, the processes $(X^s_t), (P^s_t)$ are both uniformly bounded. Further, by the assumption (3.1), the drift terms of the SDEs (5.1) are of linear growth uniformly in t. Then it follows from the standard estimate of SDE that

$$\mathbb{E}\big[\sup_{s\in[0,S]} |\Theta^s_t|^p\big] \le C\Big(\mathbb{E}\big[|\Theta^0_t|^p\big] + 1\Big) \quad \text{for all} \quad t \in [0,T],$$

where C is a constant independent of t. By the assumption (3.2), we have for any $s \in [0,S]$:

$$\int_0^T \mathbb{E}\Big[\sup_{s'\in[0,S]} |\Theta^{s'}_t - \Theta^s_t|^p\Big]dt < \infty.$$

Then (5.2) follows from the dominated convergence theorem.

Step 2. Let $(\mu^s)_{s\in[0,S]}, (\widetilde{\mu}^s)_{s\in[0,S]} \in C_p\big([0,S],\mathcal{V}\big)$ and denote $(\Theta^s_t)_{s\in[0,S]}$ and $(\widetilde{\Theta}^s_t)_{s\in[0,S]}$ the corresponding solution of the SDEs defined above, respectively. Denote $\nu^s_t := \mathrm{Law}(\Theta^s_t)$ and $\widetilde{\nu}^s_t := \mathrm{Law}(\widetilde{\Theta}^s_t)$. Denote K a common Lipschitz coefficient of $\nabla_a H$, $\nabla_x H$, $\nabla_x G$ and ϕ. We have

$$\begin{aligned} |\delta X^s_t| &:= |X^s_t(\mu) - X^s_t(\widetilde{\mu})| \le \int_0^t K\left(|X^s_r(\mu) - X^s_r(\widetilde{\mu})| + \mathcal{W}_1(\mu^s_r, \widetilde{\mu}^s_r)\right)\mathrm{d}r, \\ \text{and thus} \quad |\delta X^s_t| &\le K e^{Kt}\overline{\mathcal{W}}^t_1(\mu^s, \widetilde{\mu}^s) \le K e^{Kt}\overline{\mathcal{W}}^T_p(\mu^s, \widetilde{\mu}^s). \end{aligned} \tag{5.3}$$

Similarly, we obtain

$$\begin{aligned} |\delta P^s_t| &\le |\nabla_x G(X^s_T(\mu)) - \nabla_a G(X^s_T(\widetilde{\mu}))| + \int_0^t K\mathcal{W}_1(\mu^s_r, \widetilde{\mu}^s_r)dr + \int_0^t K\left(|\delta X^s_r| + |\delta P^s_r|\right)\mathrm{d}r \\ &\le (K^2 e^{KT} + K)\overline{\mathcal{W}}^T_p(\mu^s, \widetilde{\mu}^s) + K^2\left(\int_0^t e^{Kr}\mathrm{d}r\right)\overline{\mathcal{W}}^T_p(\mu^s, \widetilde{\mu}^s) + \int_0^t K|\delta P^s_r|\mathrm{d}r \\ &\le K(1+K)e^{KT}\overline{\mathcal{W}}^T_p(\mu^s, \widetilde{\mu}^s) + \int_0^t K|\delta P^s_r|\mathrm{d}r, \\ \text{and thus} \quad & |\delta P^s_t| \le K(1+K)e^{2KT}\overline{\mathcal{W}}^T_p(\mu^s, \widetilde{\mu}^s). \end{aligned} \tag{5.4}$$

Then define $\delta\Theta_t^s := \Theta_t^s - \widetilde{\Theta}_t^s$, and we can similarly estimate

$$\begin{aligned}\int_0^T |\delta\Theta_t^s|^p \mathrm{d}t &\le e^{Ks}K\int_0^T\int_0^s \Big(|\delta X_t^r|^p + |\delta P_t^r|^p\Big)\mathrm{d}r\mathrm{d}t \\ &\le K^2(1+K)Te^{Ks+2KT}\int_0^s \overline{\mathcal{W}}_p^T(\mu^r,\widetilde{\mu}^r)^p\mathrm{d}r.\end{aligned}$$

By taking the expectation on both sides we get

$$\overline{\mathcal{W}}_p^T(\nu^s,\widetilde{\nu}^s)^p \le K^2(1+K)STe^{KS+2KT}d_p^{T,S}(\mu,\widetilde{\mu})^p \qquad \text{for} \quad s\in[0,S].$$

Therefore, for S small enough, the mapping $(\mu^s)_{s\in[0,S]} \mapsto (\nu^s)_{s\in[0,S]}$ is a contraction. ■

Next we provide some useful estimates for the solution to the system (2.3).

Lemma 5.1. *Let Assumption 3.1 hold true. Let $(\Theta_t^0), (\tilde{\Theta}_t^0)$ be two initial values satisfying* (3.2)*, and denote by $(\nu_t^s), (\tilde{\nu}_t^s)$ the marginal laws of the solutions to the systems* (2.3)*, respectively. Then*

$$\overline{\mathcal{W}}_p^T(\nu^s,\tilde{\nu}^s) \le C\overline{\mathcal{W}}_p^T(\nu^0,\tilde{\nu}^0),$$

for some constant C possibly depending on s. Moreover, if we further assume that the functions $\phi, \nabla_x H$ are globally Lipschitz continuous in t, then we have

$$\mathcal{W}_p(\nu^s,\tilde{\nu}^s) \le C\mathcal{W}_p(\nu^0,\tilde{\nu}^0).$$

Proof This first result is a direct result of an elementary estimate of SDE. As for the second one, it is enough to note that under the additional assumption, for each $s\in\mathbb{R}^+$ the mappings $(t,a)\mapsto \phi(t,X_t^s,a,Z)$ and $(t,a)\mapsto \nabla_x H(t,X_t^s,a,P_t^s,Z)$ are both uniformly Lipschitz continuous, and thus

$$|X_t^s - \tilde{X}_t^s| \le C\mathcal{W}_1(\nu^s,\tilde{\nu}^s) \quad \text{as well as} \quad |P_t^s - \tilde{P}_t^s| \le C\mathcal{W}_1(\nu^s,\tilde{\nu}^s).$$

The rest follows again from the standard estimate of SDE. ■

Lemma 5.2. *Let Assumptions 3.1 and 3.3 hold true. Further assume that $(\Theta_t^0)_{t\in[0,T]}$ satisfies* (3.4)*. Then we have*

$$\int_0^T \mathbb{E}\Big[\sup_{s\in[0,S]}|\Theta_t^s|^p\Big]dt < \infty \quad \textit{for any } S\in\mathbb{R}^+, \tag{5.5}$$

$$\textit{as well as} \quad \int_0^T \sup_{s\in\mathbb{R}^+}\mathbb{E}\big[|\Theta_t^s|^p\big]dt < \infty. \tag{5.6}$$

Proof The result (5.5) follows from the standard SDE estimate, so its proof is omitted. By the Itô formula, we have

$$d|\Theta_t^s|^p = |\Theta_t^s|^{p-2}\Big(-p\Theta_t^s\cdot \mathsf{E}\big[\nabla_a H(t,X_t^s,\Theta_t^s,P_t^s,Z)\big] + \frac{\sigma^2}{2}p(p-1)\Big)ds + \sigma|\Theta_t^s|^{p-2}\Theta_t^s\cdot dW_s.$$

Now recall the assumptions (3.1) and (3.3) on $\nabla_a H$. We obtain:

$$\begin{aligned} d|\Theta_t^s|^p &\le |\Theta_t^s|^{p-2}\Big(C - \varepsilon|\Theta_t^s|^2 1_{\{|\Theta_t^s|\ge M\}}\Big)ds + \sigma|\Theta_t^s|^{p-2}\Theta_t^s\cdot dW_s \qquad \text{for some} \quad M>0,\\ &\le |\Theta_t^s|^{p-2}\Big((C+\varepsilon M^2) - \varepsilon|\Theta_t^s|^2\Big)ds + \sigma|\Theta_t^s|^{p-2}\Theta_t^s\cdot dW_s,\end{aligned}$$

where C does not depend on t. In the case $p=2$, it clearly leads to $\sup_{s\in\mathbb{R}^+}\mathbb{E}[|\Theta_t^s|^2] \le C(1+\mathbb{E}[|\Theta_t^0|^2])$, due to the Gronwall inequality. Then (5.6) follows. For general $p>2$, the result (5.6) is due to a simple induction. ■

Lemma 5.3. *Under the assumptions of Lemma 5.2, for each $t \in [0,T]$ the process $(X_t^s)_{s\in\mathbb{R}^+}$ in the system* (2.3) *is Lipschitz continuous in s.*

Proof Let $s' > s \geq 0$. Since the function ϕ is Lipschitz continuous in (x,a), we have

$$\begin{aligned}|\delta X_t| := |X_t^{s'} - X_t^s| &= \Big|\int_0^t \Big(\mathbb{E}\big[\phi(r, X_r^{s'}, \Theta_r^{s'}, Z)\big] - \mathbb{E}\big[\phi(r, X_r^s, \Theta_r^s, Z)\big]\Big)dr\Big| \\ &\leq \int_0^t \Big(C|\delta X_r| + \Big|\mathbb{E}\big[\phi(r, X_r^s, \Theta_r^{s'}, Z) - \phi(r, X_r^s, \Theta_r^s, Z)\big]\Big|\Big)dr.\end{aligned}$$

Further, by the Itô formula, we have

$$\begin{aligned}&\Big|\mathbb{E}\big[\phi(r, X_r^s, \Theta_r^{s'}, Z) - \phi(r, X_r^s, \Theta_r^s, Z)\big]\Big| \\ = &\Big|\mathbb{E}\Big[\int_s^{s'}\Big(-\nabla_a\phi(r, X_r^s, \Theta_r^u, Z)\cdot \mathsf{E}\big[\nabla_a H(r, X_r^u, \Theta_r^u, P_r^u, Z)\big] + \frac{\sigma^2}{2}\Delta_{aa}\phi(r, X_r^s, \Theta_r^u, Z)\Big)du\Big]\Big| \\ \leq\ & C(s'-s)\Big(1 + \sup_{u\in[s,s']}\mathbb{E}\big[|\Theta_r^u|\big]\Big).\end{aligned}$$

The last inequality is due to the boundedness of $\nabla_a\phi, \Delta_{aa}\phi$ and the uniform linear growth of $\nabla_a H$ in a. Finally, it follows from Lemma 5.2 and the Gronwall inequality that $|\delta X_t| \leq C(s'-s)$. ■

Given a solution to the system of mean-field Langevin equations (2.3), define

$$b^t(s,a) := -\mathsf{E}\big[\nabla_a H(t, X_t^s, a, P_t^s, Z)\big]. \tag{5.7}$$

It is easy to verify that under Assumptions 3.1 and 3.3, the function b^t is continuous in (s,a) and smooth in a for all $t \in [0,T]$. Due to a classical regularity result in the theory of linear PDEs (see e.g. [17, p.14-15]), we obtain the following result.

Lemma 5.4. *Let Assumptions 3.1 and 3.3 hold true. The marginal laws (ν_t^s) of the solution to* (2.3) *are weakly continuous solutions to the Fokker-Planck equations:*

$$\partial_s\nu = \nabla_a\cdot(-b^t\nu + \frac{\sigma^2}{2}\nabla_a\nu) \quad \textit{for} \quad t\in[0,T]. \tag{5.8}$$

In particular, we have that $(s,a)\mapsto \nu_t^s(a)$ belongs to $C^{1,2}\big((0,\infty)\times\mathbb{R}^m)\big)$.

Lemma 5.5. *For $t\in[0,T]$, if we assume that ν_t^0 admits a continuous density such that $-C(1+|a|^2) \leq \ln\big(\nu_t^0(a)\big) \leq C$ for some $C\geq 0$, then the solution $(s,a)\mapsto\nu_t^s(a)$ to* (5.8) *is continuous on $[0,\infty)\times\mathbb{R}^m$ and we have*

$$-\underline{C}(1+|a|^2) - \underline{\beta}s \leq \ln\big(\nu_t^s(a)\big) \leq \overline{C} + \overline{\beta}s \quad \textit{for some} \quad \underline{C},\overline{C}\geq 0 \textit{ and } \underline{\beta},\overline{\beta}\geq 0. \tag{5.9}$$

Proof We shall apply the comparison result of PDE. Under the upholding assumption we have $|b^t| \leq C_0(1+|a|)$ and $|\nabla_a\cdot b^t| \leq C_0$. On the one hand, it is easy to verify that $\overline{\nu}(s,a) := e^{\overline{C}+\overline{\beta}s}$, with $\overline{C} \geq \sup_a \ln\big(\nu_t^0(a)\big)$ and $\overline{\beta}\geq C_0$, is a supersolution of (5.8). On the other hand, define $\underline{\nu}(s,a) := e^{-\underline{C}(1+|a|^2)-\underline{\beta}s}$ with $\underline{\beta} \geq \frac{(C_0\underline{C})^2}{2\sigma^2\underline{C}^2 - 2C_0\underline{C}} + C_0 + \sigma^2\underline{C}$, and note that for $\underline{C} > \frac{C_0}{\sigma^2}$ and any $a\in\mathbb{R}^m$ we have

$$\begin{aligned}\nabla_a\cdot\left(b^t\underline{\nu} + \frac{\sigma^2}{2}\nabla_a\underline{\nu}\right) &\geq e^{-\underline{C}(1+|a|^2)-\underline{\beta}s}\big(-C_0 - 2C_0\underline{C}(1+|a|)|a| + 2\sigma^2\underline{C}^2|a|^2 - \sigma^2\underline{C}\big) \\ &\geq -e^{-\underline{C}(1+|a|^2)-\underline{\beta}s}\underline{\beta} = \partial_s\underline{\nu}.\end{aligned}$$

Therefore $\underline{\nu}$ is a subsolution to (5.8). Since we assume that $-\underline{C}(1+|x|^2) \leq \ln\nu_t^0(x) \leq \overline{C}$, it follows from the comparison result of the PDE (5.8) that $\underline{\nu}(s,a) \leq \nu_t^s(a) \leq \overline{\nu}(s,a)$ on $[0,\infty)\times\mathbb{R}^m$. ■

5.2 Gradient Flow

We first provide an estimate of the value $\nabla_a \ln(\nu_t^s)$. First, the following result ensures that $\ln(\nu_t^s)$ is well defined.

Lemma 5.6. *Assume that Assumption 3.1 and Assumption 3.3 hold true and $\nu_t^0 \in \mathcal{P}_2(\mathbb{R}^m)$ for some $t \in [0,T]$. Denote by $\mathbb{Q}_t^\sigma$ the scaled Wiener measure*[1] *with initial distribution ν_t^0 and by $(\mathcal{F}_s)_{s\in\mathbb{R}^+}$ the canonical filtration of the Wiener space. Then*

i) for any finite horizon $S > 0$, the law of the solution to (2.3), $\nu_t := \mathrm{Law}\big((\Theta_t^s)_{s\in\mathbb{R}^+}\big)$, is equivalent to $\mathbb{Q}_t^\sigma$ on $\mathcal{F}_S$ and the relative entropy

$$\int \ln\Big(\frac{\mathrm{d}\nu_t}{\mathrm{d}\mathbb{Q}_t^\sigma}\Big|_{\mathcal{F}_S}\Big) d\nu_t = \mathbb{E}\Big[\int_0^S \big|b^t(s,\Theta_t^s)\big|^2 ds\Big] < +\infty. \tag{5.10}$$

ii) the marginal law ν_t^s admits a density such that $\nu_t^s > 0$ and $\mathsf{Ent}(\nu_t^s) < +\infty$.

The proof of these results is based on the Girsanov theorem and some simple moment estimates. It is similar to the proof of Lemma 6.1 in [16] and thus omitted. Further we have the following regularity result.

Lemma 5.7. *For $t \in [0,T]$ and $(\nu_t^s)_{s\in\mathbb{R}^+}$ the marginal laws of the solution to (2.3), under the same assumptions as in Lemma 5.6, we have*

$$\nabla_a \ln(\nu_t^s(a)) = -\frac{1}{s_0}\mathbb{E}\left[\int_0^{s_0} \big(1 - r\nabla_a b^t(r,\Theta_t^{s-s_0+r})\big)\mathrm{d}W_r^{s-s_0}\Big|\Theta_t^s = a\right] \quad \text{for} \quad s_0 \in (0,s], \tag{5.11}$$

where $W_r^{s-s_0} := W_{s-s_0+r} - W_{s-s_0}$. In particular, for any $s > 0$ we have

$$C := \sup_{r\in[s,\infty)} \int_{\mathbb{R}^m} \big|\nabla \ln(\nu_t^r)\big|^2 \nu_t^r(a)\mathrm{d}a < +\infty,$$

and C only depends on the Lipschitz constant of $\nabla_a H$ with respect to a.

Proof The equality (5.11) is shown in Lemma 6.2 in [16]. The proof is based on Lemma 10.2 of the same paper and [11, Theorem 4.7 & Remark 4.13]. Further, we have for all $r \in [s,s']$:

$$\sup_{a\in\mathbb{R}^m} \big|\nabla_a \ln(\nu_t^r(a))\big|^2 \le \inf_{s_0\in(0,s)} \frac{1}{s_0^2}\mathbb{E}\Big[\int_0^{s_0} \big|1 - r\nabla_a b^t(r,\Theta_t^{s-s_0+r})\big|^2 \mathrm{d}r\Big].$$

Finally it is enough to note that $\nabla_a b^t$ is bounded under the assumptions of the present Lemma. ■

Lemma 5.8. *Assume that Assumption 3.1 and Assumption 3.3 hold true. We have*

$$\int_{\mathbb{R}^m} |\nabla_a \nu_t^s(a)|\mathrm{d}a < +\infty, \quad \int_{\mathbb{R}^m} |a\cdot\nabla_a \nu_t^s(a)|\mathrm{d}a < +\infty \quad \textit{for all} \quad s>0,$$
$$\textit{and} \quad \int_s^{s'}\int_{\mathbb{R}^m} |\Delta_{aa}\nu_t^r(a)|\mathrm{d}a\mathrm{d}r < +\infty \quad \textit{for all} \quad s' > s > 0.$$

Proof By the Young inequality, we have

$$|\nabla_a\nu_t^s(a)| \le \nu_t^s(a) + \left|\frac{\nabla_a\nu_t^s(a)}{\nu_t^s(a)}\right|^2 \nu_t^s(a) \quad \text{and} \quad |a\cdot\nabla_a\nu_t^s(a)| \le a^2\nu_t^s(a) + \left|\frac{\nabla_a\nu_t^s(a)}{\nu_t^s(a)}\right|^2 \nu_t^s(a).$$

[1] Let B be the canonical process of the Wiener space and $\mathbb{Q}$ be the Wiener measure, then the scaled Wiener measure $\mathbb{Q}^\sigma := \mathbb{Q}\circ(\sigma B)^{-1}$.

Since all the terms on the right hand sides are integrable, due to Lemma 5.7, therefore so are $\nabla_a \nu_t^s$ and $a \cdot \nabla_a \nu_t^s$. Next, in order to prove the integrability of $\Delta_{aa}\nu_t^s$, we apply Itô's formula:

$$\mathrm{d}\ln(\nu_t^s(\Theta_t^s)) = \left(\frac{\partial_t \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} + \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} \cdot b^t(s, \Theta_t^s) + \frac{\sigma^2}{2}\Delta_{aa}(\ln(\nu_t^s(\Theta_t^s)))\right)\mathrm{d}s + \sigma \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)}\mathrm{d}W_s.$$

Together with the Fokker-Planck equation (5.8), we have

$$\mathrm{d}\ln(\nu_t^s(\Theta_t^s)) = \left(\sigma^2 \frac{\Delta_{aa}\nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)} - \nabla_a \cdot b^t(s, \Theta_t^s) - \frac{\sigma^2}{2}\frac{|\nabla_a \nu_t^s(\Theta_t^s)|^2}{\nu_t^s(\Theta_t^s)}\right)\mathrm{d}s + \sigma \frac{\nabla_a \nu_t^s(\Theta_t^s)}{\nu_t^s(\Theta_t^s)}\mathrm{d}W_s. \tag{5.12}$$

By Lemma 5.7, we have $\mathbb{E}\left[\int_s^{s'} \frac{\nabla_a \nu_t^r(\Theta_t^r)}{\nu_t^r(\Theta_t^r)}\mathrm{d}W_r\right] = 0$. Also recall that $\nabla_a \cdot b^t(s, \Theta_t^s)$ is bounded. Taking expectation on both side of (5.12), we obtain

$$\begin{aligned}
\sigma^2 \int_s^{s'} \int_{\mathbb{R}^m} |\Delta_{aa}\nu_t^r(a)|\mathrm{d}a\mathrm{d}r &= \mathbb{E}\left[\int_s^{s'} \sigma^2 \frac{\Delta_{aa}\nu_t^r(\Theta_t^r)}{\nu_t^s(\Theta_t^s)}\mathrm{d}r\right] \\
&\le \mathsf{Ent}(\nu_t^{s'}) - \mathsf{Ent}(\nu_t^s) + C\mathbb{E}\left[\int_s^{s'}\left(1 + \frac{|\nabla_a \nu_t^r(\Theta_t^r)|^2}{\nu_t^r(\Theta_t^r)}\right)\mathrm{d}s\right].
\end{aligned}$$

By Lemma 5.6 and 5.7, the right hand side is finite. ■

Based on the previous integrability results, the next lemma follows from the integration by parts.

Lemma 5.9. *Under Assumption 3.1 and Assumption 3.3, we have for $s > 0$*

$$\begin{aligned}
\textstyle\int_{\mathbb{R}^m} \Delta_{aa}H(t, X_t^s, a, P_t^s, Z)\nu_t^s(a)\mathrm{d}a = -\int_{\mathbb{R}^m} \nabla_a H(t, X_t^s, a, P_t^s, Z) \cdot \nabla_a \nu_t^s(a)\mathrm{d}a \quad \textit{for all} \quad s > 0, \\
\textstyle\int_s^{s'}\int_{\mathbb{R}^m} \Delta_{aa}\big(\ln \nu_t^s(a)\big)\nu_t^r(a)\mathrm{d}a\mathrm{d}r = -\int_s^{s'}\int_{\mathbb{R}^m} |\nabla_a \ln \nu_t^s(a)|^2 \nu_t^s(a)\mathrm{d}a\mathrm{d}r \quad \textit{for all} \quad s' > s > 0.
\end{aligned}$$

Proof of Theorem 3.4. It follows from Lemma 5.3 that there exists a bounded process (U_t^s) such that $\mathrm{d}X_t^s = U_t^s \mathrm{d}s$. On the other hand, note that

$$X_t^s = X_0 + \int_0^t \mathbb{E}\big[\phi(r, X_r^s, \Theta_r^s, Z)\big]dr.$$

By the Itô formula, we get

$$\begin{aligned}
\frac{\mathrm{d}X_t^s}{\mathrm{d}s} = \int_0^t \int_{\mathbb{R}^m} \Big(&\nabla_a \phi(r, X_r^s, a, Z) b^r(s, X_r^s) \\
&+ \nabla_x \phi(r, X_r^s, a, Z) U_r^s + \frac{\sigma^2}{2}\Delta_{aa}\phi(r, X_r^s, a, Z)\Big)\nu_r^s(\mathrm{d}a)\mathrm{d}r,
\end{aligned}$$

where b^r, defined as in (5.7), is the drift term of the diffusion $(\Theta_r^s)_{s\in\mathbb{R}^+}$. In particular, we have

$$\frac{\mathrm{d}U_t^s}{\mathrm{d}t} = \int_{\mathbb{R}^m}\Big(\nabla_a \phi(t, X_t^s, a, Z) b^t(s, X_t^s) + \nabla_x \phi(t, X_t^s, a, Z) U_t^s + \frac{\sigma^2}{2}\Delta_{aa}\phi(t, X_t^s, a, Z)\Big)\nu_t^s(\mathrm{d}a) \tag{5.13}$$

Then note that

$$V(\nu^s) = \mathsf{E}\left[\int_0^T \mathbb{E}\big[L(t, X_t^s, \Theta_t^s, Z)\big]\mathrm{d}t + G(X_T^s, Z)\right].$$

Again by the Itô formula, we have

$$\begin{aligned}
\frac{\mathrm{d}V(\nu^s)}{\mathrm{d}s} &= \mathsf{E}\Big[\int_0^T \int_{\mathbb{R}^m}\Big(\nabla_a L(t, X_t^s, a, Z) \cdot b^t(s, a) + \frac{\sigma^2}{2}\Delta_{aa}L(t, X_t^s, a, Z) \\
&\qquad + \nabla_x L(t, X_t^s, a, Z) \cdot U_t^s\Big)\nu_t^s(\mathrm{d}a)\mathrm{d}t + \nabla_x G(X_T^s, Z) \cdot U_T^s\Big].
\end{aligned} \tag{5.14}$$

Recall (5.13) and the dynamic of $(P_t^s)_{t\in[0,T]}$ in (2.3). By integration by parts, we have

$$\begin{aligned}\nabla_x G(X_T^s,Z)\cdot U_T^s &= \int_0^T\int_{\mathbb{R}^m}\Big(-U_t^s\cdot\nabla_x H(t,X_t^s,a,P_t^s,Z)+P_t^s\cdot\nabla_x\phi(t,X_t^s,a,Z)U_t^s\\ &\qquad +P_t^s\cdot\nabla_a\phi(t,X_t^s,a,Z)b^t(s,a)+P_t^s\cdot\frac{\sigma^2}{2}\Delta_{aa}\phi(t,X_t^s,a,Z)\Big)\nu_t^s(\mathrm{d}a)\mathrm{d}t\\ &= \int_0^T\int_{\mathbb{R}^m}\Big(-U_t^s\cdot\nabla_x L(t,X_t^s,a,Z)\\ &\qquad +P_t^s\cdot\nabla_a\phi(t,X_t^s,a,Z)b^t(s,a)+P_t^s\cdot\frac{\sigma^2}{2}\Delta_{aa}\phi(t,X_t^s,a,Z)\Big)\nu_t^s(\mathrm{d}a)\mathrm{d}t.\end{aligned}$$

Together with (5.14), we obtain

$$\begin{aligned}\frac{\mathrm{d}V(\nu^s)}{\mathrm{d}s} &= \mathsf{E}\Big[\int_0^T\int_{\mathbb{R}^m}\Big(b^t(s,a)\cdot\nabla_a H(t,X_t^s,a,P_t^s,Z)+\frac{\sigma^2}{2}\Delta_{aa}H(t,X_t^s,a,P_t^s,Z)\Big)\nu_t^s(\mathrm{d}a)\mathrm{d}t\Big]\\ &= \int_0^T\int_{\mathbb{R}^m}\Big(-\big|b^t(s,a)\big|^2+\mathsf{E}\Big[\frac{\sigma^2}{2}\Delta_{aa}H(t,X_t^s,a,P_t^s,Z)\Big]\Big)\nu_t^s(\mathrm{d}a)\mathrm{d}t.\end{aligned}$$

Further by Lemma 5.9, we have for $s>0$

$$\frac{\mathrm{d}V(\nu^s)}{ds}=\int_0^T\int_{\mathbb{R}^m}\left(-\big|b^t(s,a)\big|^2+\frac{\sigma^2}{2}b^t(s,a)\cdot\nabla_a\ln\nu_t^s(a)\right)\nu_t^s(\mathrm{d}a)\mathrm{d}t. \tag{5.15}$$

On the other hand, by Itô formula, we get

$$\mathrm{d}\ln\nu_t^s(\Theta_t^s) = \Big(\nabla_a\ln\nu_t^s(\Theta_t^s)\cdot b^t(s,\Theta_t^s)+\frac{\sigma^2}{2}\Delta_{aa}\big(\ln\nu_t^s(\Theta_t^s)\big)\Big)\mathrm{d}s+\nabla_a\ln\nu_t^s(\Theta_t^s)\cdot\sigma dW_s.$$

It follows from Lemma 5.7 that the stochastic integral above is a martingale on any interval away from the origin. By taking expectation on both sides and applying Lemma 5.9, we obtain for any $s>0$:

$$\frac{\mathrm{d}\Big(\frac{\sigma^2}{2}\int_0^T\mathsf{Ent}(\nu_t^s)\mathrm{d}t\Big)}{\mathrm{d}s}=\int_0^T\int_{\mathbb{R}^m}\left(\frac{\sigma^2}{2}\nabla_a\ln\nu_t^s(a)\cdot b^t(s,a)-\frac{\sigma^4}{4}\big|\nabla_a\ln\nu_t^s(a)\big|^2\right)\nu_t^s(\mathrm{d}a)\mathrm{d}t. \tag{5.16}$$

Summing up (5.15) and (5.16), we finally obtain (3.5). ■

6 Proof for the Convex Case

6.1 Sufficient First Order Condition

We are going to apply a standard variational calculus argument in order to derive the sufficient condition for being the optimal control of (2.2).

Proof of Theorem 3.7. Take a $\nu\in\mathcal{V}$ such that ν is equivalent to the Lebesgue measure (otherwise $\int_0^T\mathsf{Ent}(\nu_t)dt=+\infty$), and thus equivalent to the measure ν^*. Denote X^* and X the controlled processes with ν^* and ν, respectively, and define $\delta X:=X-X^*$ and $\delta\nu:=\nu-\nu^*$. By the assumption on convexity of the coefficients, we have

$$\begin{aligned}\delta V &:= \mathsf{E}\left[\int_0^T\int_{\mathbb{R}^m}\Big(L(t,X_t,a,Z)\nu_t(da)-L(t,X_t^*,a,Z)\Big)\nu_t^*(da)dt+G(X_T,Z)-G(X_T^*,Z)\right]\\ &\ge \mathsf{E}\Big[\int_0^T\Big(\nabla_x\ell(t,X_t^*,Z)\cdot\delta X_t+\int_{\mathbb{R}^m}L(t,X_t^*,a,Z)\delta\nu_t(da)\Big)dt+\nabla_x G(X_T^*,Z)\cdot\delta X_T\Big]\end{aligned} \tag{6.1}$$

Recall the adjoint process P^* defined in (3.8). By integration by parts, we have

$$\nabla_x G(X_T^*) \cdot \delta X_T = P_T^* \cdot \delta X_T = \int_0^T \left(\int_{\mathbb{R}^m} P_s^* \cdot \phi(s,a) \delta\nu_s(da) - \nabla_x \ell(s, X_s^*, Z) \cdot \delta X_s \right) ds.$$

Together with (6.1), it leads to

$$\delta V \quad \geq \quad \mathsf{E}\left[\int_0^T \int_{\mathbb{R}^m} \Big(P_t^* \cdot \phi(t,a) + L(t, X_t^*, a, Z) \Big) \delta\nu_t(da) dt \right].$$

Further, we are going to compute the difference of the relative entropies. Since ν and ν^* are equivalent, we may define $f_t := \frac{\nu_t}{\nu_t^*}$. Denote $h(x) = x\ln(x)$ and note that $h(x) \geq x - 1$ for all $x \in \mathbb{R}^+$. We have

$$\begin{aligned}
\mathsf{Ent}(\nu_t) - \mathsf{Ent}(\nu_t^*) &= \int_{\mathbb{R}^d} \Big(\nu_t \ln \nu_t - \nu_t^* \ln \nu_t^* \Big) dx \\
&= \int_{\mathbb{R}^d} (\nu_t - \nu_t^*) \ln \nu_t^* \, dx + \int_{\mathbb{R}^d} \nu_t \Big(\ln \nu_t - \ln \nu_t^* \Big) dx \\
&= \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* \ln \nu_t^* \, dx + \int_{\mathbb{R}^d} h(f_t) \nu_t^* \, dx \\
&\geq \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* \ln \nu_t^* \, dx + \int_{\mathbb{R}^d} (f_t - 1) \nu_t^* \, dx \\
&= \int_{\mathbb{R}^m} \ln(\nu_t^*(a)) \delta\nu_t(da).
\end{aligned}$$

The last equality is due to $\int_{\mathbb{R}^d} (f_t - 1)\nu_t^* \, dx = \int_{\mathbb{R}^d} (\nu_t - \nu_t^*) \, dx = 0$. Finally, by (3.7) we have

$$\begin{aligned}
V^\sigma(\nu) - V^\sigma(\nu^*) \quad &\geq \quad \mathsf{E}\left[\int_0^T \int_{\mathbb{R}^m} \Big(P_t^* \cdot \phi(t,a) + L(t, X_t^*, a, Z) + \frac{\sigma^2}{2} \ln(\nu_t^*(a)) \Big) \delta\nu_t(da) dt \right] \\
&= \quad \mathsf{E}\left[\int_0^T \int_{\mathbb{R}^m} \Big(H(t, X_t^*, a, P_t^*, Z) + \frac{\sigma^2}{2} \ln(\nu_t^*(a)) \Big) \delta\nu_t(da) dt \right] = 0
\end{aligned}$$

■

6.2 Convergence Towards the Invariant Measure

In order to prove that there exists an invariant measure of (2.3) equal to the minimizer of V^σ, we follow the same strategy as in [16]. For readers' convenience, we shall provide a brief proof. The main ingredients of the proof are LaSalle's invariance principle (see e.g. [13, Theorem 4.3.3]) and the HWI inequality (see [26, Theorem 3]). Let $(\nu^s)_{s \in \mathbb{R}^+}$ be the flow of marginal laws of the solution of (2.3), given an initial law ν^0. Define a dynamic system $\mathcal{S}(s)\left[\nu^0\right] := \nu^s$. We shall consider the following ω-limit set:

$$\omega(\nu^0) := \left\{ \nu \in \mathcal{V} : \quad \text{there exists } s_n \to +\infty \text{ such that } \mathcal{W}_2\left(\mathcal{S}(s_n)\left[\nu^0\right], \nu \right) \to 0 \right\}.$$

Proposition 6.1 (Invariance Principle)**.** *Assume that Assumption 3.1 and Assumption 3.3 hold true and ν^0 satisfies (3.4) for some $p > 2$. Then the set $\omega(\nu^0)$ is non-empty, compact and invariant, that is*

i) for any $\nu \in \omega(\nu^0)$, we have $S(s)[\nu] \in \omega(\nu^0)$ for all $s \in \mathbb{R}^+$;

ii) for any $\nu \in \omega(\nu^0)$ and all $s \in \mathbb{R}^+$, there exists $\nu' \in \omega(\nu^0)$ such that $S(s)[\nu'] = \nu$.

Proof It is important to note that

- the mapping $\nu^0 \mapsto \mathcal{S}(s)\left[\nu^0\right]$ is $\mathcal{W}_2$-continuous, due to Lemma 5.1;
- the mapping $s \mapsto \mathcal{S}(s)\left[\nu^0\right]$ belongs to $C_2\big(\mathbb{R}^+, \mathcal{V}\big)$, due to Theorem 3.2;
- the set $\big\{\mathcal{S}(s)[\nu^0],\ s \in \mathbb{R}^+\big\}$ belongs to a $\mathcal{W}_2$-compact set, due to Lemma 5.2.

The rest follows the standard argument for LaSalle's invariance principle (see e.g. [13, Theorem 4.3.3] or [16, Proposition 6.5]). ■

Proof of Theorem 3.8. As in the Step 1 of the proof to [16, Theorem 2.10], using the invariance principle we can prove the existence of a convergent subsequence of the measure flow $(\nu^{s_n})_{n\in\mathbb{N}}$ such that $\mathcal{W}_2\big(\nu^{s_n}, \nu^*\big) \to 0$, where $\nu^* = \arg\min_\nu V^\sigma(\nu)$ and satisfies

$$\nu^*(t,a) = C \exp\Big(-\frac{2}{\sigma^2} H(t, X_t^*, a, P_t^*, Z)\Big).$$

In particular, ν^* is log-semiconcave, because one may easily verify that the gradient of the mapping $(t,a) \mapsto H(t, X_t^*, a, P_t^*, Z)$ is Lipschitz continuous. By the HWI inequality we have

$$\int \big(\ln \nu^{s_n} - \ln \nu^*\big)\nu^{s_n}(dt, da) \le \mathcal{W}_2\big(\nu^{s_n}, \nu^*\big) \left(\sqrt{\mathcal{I}_n} + C\mathcal{W}_2\big(\nu^{s_n}, \nu^*\big)\right),$$

where $\mathcal{I}_n$ is the relative Fisher information defined as

$$\begin{aligned} \mathcal{I}_n &:= \int \Big|\nabla_a \ln \nu^{s_n} - \nabla_a \ln \nu^*\Big|^2 \nu^{s_n}(dt, da) \\ &= \int \Big|\nabla_a \ln \nu^{s_n} + \frac{2}{\sigma^2}\nabla_a H(t, X_t^*, a, P_t^*, Z)\Big|^2 \nu^{s_n}(dt, da) \\ &\le 2\int \big|\nabla_a \ln \nu^{s_n}\big|^2 \nu^{s_n}(dt, da) + C\Big(1 + \int |a|^2 \nu^{s_n}(dt, da)\Big). \end{aligned}$$

Then it follows from Lemma 5.7 and 5.2 that $\sup_n \mathcal{I}_n < \infty$. Together with the fact that $\mathcal{W}_2\big(\nu^{s_n}, \nu^*\big) \to 0$, we have

$$\overline{\lim_{n\to\infty}}\, \mathsf{Ent}(\nu^{s_n}) - \mathsf{Ent}(\nu^*) = \overline{\lim_{n\to\infty}} \int \big(\ln \nu^{s_n} - \ln \nu^*\big)\nu^{s_n}(dt, da) \le 0.$$

Since Ent is $\mathcal{W}_2$-lower-semicontinuous, we have $\lim\limits_{n\to\infty} \mathsf{Ent}(\nu^{s_n}) = \mathsf{Ent}(\nu^*)$, and thus $\lim\limits_{s\to\infty} V^\sigma(\nu^s) = V^\sigma(\nu^*)$. Till now, we have proved that V^σ is a continuous Lyapunov function along the trajectory of (ν^s). Further we can conclude the proof using the standard argument (see e.g. Step 3 of the proof of Theorem 2.10 [16]).

7 Proofs for the Contraction Case

We first provide a sufficient condition for the regularized control problem (2.2) to have at least one optimal control.

Lemma 7.1. *Assume that there exists $\nu \in \mathcal{V}$ such that $V^\sigma(\nu) < \infty$, and that there is a function $U : \mathbb{R}^m \to \mathbb{R}$ such that $\int_{\mathbb{R}^m} e^{-U(a)} da < \infty$ and $\bar V(\nu) := V(\nu) - \frac{\sigma^2}{2}\int_0^T \int_{\mathbb{R}^m} U(a)\nu_t(da)dt$ is bounded from below and weakly lower-semicontinuous. Then* $\operatorname{argmin}_\nu V^\sigma(\nu) \neq \emptyset$.

Proof Let $\bar\nu \in \mathcal{V}$ such that $V^\sigma(\bar\nu) < \infty$. Denote $C_0 := \inf_\nu \bar V(\nu)$ and $\bar C := V^\sigma(\bar\nu) - C_0$. Recall that

$$V^\sigma = \bar V + \frac{\sigma^2}{2}\int_0^T \Big(\mathsf{Ent}(\nu_t) + \int_{\mathbb{R}^m} U(a)\nu_t(da)\Big)dt = \bar V + \frac{\sigma^2}{2} I(\nu),$$

$$\text{where} \qquad I(\nu) := \int_0^T \int_{\mathbb{R}^m} \ln\left(\frac{\nu_t(a)}{e^{-U(a)}}\right)\nu_t(da)dt.$$

Note that $I(\nu)$ is the relative entropy of ν with respect to the measure $dt \times e^{-U(a)}da$. Therefore, I is weakly lower-semicontinuous (so is V^σ) and the sublevel set $\mathcal{K} := \{\nu \in \mathcal{V} : \; I(\nu) \le \bar{C}\}$ is weakly compact (see e.g. [5, Lemma 1.4.3]). Further note that $\{\nu \in \mathcal{V} : \; V^\sigma(\nu) \le V^\sigma(\bar{\nu})\} \subset \mathcal{K}$, so

$$\inf_{\nu \in \mathcal{V}} V^\sigma(\nu) = \inf_{\nu \in \mathcal{K}} V^\sigma(\nu).$$

Since V^σ is weakly lower-semicontinuous and $\mathcal{K}$ is weakly compact, there exists a global minimizer in $\mathcal{K}$. ■

Next we prove the necessary condition of being an optimal control.

Proof of Proposition 3.9. Since $V^\sigma(\nu^*) < \infty$ and V is bounded from below, we know that $\int_0^T \mathsf{Ent}(\nu_t^*)dt < \infty$. In particular, ν^* is absolutely continuous with respect to the Lebesgue measure.

Step 1. Let $\nu \in \mathcal{V}$ be a measure such that $\int_0^T \mathsf{Ent}(\nu_t)dt < \infty$, in particular, it is absolutely continuous with respect to the Lebesgue measure. Define $\nu^\varepsilon := (1-\varepsilon)\nu^* + \varepsilon\nu$ for $\varepsilon > 0$. By standard variational calculus we have

$$\lim_{\varepsilon \to 0} \frac{V(\nu^\varepsilon) - V(\nu^*)}{\varepsilon} \quad = \quad \int_0^T \int_{\mathbb{R}^m} \mathsf{E}\big[H(t, X_t^*, a, P_t^*, Z)\big]\big(\nu(da) - \nu_t^*(da)\big)dt.$$

Further, define the function $h(x) := x \ln x$. We have

$$\frac{1}{\varepsilon}\int_0^T (\mathsf{Ent}(\nu_t^\varepsilon) - \mathsf{Ent}(\nu_t^*))\,dt \quad = \quad \frac{1}{\varepsilon}\int_0^T \int_{\mathbb{R}^m} \Big(h\left(\nu_t^\varepsilon(a)\right) - h\left(\nu_t^*(a)\right)\Big)dadt$$

Since the function h is convex, we note that

$$\frac{1}{\varepsilon}\big(h\left(\nu_t^\varepsilon(a)\right) - h\left(\nu_t^*(a)\right)\big) \le h(\nu_t(a)) - h(\nu_t^*(a)) \quad \text{for all} \quad \varepsilon \in (0,1).$$

The right hand side of the inequality above is integrable because both $\int_0^T \mathsf{Ent}(\nu_t)dt$ and $\int_0^T \mathsf{Ent}(\nu_t^*)dt$ are finite. Therefore, by Fatou's lemma we obtain

$$\begin{aligned} 0 \le \overline{\lim_{\varepsilon\to 0}} & \frac{V^\sigma(\nu^\varepsilon) - V^\sigma(\nu^*)}{\varepsilon} \\ & \le \int_0^T \int_{\mathbb{R}^m} \left(\mathsf{E}\big[H(t, X_t^*, a, P_t^*, Z)\big] + \frac{\sigma^2}{2}\ln\big(\nu_t^*(a)\big)\right)\big(\nu_t(da) - \nu_t^*(da)\big)dt. \quad (7.1) \end{aligned}$$

Step 2. We are going to show that for Leb-a.s. t

$$\Sigma_t(a) := \mathsf{E}\big[H(t, X_t^*, a, P_t^*, Z)\big] + \frac{\sigma^2}{2}\ln\big(\nu_t^*(a)\big) \quad \text{is equal to a constant } \nu_t^*\text{-a.s.} \qquad (7.2)$$

Define the mean value $\overline{c}_t := \int_{\mathbb{R}^m} \Sigma_t(a)\nu_t^*(da)$ and let $\varepsilon, \varepsilon' > 0$. Consider the measure $\nu \in \mathcal{V}$ absolutely continuous with respect to ν^* such that $\nu_t = \nu_t^*$ if $\nu_t^*[\Sigma_t \le \overline{c}_t - \varepsilon] < \varepsilon'$, otherwise

$$\frac{d\nu_t}{d\nu_t^*} = \frac{1_{\Sigma_t \le \overline{c}_t - \varepsilon}}{\nu_t^*[\Sigma_t \le \overline{c}_t - \varepsilon]}.$$

Note that $\Sigma_t \le \overline{c}_t - \varepsilon$, ν_t-a.s. for t such that $\nu_t^*[\Sigma_t \le \overline{c}_t - \varepsilon] \ge \varepsilon'$. Then we have

$$\int_0^T \int_{\mathbb{R}^m} \Sigma_t(a)\big(\nu_t(da) - \nu_t^*(da)\big)dt \le -\varepsilon \int_0^T 1_{\nu_t^*[\Sigma_t \le \overline{c}_t - \varepsilon] \ge \varepsilon'}dt.$$

Together with (7.1), we conclude

$$\nu_t^*[\Sigma_t \le \bar{c}_t - \varepsilon] < \varepsilon', \quad \text{for Leb-a.s. } t \in [0,T].$$

Since this holds true for arbitrary $\varepsilon', \varepsilon > 0$, we obtain (7.2).

Step 3. We are going to show that ν^* is equivalent to the Lebesgue measure. First we provide an estimate for the constant $\bar{c}_t$ above. Since ν_t^* is a probability measure, we have

$$\int_{\mathbb{R}^m} \exp\left(\frac{2\Big(\bar{c}_t - \mathsf{E}\big[H(t, X_t^*, a, P_t^*, Z)\big]\Big)}{\sigma^2}\right) da = 1. \tag{7.3}$$

Moreover, since $(t,z) \mapsto H(t,0,0,0,z)$ is bounded and $a \mapsto \nabla_a H(t,x,a,p,z)$ is uniformly Lipschitz continuous, we have

$$\sup_{t,z} \big|H(t, X_t^*, a, P_t^*, z)\big| \le C(1 + |a|^2). \tag{7.4}$$

On the other hand, following the dissipative assumption (3.3), we may easily prove that there are constants $C, C' > 0$ such that for all (t,a)

$$H(t, X_t^*, a, P_t^*, z) \ge -C + C'|a|^2. \tag{7.5}$$

Together with (7.3) and (7.4), we prove that $(\bar{c}_t)_{t\in[0,T]}$ is bounded.

Now suppose that ν^* is not equivalent to the Lebesgue measure. Then there is a set $\mathcal{K} \in [0,T] \times \mathbb{R}^m$ such that $\nu^*(\mathcal{K}) = 0$ (so $\ln \nu^* = -\infty$ on $\mathcal{K}$) and $\text{Leb}[\mathcal{K}] > 0$. It follows from (7.1) that

$$0 \le C - \int_{\mathcal{K}} \infty d\nu.$$

Since we may choose ν having positive mass on $\mathcal{K}$, it is a contradiction. Therefore ν^* must be equivalent to the Lebesgue measure.

Step 4. Since ν^* is equivalent to the Lebesgue measure, together with (7.2) we obtain (3.9). The bounds in (3.12) are obtained from the estimates (7.4) and (7.5). ■

We are going to show that Corollary 3.10 is a direct result of Theorem 3.4.

Proof of Corollary 3.10. Let $\nu^* \in \mathcal{V}$ be a local minimizer of V^σ. Denote by $(\nu^s)_{s\in\mathbb{R}^+}$ the marginal laws of the solution to the system of mean-field Langevin equations (2.3) given $\nu_0 = \nu^*$.

Step 1. We are going to prove that given $\nu_0 = \nu^*$ the process $s \mapsto V^\sigma(\nu^s)$ is continuous at 0. First note that under the assumptions of the present Corollary, (3.12) holds true. Together with Lemma 5.5, we obtain that the function $(s,a) \mapsto \nu_t^s(a)$ is continuous on $[0,\infty) \times \mathbb{R}^m$ and we have the estimate (5.9). Recall that by Proposition 3.9, the optimal control ν^* has finite 2-moment. So, by Lemma 5.2 we have (5.5) for $p = 2$. Then, by the dominated convergence theorem, we obtain

$$\lim_{s\to 0} \int_0^T \mathsf{Ent}(\nu_t^s) dt = \lim_{s\to 0} \int_0^T \mathbb{E}\big[\ln \nu_t^s(\Theta_t^s)\big] dt = \int_0^T \mathbb{E}\big[\ln \nu_t^*(\Theta_t^0)\big] dt = \int_0^T \mathsf{Ent}(\nu_t^*) dt.$$

Finally, since $s \mapsto V(\nu^s)$ is continuous, we conclude that $s \mapsto V^\sigma(\nu^s)$ is continuous at 0. In particular, the dynamics (3.5) holds true for $s' > s \ge 0$ (s no longer needs to be strictly positive).

Step 2. Since ν^* is a local minimizer of V^σ, there exists $s_0 > 0$ such that for all $s \in [0, s_0]$ we have $V^\sigma(\nu^s) \geq V^\sigma(\nu^*)$. Since $s \mapsto V^\sigma(\nu^s)$ is non-increasing on $[0, s_0]$ (due to the result of *Step 1*), we have $V^\sigma(\nu^s) = V^\sigma(\nu^*)$ for all $s \in [0, s_0]$. Due to Theorem 3.4, it leads to

$$
\begin{aligned}
&- b^t(s,a) + \frac{\sigma^2}{2}\nabla_a \ln\big(\nu_t^s(a)\big) \\
&= \mathsf{E}\big[\nabla_a H(t, X_t^s, a, P_t^s, Z)\big] + \frac{\sigma^2}{2}\nabla_a \ln\big(\nu_t^s(a)\big) = 0, \quad \text{for all } s \in (0, s_0], \quad \text{for } t\text{-a.s. in } [0,T],
\end{aligned}
$$

where we recall the function b^t defined as in (5.7). Moreover, for such $t \in [0,T]$, it follows from Lemma 5.4 that $(\nu_t^s)_{s>0}$ is a solution to the Fokker-Planck equation, so we have

$$
\partial_s \nu_t = \nabla_a \cdot (-b^t \nu_t + \frac{\sigma^2}{2}\nabla_a \nu_t) = \nabla_a \cdot \Big(\nu_t(-b^t + \frac{\sigma^2}{2}\nabla_a \ln \nu_t)\Big) = 0.
$$

Therefore, for t-a.s. in $[0,T]$, we have $\nu_t^s = \nu_t^{s'}$ for all $s, s' \in (0, s_0]$, and thus $\nu_t^s = \nu_t^*$ for all $s \in [0, s_0]$, i.e. ν^* is an invariant measure to the system of mean-field Langevin equations (2.3). ■

In order to prove the main Theorem 3.12, we are going to use some coupling technique. The main ingredient is the reflection coupling in Eberle [7]. For the mean-field system, we shall adopt the reflection-synchronous coupling similar to [8].

We fix a parameter $\varepsilon > 0$. Introduce the Lipschitz functions $\mathrm{rc} : \mathbb{R}^m \times \mathbb{R}^m \to [0,1]$ and $\mathrm{sc} : \mathbb{R}^m \times \mathbb{R}^m \to [0,1]$ satisfying

$$
\mathrm{sc}^2(x,y) + \mathrm{rc}^2(x,y) = 1, \quad \mathrm{rc}(x,y) = 1 \ \text{ for } |x-y| \geq \varepsilon, \quad \mathrm{rc}(x,y) = 0 \ \text{ for } |x-y| \leq \varepsilon/2.
$$

Let $(\nu_t^0)_{t\in[0,T]}$ and $(\tilde\nu_t^0)_{t\in[0,T]}$ be two initial measures, and $(W_s^1), (W_s^2)$ be two independent Brownian motions. For the given $(\nu_t^0), (\tilde\nu_t^0)$ we construct the drift coefficients $(b^t), (\tilde b^t)$ as in (5.7), respectively. Further, for a fixed[2] $t \in [0,T]$, define the coupling $\Sigma_t = (\Theta_t, \tilde\Theta_t)$ as the solution to the standard SDE

$$
\begin{aligned}
d\Theta_t^s &= b^t(s, \Theta_t^s)ds + \mathrm{rc}(\Sigma_t^s)\sigma dW_s^1 + \mathrm{sc}(\Sigma_t^s)\sigma dW_s^2, \\
d\tilde\Theta_t^s &= \tilde b^t(s, \tilde\Theta_t^s)ds + \mathrm{rc}(\Sigma_t^s)\big(\mathrm{Id} - 2e_s\langle e_s, \cdot\rangle\big)\sigma dW_s^1 + \mathrm{sc}(\Sigma_t^s)\sigma dW_s^2,
\end{aligned}
$$

where $e_s := \frac{\Theta_t^s - \tilde\Theta_t^s}{|\Theta_t^s - \tilde\Theta_t^s|}$ for $\Theta_t^s \neq \tilde\Theta_t^s$, otherwise $e_s := \hat e$ some arbitrary fixed unit vector in $\mathbb{R}^m$. Next, we construct a concave increasing function f as in the proof of [8, Theorem 2.3]. Let

$$
f(r) := \int_0^r \varphi(s) g(s \wedge R_2) ds, \quad \text{where} \quad g(r) := 1 - \frac{c}{2}\int_0^r \Phi(s)\varphi(s)^{-1} ds,
$$

and the function φ and the constant R_2 are defined as in the statement of Theorem 3.12. Note that by definition $\kappa^+(r) = 0$ for any $r \geq R_1$, so $(\varphi(r))_{r\geq R_1}$ is a constant and the function f is linear on $[R_2, \infty)$. Furthermore, f is twice continuously differentiable on $(0, R_2)$ and

$$
2\sigma^2 f''(r) = -r\kappa^+(r) f'(r) - c\sigma^2 \Phi(r) \leq -r\kappa^+(r) f'(r) - c\sigma^2 f(r) \leq 0, \tag{7.6}
$$

$$
r\varphi(R_1) \leq \Phi(r) \leq 2f(r) \leq 2\Phi(r) \leq 2r. \tag{7.7}
$$

We next prove that an inequality similar to (7.6) holds for $r \in (R_2, \infty)$. Recall that $(\varphi(r))_{r\geq R_1}$ is a constant and thus we have $\Phi(r) = \Phi(R_1) + \varphi(R_1)(r - R_1)$. Since $\Phi(R_1) \geq \varphi(R_1)R_1$, we have

$$
\frac{\Phi(r)}{\Phi(R_2)} = \frac{\Phi(R_1) - \varphi(R_1)R_1 + \varphi(R_1)r}{\Phi(R_1) - \varphi(R_1)R_1 + \varphi(R_1)R_2} \leq \frac{r}{R_2}, \quad \text{for } r \geq R_2. \tag{7.8}
$$

[2] We are not defining the coupling for the system of SDE's, but for a single SDE with the fixed label t.

Furthermore, it is easy to verify that

$$c^{-1} \geq \int_{R_1}^{R_2} \Phi(s)\varphi^{-1}(s)ds \geq \frac{\varphi^{-1}(R_1)\Phi(R_2)(R_2 - R_1)}{2}. \tag{7.9}$$

Also note that $g(R_2) = \frac{1}{2}$ due to the definition of c, and thus $f'(r) = \frac{\varphi(R_1)}{2}$ for $r \geq R_2$. Together with (7.8), (7.9) and the definition of R_2, we have

$$r\kappa(r)f'(r) \leq -2\sigma^2 \frac{r\varphi(R_1)}{(R_2 - R_1)R_2} \leq -2\sigma^2 \frac{\Phi(r)\varphi(R_1)}{(R_2 - R_1)\Phi(R_2)} \leq -c\sigma^2\Phi(r) \leq -c\sigma^2 f(r).$$

Since on (R_2, ∞) the function f is linear, i.e. $f'' = 0$, the inequality (7.6) holds true on (R_2, ∞).

Proof of Theorem 3.12. *Step 1.* We first use an argument similar to that of the proof of Theorem 2.3 in [8] to obtain some estimates concerning the coupling. As usual in the contraction result, we assume

$$\mathcal{W}_1(\nu_t^0, \tilde{\nu}_t^0) = \mathbb{E}\big[|\Theta_t^0 - \tilde{\Theta}_t^0|\big] \geq \mathbb{E}\Big[f\big(|\Theta_t^0 - \tilde{\Theta}_t^0|\big)\Big]. \tag{7.10}$$

The last inequality is due to (7.7). On the other hand, for all $s \geq 0$ we have

$$\mathcal{W}_1(\nu_t^s, \tilde{\nu}_t^s) \leq \mathbb{E}\big[|\Theta_t^s - \tilde{\Theta}_t^s|\big] \leq \frac{2}{\varphi(R_1)}\mathbb{E}\Big[f\big(|\Theta_t^s - \tilde{\Theta}_t^s|\big)\Big]. \tag{7.11}$$

Denote $\delta\Theta_t^s := \Theta_t^s - \tilde{\Theta}_t^s$. By the definition of the coupling above, we have

$$d\delta\Theta_t^s = \Big(b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s)\Big)ds + 2\mathrm{rc}(\Sigma_t^s)\sigma d\bar{W}_s,$$

where $\bar{W}_s := \int_0^s e_r \cdot dW_r^1$ is a one-dimensional Brownian motion. Denote $r_s := |\delta\Theta_t^s|$ and note that by the definition of rc we have $\mathrm{rc}(U_t^s) = 0$ whenever $r_s \leq \varepsilon/2$. Therefore, one may show that

$$dr_s = e_s \cdot \Big(b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s)\Big)ds + 2\mathrm{rc}(\Sigma_t^s)\sigma d\bar{W}_s.$$

Define $\mathcal{L}_s^x$ as the right-continuous local time of (r_s) and μ_f as the nonpositive measure representing the second derivative of f. Then it follows from the Itô-Tanaka formula that

$$\begin{aligned} & f(r_s) - f(r_0) \\ = & \int_0^s f'(r_u)e_u \cdot \big(b^t(u, \Theta_t^u) - \tilde{b}^t(u, \tilde{\Theta}_t^u)\big)du + \frac{1}{2}\int_{\mathbb{R}} \mathcal{L}_s^x \mu_f(dx) + M_s, \\ \leq & \int_0^s \Big(f'(r_u)e_u \cdot \big(b^t(u, \Theta_t^u) - \tilde{b}^t(u, \tilde{\Theta}_t^u)\big) + 2\mathrm{rc}(\Sigma_t^u)^2\sigma^2 f''(r_u)\Big)\,du + M_s, \end{aligned}$$

where $M_s := 2\int_0^s \mathrm{rc}(\Sigma_t^u)f'(r_u)\sigma d\bar{W}_u$ is a martingale, and the last inequality is due to the concavity of f. Now it is important to note that under the assumptions of the present Theorem we have

$$e_s \cdot \big(b^t(s, \Theta_t^s) - \tilde{b}^t(s, \tilde{\Theta}_t^s)\big) \leq 1_{\{r_s \geq \varepsilon\}} r_s\kappa(r_s) + 1_{\{r_s < \varepsilon\}}\gamma\varepsilon + \gamma\overline{\mathcal{W}}_1^T(\nu^s, \tilde{\nu}^s),$$

where $\gamma := K^2(1+K)\exp(2KT)$ with K a common Lipschitz coefficient of $\nabla_a H$, $\nabla_x H$, $\nabla_x G$ and ϕ, which is computed explicitly in the proof of Theorem 3.2. Further, since $f'' \leq 0$ and $\mathrm{rc}(\Sigma_t^s) = 1$ whenever $r_s \geq \varepsilon$ and $f' \leq 1$, we have

$$\begin{aligned} f(r_s) - f(r_0) \leq & \int_0^s \Big(1_{\{r_u \geq \varepsilon\}}\big(f'(r_u)r_u\kappa(r_u) + 2\sigma^2 f''(r_u)\big) + 1_{\{r_u < \varepsilon\}}\gamma\varepsilon + \gamma\overline{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u)\Big)du + M_s \\ \leq & \int_0^s \Big(-1_{\{r_u \geq \varepsilon\}}c\sigma^2 f(r_u) + 1_{\{r_u < \varepsilon\}}\gamma\varepsilon + \gamma\overline{\mathcal{W}}_1^T(\nu^u, \tilde{\nu}^u)\Big)du + M_s. \end{aligned}$$

The last inequality is due to (7.6). It clearly leads to

$$\mathbb{E}\Big[e^{c\sigma^2 s}f(r_s)-f(r_0)\Big]\le\gamma\int_0^s e^{c\sigma^2 u}\Big(\varepsilon+\overline{\mathcal{W}}_1^T(\nu^u,\tilde{\nu}^u)\Big)du.$$

Recall (7.10) and (7.11). Together with the estimate above, we obtain

$$\frac{\varphi(R_1)}{2}e^{c\sigma^2 s}\mathcal{W}_1(\nu_t^s,\tilde{\nu}_t^s)-\mathcal{W}_1(\nu_t^0,\tilde{\nu}_t^0)\le\gamma\int_0^s e^{c\sigma^2 u}\Big(\varepsilon+\overline{\mathcal{W}}_1^T(\nu^u,\tilde{\nu}^u)\Big)du. \tag{7.12}$$

Step 2. Since for each $t\in[0,T]$ one may obtain the estimate (7.12) through the previous coupling argument, we have

$$\frac{\varphi(R_1)}{2}e^{c\sigma^2 s}\overline{\mathcal{W}}_1^T(\nu^s,\tilde{\nu}^s)-\overline{\mathcal{W}}_1^T(\nu^0,\tilde{\nu}^0)\le\gamma T\int_0^s e^{c\sigma^2 u}\Big(\varepsilon+\overline{\mathcal{W}}_1^T(\nu^u,\tilde{\nu}^u)\Big)du.$$

By the Gronwall inequality, we have

$$\overline{\mathcal{W}}_1^T(\nu^s,\tilde{\nu}^s)\le e^{\left(\frac{2\gamma T}{\varphi(R_1)}-c\sigma^2\right)s}\frac{2}{\varphi(R_1)}\overline{\mathcal{W}}_1^T(\nu^0,\tilde{\nu}^0)+\frac{2\gamma T}{\varphi(R_1)}e^{\frac{2\gamma T}{\varphi(R_1)}s}\varepsilon.$$

This holds true for all $\varepsilon>0$, so finally we obtain (3.13). ■

References

[1] Boltyanskii, V.G., Gamkrelidze, R.V., Pontryagin, L.S.: The theory of optimal processes. I. The maximum principle. TRW Space Technology Labs, Los Angeles, California (1960)

[2] Bryson, A.E., Denham, W.F.: A steepest ascent method for solving optimum programming problems. Journal of Applied Mechanics **29**(2), 247 (1962)

[3] Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. arXiv:1806.07366 (2019)

[4] Cuchiero, C., Larsson, M., Teichmann, J.: Deep neural networks, generic universal interpolation, and controlled ODEs. arXiv:1908.07838 (2019)

[5] Dupuis, P., Ellis, R.S.: A Weak Convergence Approach to the Theory of Large Deviations. Wiley (1997)

[6] E, W., Han, J., Li, Q.: A mean-field optimal control formulation of deep learning. Res Math Sci **6**(10) (2019)

[7] Eberle, A.: Reflection couplings and contraction rates for diffusions. Probability Theory and Related Fields **166**(3-4), 851–886 (2016)

[8] Eberle, A., Guillin, A., Zimmer, R.: Quantitative Harris-type theorems for diffusions and McKean–Vlasov processes. Transactions of the American Mathematical Society **371**(10), 7135–7173 (2019)

[9] El Karoui, N., Nguyen, D., Jeanblanc-Picqué, M.: Compactification methods in the control of degenerate diffusions: existence of an optimal control. Stochastics **20**(3), 169–219 (1987)

[10] Fleming, W.H.: Generalized solutions in optimal stochastic control. Differential Games and Control theory II, Proceedings of 2nd Conference, Univ. of Rhode Island, Kingston, RI, 1976, Lect. Notes in Pure and Appl. Math. **30**, 147–165 (1977)

[11] Föllmer, H.: Time reversal on Wiener space. In: S.A. Albeverio, P. Blanchard, L. Streit (eds.) Stochastic Processes - Mathematics and Physics, pp. 119–129. Springer (1986)

[12] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016)

[13] Henry, D.: Geometric Theory of Semilinear Parabolic Equations. Springer (1981)

[14] Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997). DOI 10.1162/neco.1997.9.8.1735. URL `https://doi.org/10.1162/neco.1997.9.8.1735`

[15] Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Networks **4**(2), 251–257 (1991)

[16] Hu, K., Ren, Z., Siska, D., Szpruch, L.: Mean-field Langevin dynamics and energy landscape of neural networks. Preprint arXiv:1905.07769 (2019)

[17] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)

[18] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: F. Pereira, C.J.C. Burges, L. Bottou, K.Q. Weinberger (eds.) Advances in Neural Information Processing Systems 25, pp. 1097–1105. Curran Associates, Inc. (2012). URL `http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural`

[19] LeCun, Y., Touresky, D., Hinton, G., Sejnowski, T.: A theoretical framework for back-propagation. Proceedings of the 1988 connectionist models summer school **1**, 21–28 (1988)

[20] Liu, G.H., Theodorou, E.A.: Deep learning theory review: An optimal control and dynamical systems perspective. arXiv: 1908.10920 (2019)

[21] Liu, H., Markowich, P.: Selection dynamics for deep neural networks. arXiv:1905.09076 (2019)

[22] Ma, J., Wu, Z., Zhang, D., Zhang, J.: On well-posedness of forward-backward SDEs - a unified approach. The Annals of Applied Probability **25**(4), 2168–2214 (2015)

[23] Mei, S., Misiakiewicz, T., Montanari, A.: Mean-field theory of two-layers neural networks: dimension-free bounds and kernel limit. arXiv:1902.06015 (2019)

[24] Mei, S., Montanari, A., Nguyen, P.M.: A mean field view of the landscape of two-layer neural networks. Proceedings of the National Academy of Sciences **115**(33), E7665–E7671 (2018)

[25] Mitter, S.K.: Successive approximation methods for the solution of optimal control problems. Automatica **3**(3-4), 135–149 (1966)

[26] Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. Journal of Functional Analysis **173**, 361–400 (2000)

[27] Pearlmutter, B.A.: Gradient calculations for dynamic recurrent neural networks: A survey. IEEE Transactions on Neural networks **6**(5), 1212–1228 (1995)