



**HAL**  
open science

# Forward and Backward Stochastic Differential Equations with normal constraint in law

Philippe Briand, Pierre Cardaliaguet, Paul-Eric Chaudru de Raynal, Ying Hu

► **To cite this version:**

Philippe Briand, Pierre Cardaliaguet, Paul-Eric Chaudru de Raynal, Ying Hu. Forward and Backward Stochastic Differential Equations with normal constraint in law. *Stochastic Processes and their Applications*, 2020, 130 (12), pp.7021-7097. 10.1016/j.spa.2020.07.007 . hal-02053777

**HAL Id: hal-02053777**

**<https://hal.science/hal-02053777v1>**

Submitted on 1 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# FORWARD AND BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS WITH NORMAL CONSTRAINTS IN LAW

PHILIPPE BRIAND, PIERRE CARDALIAGUET, PAUL-ÉRIC CHAUDRU DE RAYNAL, AND YING HU

ABSTRACT. In this paper we investigate the well-posedness of backward or forward stochastic differential equations whose law is constrained to live in an a priori given (smooth enough) set and which is reflected along the corresponding “normal” vector. We also study the associated interacting particle system reflected in mean field and asymptotically described by such equations. The case of particles submitted to a common noise as well as the asymptotic system is studied in the forward case. Eventually, we connect the forward and backward stochastic differential equations with normal constraints in law with partial differential equations stated on the Wasserstein space and involving a Neumann condition in the forward case and an obstacle in the backward one.

## 1. Introduction.

In this paper, we are concerned with reflected (forward or backward) Stochastic Differential Equations (SDE) in the case where the constraint is on the law of the solution rather than on its paths. This kind of equations have been introduced in their backward form in [4] in the scalar case and when the constraint is of the form  $\int h d\mu \geq 0$  for some map  $h : \mathbb{R} \rightarrow \mathbb{R}$  satisfying suitable assumptions and where  $\mu$  denotes the law of the considering process. Such a system being reflected according to the mean of (a functional of) the process, the authors called it a *Mean Reflected Backward Stochastic Differential Equation* (MR BSDE). In [3], the authors studied the forward version (hence called MR SDE) in the same setting as well as its approximation by an appropriate interacting particle system and numerical schemes. In [20], weak solution to related forward equations with constraint<sup>1</sup> are built. In the same framework, let us also mention the work [6] where the approximation of MR BSDE by an interacting particle system is studied, [19] where MR BSDE with quadratic generator are investigated and [5] where MR SDE with jumps are considered.

The aim of this work consists in enlarging the results of [3, 4, 6] to the multi-dimensional case and for rather general constraint sets on the law in the backward and forward cases. Mean-field interacting particles counterpart of such systems are also investigated as well as the so called *common noise* setting. Eventually, we also aim at introducing the deterministic counterpart of such reflected stochastic system through Partial Differential Equation (PDE) stated on Wasserstein space of Neumann (in the forward case) or obstacle (in the backward case) type.

### 1.1. Outline of the paper.

**On SDE with normal constraint in law.** Given coefficients<sup>2</sup>  $(b, \sigma) : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^{n \times d}$ , a MR SDE is an equation of the following form

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s + K_t, & t \geq 0, \\ \mathbb{E}[h(X_t)] \geq 0, \quad \int_0^t \mathbb{E}[h(X_s)] dK_s = 0, & t \geq 0, \end{cases} \quad (1)$$

---

The second and fourth authors have been partially supported by the ANR project ANR-16-CE40-0015-01. For the third Author, this work has been partially supported by the ANR project ANR-15-IDEX-02 .

<sup>1</sup>Therein, the author consider SDE with constraint in law but without the Skorokhod condition.

<sup>2</sup>which could be assumed to be random, see the results below.

where  $B$  is a Brownian motion,  $h$  is a function from  $\mathbb{R}^n$  to  $\mathbb{R}$ , the law of the initial condition  $X_0$  is such that  $\mathbb{E}[h(X_0)] \geq 0$  and where, for the time being,  $n = d = 1$ .

When focusing on this forward system, there are several ways to understand the mean reflected SDE. One striking example lies into the equation satisfied by the law of the solution  $X$ , which turns out to be a reflected Fokker-Planck equation. In other words, solving the above system translates into searching for solution to the Skorohod problem stated on Partial Differential Equation of Fokker-Planck type. Indeed, let  $(X, K)$  be a solution of the above system (which, according to [3], under suitable assumptions, exists, is unique and where the process  $K$  is supposed to be a deterministic increasing process starting from 0). By Itô's formula, the law  $\mu_t = [X_t]$  of  $X_t$  satisfies, in the sense of distributions,

$$\begin{cases} d_t \mu_t(x) = \left\{ D_x(\mu_t(x)b(t, x)) + \frac{1}{2} D_x^2(\mu_t(x)a(t, x)) \right\} dt + D_x(\mu_t(x)) dK_t, \\ \int h d\mu_t \geq 0 \quad \int_0^t \left( \int h d\mu_s \right) dK_s = 0, \quad t \geq 0, \end{cases}$$

where  $a = \sigma^2$ . Note that, as this system is deterministic, the deterministic assumption assumed on the reflexion term  $K_t$  really makes sense.

From this perspective it is natural to look for a solution to the more general problem (allowing  $d$  and  $n$  to be greater than 1):

$$\begin{cases} d_t \mu_t(x) = \left\{ \operatorname{div}(\mu_t(x)b(t, x)) + \frac{1}{2} \sum_{i,j} D_{i,j}^2(\mu_t(x)a_{ij}(t, x)) \right\} dt + \operatorname{div}(\mu_t(x)D_\mu H(\mu_t)(x)) dK_t, \\ H(\mu_t) \geq 0 \quad \int_0^t H(\mu_s) dK_s = 0, \quad t \geq 0, \end{cases} \quad (2)$$

where  $H$  is a map from  $\mathcal{P}^2(\mathbb{R}^n)$  to  $\mathbb{R}$  (where we denote by  $\mathcal{P}^2(\mathbb{R}^n)$  the set of Borel probability measures on  $\mathbb{R}^n$  with finite second order moment) and  $D_\mu H$  denotes the Lions' derivative (cf. [12, 24] and the discussion at the end of the Introduction). Of course, when  $H(\mu) = \int_{\mathbb{R}^n} h(x) \mu(dx)$ ,  $D_\mu H(\mu)(x) = \nabla h(x)$ . The map  $D_\mu H(\mu)(\cdot)$  can be viewed as a gradient of  $H$  in the space  $\mathcal{P}(\mathbb{R}^n)$  (see [12, 24] and the discussion below), so that the “outward normal” to the set  $\mathcal{O} := \{\mu, H(\mu) > 0\}$  at a point  $\mu \in \partial\mathcal{O}$  is, at least formally,  $D_\mu H(\mu)(\cdot)$ . For this reason (2) can be understood as the Fokker-Planck equation with a *normal reflexion*. The probabilistic counterpart of (2) then writes as the following reflected SDE:

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s + \int_0^t D_\mu H([X_s])(X_s) dK_s, \quad t \geq 0, \\ H([X_t]) \geq 0, \quad \int_0^t H([X_s]) dK_s = 0, \quad t \geq 0. \end{cases} \quad (3)$$

Indeed, if one denotes by  $(X, K)$  a solution of the above system, by Itô's formula, the law  $\mu_t = [X_t]$  of  $X_t$  satisfies system (2) in the sense of distributions. The above problem is actually no longer a “mean reflected” SDE but a SDE with *normal constraint in law* since the constraint can now be written in a general form.

In comparison with [3], we are led to deal with general constraints and possibly multi-dimensional valued processes. Considering this more general setting explains why we now specify the direction of the reflection (which is done along the “outward normal”) while in [3] the reflection is oblique along the unitary vector. Indeed, when working in this framework, the explicit formula obtained in the aforementioned work for the process  $K$  is not valid anymore, so that we cannot “just add it” to ensure that the constraint is satisfied.

In (2) the Fokker-Planck equation is deterministic, which explains why the reflection term  $K$  is naturally deterministic. A further generalization of (2) consists in considering instead a *stochastic* Fokker-Planck equation reflected on the set  $\mathcal{O} = \{\mu, H(\mu) > 0\}$ . Given the coefficients  $(b, \sigma_0, \sigma_1) : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow$

$\mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathbb{R}^{n \times d}$  and a  $d$ -dimensional Brownian motion  $W$ , one may look for a solution to the system

$$\begin{cases} d_t \mu_t(x) = \left\{ \operatorname{div}(\mu_t(x)b(t, x)) + \frac{1}{2} \sum_{i,j} D_{i,j}^2(\mu_t(x)a_{ij}(t, x)) \right\} dt \\ \quad + \operatorname{div}(\mu_t(x)\sigma_1(t, x)dW_t) + \operatorname{div}(\mu_t(x)D_\mu H(\mu_t)(x)) dK_t, \\ H(\mu_t) \geq 0 \quad \int_0^t H(\mu_s) dK_s = 0, \quad t \geq 0, \end{cases} \quad (4)$$

where  $a = (\sigma_0\sigma_0^* + \sigma_1\sigma_1^*)$  and  $(K_s)$  is a continuous nondecreasing process adapted to the filtration  $\mathcal{F}^W$  associated with  $W$ . The way we added the stochastic perturbation in the above reflected stochastic Fokker-Planck equation relies on the probabilistic interpretation of this problem. Indeed, given a Brownian motion  $B$  (supposed to be independent of  $W$ ), the above system precisely gives the dynamic of the *conditional* law (conditioned by  $\mathcal{F}^W$ )  $\mu_t = [X_t|W]$  of the solution of the system

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma_0(s, X_s) dB_s + \int_0^t \sigma_1(s, X_s) dW_s + \int_0^t D_\mu H([X_s|W])(X_s) dK_s, \\ H([X_t|W]) \geq 0, \quad \int_0^t H([X_s|W]) dK_s = 0, \quad t \geq 0. \end{cases} \quad (5)$$

This is a reflected SDE with *normal constraint on its conditional law*. For reasons that will be clear in few lines (and relies on the mean field interacting particles counterpart of such a system), we call this case *the normal reflexion in law with a common noise*.

Let us now discuss the interpretation of (5) (or (3)) at the particles level. Consider the system  $(X_t^i)$  of  $N$  particles, where, for  $i = 1, \dots, N$ ,

$$\begin{cases} X_t^i = X_0^i + \int_0^t b(s, X_s^i) ds + \int_0^t \sigma_0(s, X_s^i) dB_s^i + \int_0^t \sigma_1(s, X_s^i) dW_s + \int_0^t D_\mu H(\mu_s^N)(X_s^i) dK_s^N, \\ \mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}, \quad H(\mu_t^N) \geq 0, \quad \int_0^t H(\mu_s^N) dK_s^N = 0, \quad t \geq 0. \end{cases} \quad (6)$$

Here the  $B^i$  and  $W$  are independent Brownian motions. The initial conditions of the particles  $X_0^i$  are i.i.d. random variable with law  $\mu_0$  and are independent of the  $B^i$  and  $W$ . Eventually,  $K^N$  is a continuous, nondecreasing process adapted to the filtration  $\mathcal{F}^N$  generated by the  $B^i$ ,  $X_0^i$  and  $W$ . Assuming that  $H(\mu_0) > 0$ , this system is, conditionally to  $\Omega_N = \{H(\mu_0^N) \geq 0\}$ , a classical reflected SDE in  $(\mathbb{R}^n)^N$ , with normal reflexion on the boundary of the constraint

$$\mathcal{O}_N = \left\{ (x_1, \dots, x_N) \in (\mathbb{R}^n)^N, H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right) > 0 \right\}.$$

Following [24], we understand  $W$  as a *common noise*, since it affects all the players. Note that we have chosen to describe the mean field limit in the common noise case, but it reduces to the case without common noise if we let  $\sigma_1 \equiv 0$ . It turns out that the limit system of (6) is nothing but our SDE with normal reflexion in its condition law (5): as the number of players tends to infinity, the reflection term  $K^N$  in (6) no longer depends on the position of the other particles, but only on their statistical distribution conditioned by the common noise.

Let us just emphasize that, above, the law of the initial condition is assumed to satisfy the constraint in order to work conditionally to the event  $\Omega_N$ . This assumption relies on the fact that  $H(\mu_0) \geq 0$  is not a sufficient condition to ensure that  $H(\mu_0^N) \geq 0$ . In order to remove such a strict inequality, one may either work conditionally to the event  $\Omega_{N,\eta} = \{H(\mu_0^N) \geq -\eta_N\}$  which, provided that  $(\eta_N)_{N \geq 0}$  is chosen appropriately, becomes of full measure as  $N$  tends to the infinity (so that the particle system solves the “original” Skorohod problem asymptotically) or either construct suitable initial conditions  $\tilde{X}_0^i$  whose empirical measure  $\tilde{\mu}_0^N$  satisfies (under suitable assumptions)  $H(\tilde{\mu}_0^N) \geq 0$  and which is such that  $\tilde{\mu}_0^N \rightarrow \mu$  as  $N \rightarrow \infty$ . In this last case, for any  $N \geq 1$ , system (6) with  $\tilde{X}_0^i$  instead of  $X_0^i$  is a solution of the Skorohod problem in mean field on  $\mathcal{O}_N$ . We refer to Sections 3.2 and 4.2 respectively for more details.

We eventually conclude this outline on SDE with normal constraint in law by giving another interpretation of (6) (or of (3) when  $\sigma_1 \equiv 0$ ) through the dynamic of the family of mapping  $(U_t)_{t \geq 0}$  from  $\mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  defined by  $\forall t_0 \geq 0, U_{t_0}(\mu_0) = G([X_T^{t_0, \mu_0} | W])$ , for some given  $G : \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$ . Enlarging naively the definition of viscosity solution on finite dimensional space, one can show that the dynamic of  $U$  is given (in the viscosity sense) by the the following (backward) Neumann problem in the set  $\mathcal{O} = \{\mu, H(\mu) > 0\}$ :

$$\left\{ \begin{array}{l} (i) \quad -\partial_t U(t, \mu) - \frac{1}{2} \int_{\mathbb{R}^n} \text{Tr}((a(t, y)) \partial_y D_\mu U(t, \mu)(y)) \mu(dy) \\ \quad - \int_{\mathbb{R}^n} D_\mu U(t, \mu)(y) \cdot b(t, y) \mu(dy) \\ \quad - \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} \text{Tr}(D_{\mu\mu}^2 U(t, \mu)(x, y) \sigma_1(t, x) \sigma_1^T(t, y)) \mu(dx) \mu(dy) = 0 \\ \quad \text{in } (0, T) \times \mathcal{O}, \\ (ii) \quad \int_{\mathbb{R}^n} D_\mu U(t, \mu)(y) \cdot D_\mu H(\mu)(y) \mu(dy) = 0 \text{ in } (0, T) \times \partial \mathcal{O}, \\ (iii) \quad U(T, \mu) = G(\mu) \text{ in } \mathcal{O}, \end{array} \right. \quad (7)$$

where  $a = \sigma_0 \sigma_0^* + \sigma_1 \sigma_1^*$ . Condition (ii) is exactly the Neumann boundary condition associated with the set  $\mathcal{O}$ . Conversely, any smooth solution to (7) can be written as  $\mathbb{E}[G([X_T^{t, \mu} | W])]$ .

**BSDE with normal constraint in law.** The above discussion leads to investigate whenever the original constrained in law problem stated on BSDE from [4] can be generalized. We here aim at showing that the generalization done in the case of forward SDE remains valid, up to the *common noise* setting (and under a slightly different set of assumptions, see the discussion in the next paragraph). Concerning this last framework, it seems that the results remain valid, provided additional assumptions (insuring suitable boundedness conditions on  $Z$ ) are satisfied. Nevertheless we do not explore it in order to short the current paper.

As underlined in the previous paragraph, we aim at extending the results of [4] in a multi-dimensional setting. Following (3), we are lead to investigate the problem

$$\left\{ \begin{array}{l} Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s + \int_t^T D_\mu H([Y_s])(Y_s) dK_s, \quad 0 \leq t \leq T, \\ H([Y_t]) \geq 0, \quad 0 \leq t \leq T, \quad \int_0^T H([Y_s]) dK_s = 0, \end{array} \right. \quad (8)$$

where  $\xi \in \mathbb{L}_2(\mathcal{F}_T)$ , where the processes  $Y, Z$  are respectively of  $n$  and  $n \times d$  dimension and where  $f : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \rightarrow \mathbb{R}^n$ . Here again,  $H$  is a map from  $\mathcal{P}^2(\mathbb{R}^n)$  to  $\mathbb{R}$  and  $D_\mu H$  denotes the Lions' derivative so that the "outward normal" to the set  $\mathcal{O} := \{\mu, H(\mu) > 0\}$  at a point  $\mu \in \partial \mathcal{O}$  is again, at least formally,  $D_\mu H(\mu)(\cdot)$ . For this reason (8) is now called a *BSDE with normal constraint in law*.

Also we aim at investigating the mean field counterpart of such a system, extending to the current framework the results obtained in [6]. Namely, we consider the interacting particle system

$$\left\{ \begin{array}{l} Y_t^i = \xi^i + \int_t^T f(s, Y_s^i, Z_s^{i,j}) ds - \int_t^T \sum_{j=1}^N Z_s^{i,j} dB_s^j + \int_t^T D_\mu H(\mu_s^N)(Y_s^i) dK_s^N, \\ \forall t \in [0, T] : \quad \mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}, \quad H(\mu_t^N) \geq 0, \quad \int_0^T H(\mu_s^N) dK_s^N = 0, \quad 1 \leq i \leq N, \end{array} \right. \quad (9)$$

where for each  $i, j$ ,  $Z_s^{i,j}$  is a  $n \times d$  matrix,  $\{B^i\}_{1 \leq i \leq N}$  are  $N$  independent  $d$ -dimensional Brownian motions and  $K^N$  is a continuous non decreasing process. Assuming that  $H(\mu_T) > 0$  this system is, conditionally to  $\Omega_N = \{H(\mu_T^N) \geq 0\}$ , a classical reflected BSDE in  $(\mathbb{R}^n)^N$ , with normal reflexion on the boundary of the constraint

$$\mathcal{O}_N = \left\{ (y_1, \dots, y_N) \in (\mathbb{R}^n)^N, H\left(\frac{1}{N} \sum_{i=1}^N \delta_{y_i}\right) > 0 \right\},$$

and (8) is precisely the asymptotic dynamic (as  $N \rightarrow +\infty$ ) of one of the particles in (9). As in the forward case, the strict inequality assumed on the law of the terminal condition comes from the the fact

that  $H(\mu_T) \geq 0$  is not a sufficient condition to ensure that  $H(\mu_T^N) \geq 0$ . We refer to the corresponding point in the above discussion on forward SDE with constraint in normal law and to Sections 3.2 and 4.2 respectively for more details to handle properly such a problem.

Eventually, one may wonder if, for the one dimensional case when  $Y$  is a real process and in a Markovian set up (*i.e.* when  $\xi = \phi(X_T)$  for some  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  and some diffusion process  $X$  evolving according to coefficients  $b, \sigma : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$  and to the scalar valued Brownian motion  $B$ ), the solution of the system (8) can be written in term of the position  $X$  and its law thanks to a *decoupling field*  $u : \mathbb{R}_+ \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  and if the dynamic of such a decoupling field is given thanks to an obstacle problem, as shown in [22] when the reflection is on the path. Assuming that the generator  $f$  is deterministic we show that such a decoupling field exists and, when  $f$  does not depend on the  $Z$  argument, we obtain that  $u$  solves, in the viscosity sense, the following obstacle problem on the Wasserstein space:

$$\begin{cases} \min \left\{ \{(\partial_t + \mathcal{L})u(t, x, \mu) + f(t, x, u(t, x, \mu))\}; \right. \\ \left. H(u(t, \cdot, \mu) \# \mu) \right\} = 0, \quad \text{on } [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n), \\ u(T, \cdot, \cdot) = \phi, \end{cases} \quad (10)$$

where for any probability measures  $\nu$  and measurable function  $\varphi$ ,  $\varphi \# \nu$  denotes the push-forward of the measure  $\nu$  by the map  $\varphi$  and where  $\mathcal{L}$  is given by, for all smooth  $\varphi : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$

$$\begin{aligned} \mathcal{L}\varphi(t, x, \mu) &= \frac{1}{2} \int_{\mathbb{R}^n} \text{Tr} [((\sigma\sigma^*)(t, y)) \partial_y D_\mu \varphi(t, x, \mu)(y)] \mu(dy) + \int_{\mathbb{R}^n} D_\mu \varphi(t, x, \mu)(y) \cdot b(t, y) \mu(dy) \\ &+ \frac{1}{2} \text{Tr} [((\sigma\sigma^*)(t, x)) D_x^2 \varphi(t, x, \mu)] + D_x \varphi(t, x, \mu) \cdot b(t, x) \end{aligned}$$

**Main results.** We now describe our main results. In all the cases (forward setting with and without common noise and backward setting), the existence part differs from the ones in [4] or [3]: we build the solution by a penalization technique inspired from [25, 26]. Also, and still in all the cases considered, our approach for the existence part heavily relies on suitable bi-Lipschitz property of the map  $H$  which is assumed to hold in a neighbourhood of the boundary of the constraint set.

Concerning the forward equations, we first prove in Theorem 6 that (1) is well posed when  $K_t = \int_0^t \nabla h(X_s) d|K|_s$  for stochastic Lipschitz in space coefficients bounded in  $\mathbb{L}_2$  at point 0 uniformly in time. We start by handling this specific case for the following reasons: firstly, this allows to introduce the main tools needed for our analysis in a simple setting; secondly, the proof is done under some concavity assumption assumed on  $H(\mu) = \int h d\mu$  (which will be no longer assumed for the rest of the forward part) that allows to work with diffusion coefficients not necessarily globally bounded in space (and then enables to sketch what will be useful in the backward setting); thirdly, this allows to weaken a little bit the bi-Lipschitz assumption on the constraint function  $H$  (through the map  $h$ ) and finally, thanks to a suitable transformation, this allows to connect the results obtained in this case of normal reflection with the previous results of [3] where the reflection is oblique.

We then enter our framework and we prove in Theorem 8 and Theorem 11 the well-posedness of (3) and (5) respectively: we show that, under suitable assumptions on the data, there exists a unique solution to these equations. We emphasize that in both cases the uniqueness is shown to hold pathwise and in law. We present both results separately because they involve different sets of assumptions: for instance, for (3), we can allow the diffusion  $b$  to be unbounded while it is not the case for problems with common noise. The case with common noise also requires stronger assumptions on the constraint  $H$ . One reason for this is that, as can be seen on the Kolmogorov equations, (2) (which is associated with (3)) is a *deterministic* reflexion problem, while (4) (associated with (5)) involves a diffusion term. As a consequence, for (3), the process  $(K_t)$  is deterministic and Lipschitz continuous, while it is random and merely continuous for (5).

Let us again emphasize that in a recent paper [20], Jabir considers process very similar to (3) (*i.e.* satisfying same type of constraint but not enunciated or either identified as a solution of Skorohod problem on the law). Using a different penalization approach and working under weaker regularity assumptions,

provided that the constraint set is convex and the diffusion does not degenerate, he builds weak solutions to the problem through tightness of the law of the penalization. However no uniqueness result is stated in [20] and propagation of chaos as well as the common noise setting are not discussed.

Next, we present in Theorem 17 a conditional propagation of chaos for the particle system (6). In other words, we show that system (6) converges, as  $N \rightarrow +\infty$ , to the reflected SDEs with common noise (5). Let us recall that [3] already studied the propagation of chaos in the one dimensional setting. The main differences with [3] is that we work here with a normal reflexion (in [3] it was oblique) and in a common noise setup. The convergence rate are also explicitly given.

Eventually, we make the connection with the Neumann problem on the Wasserstein space. This connection remains at a basic level, as we only prove is that the map  $U$  defined as  $U(t_0, [X_0]) = \mathbb{E}[G([X_T|W])]$  is a viscosity solution of (7) in Theorem 20 and that any smooth solution of (7) can be written as  $U(t_0, [X_0]) = \mathbb{E}[G([X_T|W])]$  in Proposition 21. Here, the notion of viscosity solution we adopt consists in natural extension of the finite dimensional definition by testing the map  $U$  against smooth functions from  $\mathcal{P}_2(\mathbb{R}^n)$  to  $\mathbb{R}$ , the smoothness being understood in the Lions's sense. In that respect, we follow the definition proposed in [13] and extend it to the case of Neumann condition. Nevertheless, we do not prove a uniqueness result which, in infinite dimensional metric space, appears to be involved and definitely out of the scope of this paper. Let us just mention that, in that case, it seems to us that the map  $U$  could be shown to be the unique classical solution of (7), avoiding then the difficulties relying on the notion of viscosity solution. Again, such an investigation exceeds the scope of this paper.

Concerning the Backward equation, we first give in Theorem 24 an existence result for (8). This result relies on the concavity property assumed on  $H$  in the current setting and is done through a three steps scheme: we first tackle the case of deterministic bounded and space homogeneous driver; we then extend the result to a stochastic generator in  $L^2(\Omega, L^2([0, T])$  and eventually to a space dependent generator. As in the forward case without common noise, the process  $(K_t)$  is shown to enjoy a Lipschitz property. Again, uniqueness holds pathwisely and in law.

Then we construct in Section 4.2 an interacting particle system satisfying a Skorohod problem in mean field of type (9) and prove in Theorem 32 that the interacting particle system converges to (8) (meaning that (8) precisely gives the asymptotic dynamic of one of the interacting particles). Again, we specify the rate at which system (9) converges to (8).

Finally, thanks to our well-posedness result, we investigate the obstacle problem (10) through (decoupled) Forward-Backward SDE (FBSDE) with normal constraint in law (on the backward part). When the driver  $f$  no longer depends on the  $Z$ -variable, we show in Theorem 36 that the FBSDE with normal constraint in law admits a decoupling field  $u$  which satisfies, in the viscosity sense, the PDE (10). The reason why the driver is not allowed to depend on the variable  $Z$  relies to the lack of comparison principle when considering non-linear equations which makes the usual procedure to tackle this term unavailable. For the same reasons as pointed out for the Neumann problem, we do not prove any uniqueness result of viscosity solutions and let it for future considerations.

**1.2. Organization of this work.** The paper is organized in the following way. We complete the introduction by describing our notations, discussing the notion of Lions's derivative and presenting the various Itô's formulas used in the text as well as the tools needed to ensure the convergence of the particle systems. Section 2 is devoted to the SDEs with normal reflexion in law: starting with the simplest setting (1) to illustrate the method of proof, we then show the existence and the uniqueness of the solution of (3) (Theorem 8). SDEs with normal reflexion in conditional law (i.e., the "common noise" case) are studied in Section 3, where we prove the well-posedness of (5) (Theorem 11), the propagation of chaos (Theorem 17) and make the link with the Neumann problem (Theorem 20 and Proposition 21). Finally, the case of backward equation is treated in Section 4, where we also prove the existence and the uniqueness of the solution to (8) (Theorem 24), the propagation of chaos (Theorem 32) and a Feynman-Kac formula (Proposition 36).

**1.3. Notations and Mathematical tools.** Throughout the paper we will make an intensive use of Lions's derivative, introduced in [24] and later discussed in [8, 11, 12, 14]. Let us recall that other notions of derivatives in the space of measure have been developed in the literature: see for instance [1, 23, 27].

Related to Lions's derivatives, we also need several Itô's formulas, allowing to compute the value of a map depending on the law (or the conditional law) of a process. These formulas have been developed in various stages of generality in [8, 11, 12, 14]. Nevertheless, it is not yet clearly stated that this formula holds for reflected processes, especially when they are submitted to a common noise as it is the case in (5). Therefore, we prove below that such Itô's formulas can be generalized to our framework. Eventually, in order to ensure the convergence of our particle system (6) and (9), we provide at the end of this part a Lemma allowing to control the uniform convergence in 2-Wasserstein metric of an empirical measure of conditionally i.i.d. processes.

**Notations.** Throughout the paper, we work on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with a complete and right-continuous filtration  $\{\mathcal{F}_t\}_{t \geq 0}$ . We denote by  $\mathcal{E}$  the  $\sigma$ -algebra of progressive sets of  $\Omega \times \mathbb{R}_+$  and let  $B$  and  $W$  be two independent  $d$ -dimensional Brownian motion with respect to this filtration. We denote by  $\mathcal{F}^W$  the completed filtration generated by  $W$  and, given a random variable  $X$  on  $(\Omega, \mathcal{F}, \mathbb{P})$ , let  $[X]$  be its law and  $[X|W]$  be its law given  $\mathcal{F}^W$ .

Let  $\mathcal{P}(\mathbb{R}^n)$  be the set of Borel probability measures on  $\mathbb{R}^n$  and, given  $q \geq 1$ , let  $\mathcal{P}^q(\mathbb{R}^n)$  be the subset of measures  $m$  in  $\mathcal{P}(\mathbb{R}^n)$  with finite  $q$  order moment:

$$\mathcal{M}_q(\mu) := \int_{\mathbb{R}^n} |x|^q \mu(dx) < +\infty, \quad \mu \in \mathcal{P}^q(\mathbb{R}^n).$$

We denote by  $W_2(\mu, \mu')$  the Wasserstein distance between two measures of  $\mathcal{P}^2(\mathbb{R}^n)$ . Let us recall that

$$W_2(\mu, \mu') = \inf_{X, X'} (\mathbb{E}[|X - X'|^2])^{1/2},$$

where the infimum is taken over all random variables  $(X, X')$  such that  $[X] = \mu$ ,  $[X'] = \mu'$ .

**Derivatives.** Given a map  $H : \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$ , we denote by  $D_\mu H : \mathcal{P}^2(\mathbb{R}^n) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  its derivative, when it exists, in Lions's sense (or "L-derivative"): see [24] or Definition 5.22 in [12]. Let us recall that  $H$  has an L-derivative at a measure  $\mu_0 \in \mathcal{P}^2(\mathbb{R}^n)$  if there exists a random variable  $X_0$  with  $[X_0] = \mu_0$  at which the lifted map  $\tilde{H} : L^2(\Omega) \rightarrow \mathbb{R}$ , defined by  $\tilde{H}(X) = H([X])$ , is differentiable. If  $H$  is L-differentiable on  $\mathcal{P}^2(\mathbb{R}^n)$ , then the derivative of  $\tilde{H}$  takes the form  $\nabla \tilde{H}(X) = D_\mu H([X])(X)$  ([24] or [12, Proposition 5.25]). As explained in [12, Remark 5.27], the derivative  $D_\mu H$  allows to quantify the Lipschitz regularity of  $H$ . Namely,  $H$  is Lipschitz continuous on  $\mathcal{P}^2(\mathbb{R}^n)$  if and only if

$$\mathcal{H}_p := \sup_{\mu \in \mathcal{P}^p(\mathbb{R}^n)} \int_{\mathbb{R}^n} |D_\mu H(\mu)(x)|^p \mu(dx) < +\infty,$$

with  $p = 2$ . In this case,  $\mathcal{H}_2$  is the Lipschitz constant of  $H$ . We will often require that the map  $\nabla \tilde{H}$  itself is Lipschitz continuous on  $L^2(\Omega)$ : in order words,

$$\forall X, Y \in L^2(\Omega), \quad \mathbb{E} \left[ |D_\mu H([X])(X) - D_\mu H([Y])(Y)|^2 \right] \leq C^2 \mathbb{E} [|X - Y|^2].$$

Then it is proved in [12, Proposition 5.36] that (up to redefining  $D_\mu H$ ),

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n), \forall x, x' \in \mathbb{R}^n, \quad |D_\mu H(\mu)(x) - D_\mu H(\mu)(x')| \leq C|x - x'|.$$

**Second order derivatives.** We now discuss further regularity of  $H$ . We say that  $H$  is *partially*  $C^2$  if the map  $y \rightarrow D_\mu H(\mu)(y)$  is continuously differentiable with a derivative  $(\mu, y) \rightarrow \partial_y D_\mu H(\mu)(y)$  jointly continuous on  $\mathcal{P}^2(\mathbb{R}^n) \times \mathbb{R}^n$  (see also [12, Definition 5.95]). We say that  $H$  is *globally*  $C^2$  if it is partially  $C^2$  and if the map  $\mu \rightarrow D_\mu H(\mu)(y)$  is  $L$ -differentiable with  $(\mu, y, y') \rightarrow D_\mu^2 H(\mu)(y, y')$  continuous (see also [12, Definition 5.82]).

**Itô's formulas.** When  $H$  is partially  $C^2$ , Itô's formula holds for the law of a diffusion of the form

$$dX_t = b_t dt + \sigma_t dB_t$$

where  $(b_t)$  and  $(\sigma_t)$  are progressively measurable with values in  $\mathbb{R}^n$  and  $\mathbb{R}^{n \times d}$  respectively and satisfy

$$\mathbb{E} \left[ \int_0^T (|b_s|^2 + |\sigma_s|^4) ds \right] < +\infty.$$



Namely, if  $H$  is partially  $C^2$  with

$$\sup_{\mu \in \mathcal{P}^2(\mathbb{R}^n)} \int_{\mathbb{R}^n} \left| \partial_y D_\mu H(\mu)(y) \right|^2 \mu(dy) < +\infty, \quad (11)$$

then

$$H([X_t]) = H([X_0]) + \int_0^t \mathbb{E} [D_\mu H([X_s])(X_s) \cdot b_s] ds + \frac{1}{2} \int_0^t \mathbb{E} [\text{Tr} (a_s \partial_y D_\mu H([X_s])(X_s))] ds, \quad (12)$$

where  $a_s = \sigma_s \sigma_s^*$  [12, Theorem 5.98].

A similar Itô's formula holds for the *conditional law* of an Itô's process of the form,

$$dX_t = b_t dt + \sigma_t^0 dB_t + \sigma_t^1 dW_t,$$

where  $B$  and  $W$  are independent  $d$ -dimensional Brownian motions living on different probability spaces  $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$  and  $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$  and where  $b$ ,  $\sigma^0$  and  $\sigma^1$  are progressively measurable with respect to the filtration generated by  $W$  and  $B$ , with

$$\mathbb{E} \left[ \int_0^T (|b_s|^2 + |\sigma_s^0|^4 + |\sigma_s^1|^4) ds \right] < +\infty.$$

We assume that  $H$  is globally  $C^2$  with

$$\sup_{\mu \in \mathcal{P}^2} \int_{\mathbb{R}^n} |D_\mu H(\mu)(x)|^2 \mu(dx) + \int_{\mathbb{R}^n} |\partial_x D_\mu H(\mu)(x)|^2 \mu(dx) + \int_{\mathbb{R}^n \times \mathbb{R}^n} |D_{\mu\mu}^2 H(\mu)(x, y)|^2 \mu(dx) \mu(dy) < +\infty.$$

Then, letting  $\mu_t(\omega^1) = [X_t|W](\omega^1)$ , we have,  $\mathbb{P}^1$ -a.s.,

$$\begin{aligned} H(\mu_t) &= H(\mu_0) + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s)(X_s) \cdot b_s] ds + \int_0^t \mathbb{E}^0 [(\sigma_s^1)^* D_\mu H(\mu_s)(X_s)] \cdot dW_s \\ &\quad + \frac{1}{2} \int_0^t \mathbb{E}^0 [\text{Tr} (a_s \partial_y D_\mu H(\mu_s)(X_s))] ds + \frac{1}{2} \int_0^t \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr} (D_{\mu\mu}^2 H(\mu_s)(X_s, \tilde{X}_s) \sigma_s^1 (\tilde{\sigma}_s^1)^*)] ds \end{aligned} \quad (13)$$

where  $\tilde{X}$  and  $\tilde{\sigma}^1$  are independent copies of  $X$  and  $\sigma^1$  is defined on the space  $(\tilde{\Omega}^0 \times \tilde{\Omega}^1, \tilde{\mathbb{P}}^0 \otimes \tilde{\mathbb{P}}^1)$ , while  $a_s := (\sigma_s^0 (\sigma_s^0)^* + \sigma_s^1 (\sigma_s^1)^*)$ . See [13, Theorem 11.13].

We actually need below more general versions of Itô's formulas (12) and (13) in which there is an additional drift term  $\beta_s dK_s$ . In the generalization of (12), we assume that  $\beta$  is a continuous process adapted to the filtration generated by  $B$  and that  $K$  is a deterministic, continuous and nondecreasing process with  $K_0 = 0$ . We are then interested in the law  $(\mu_t)$  of an Itô process

$$dX_t = b_t dt + \sigma_t dB_t + \beta_t dK_t.$$

Assuming, in addition to the conditions on  $b$ ,  $\sigma$  and  $H$  given above, that

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |\beta_t|^2 \right] < +\infty,$$

we have

$$\begin{aligned} H([X_t]) &= H([X_0]) + \int_0^t \mathbb{E} [D_\mu H([X_s])(X_s) \cdot b_s] ds + \int_0^t \mathbb{E} [D_\mu H([X_s])(X_s) \cdot \beta_s] dK_s \\ &\quad + \frac{1}{2} \int_0^t \mathbb{E} [\text{Tr} (a_s \partial_y D_\mu H([X_s])(X_s))] ds, \end{aligned} \quad (14)$$

where  $a_s = \sigma_s \sigma_s^*$ . For the generalization of (13), we assume instead that  $\beta$  is a continuous process, adapted to the filtration generated by  $B$  and  $W$  and that  $K$  is a continuous and nondecreasing process adapted to the filtration generated by  $W$  with  $K_0 = 0$ . We also assume

$$\mathbb{E} \left[ \sup_{t \in [0, T]} \mathbb{E}^0 [|\beta_t|^2] K_T^2 \right] < +\infty, \quad (15)$$

We are then interested in the conditional law, given  $W$ , of the process

$$dX_t = b_t dt + \sigma_t^0 dB_t + \sigma_t^1 dW_t + \beta_t dK_t.$$

Under the above conditions on  $b$ ,  $\sigma^0$ ,  $\sigma^1$ ,  $H$ ,  $\beta$  and  $K$ , we have,  $\mathbb{P}^1$ -a.s.,

$$\begin{aligned} H(\mu_t) &= H(\mu_0) + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s)(X_s) \cdot b_s] ds + \int_0^t \mathbb{E}^0 [(\sigma_s^1)^* D_\mu H(\mu_s)(X_s)] \cdot dW_s \\ &\quad + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s)(X_s) \cdot \beta_s] dK_s + \frac{1}{2} \int_0^t \mathbb{E}^0 [\text{Tr}(a_s \partial_y D_\mu H(\mu_s)(X_s))] ds \\ &\quad + \frac{1}{2} \int_0^t \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr}(D_{\mu\mu}^2 H(\mu_s)(X_s, \tilde{X}_s) \sigma_s^1 (\tilde{\sigma}_s^1)^*)] ds. \end{aligned} \quad (16)$$

*Proof of (14) and (16).* Formula (14) and (16) can be derived from (12) and (13) respectively by regularizing the process  $K$ . We give the details for (16) to fix the idea, as (14) can be proved with the same argument. We first extend  $K$  to  $\mathbb{R}$  by setting  $K_t = 0$  for  $t \leq 0$  and then let  $K^\varepsilon := \phi^\varepsilon * K$ , where  $\phi^\varepsilon(s) = \varepsilon^{-1} \phi(s/\varepsilon)$ ,  $\phi$  being a smooth nonnegative kernel with compact support in  $\mathbb{R}_+$ . Then,  $K^\varepsilon$  is a smooth nondecreasing process adapted to  $W$ . Let us set

$$X_t^\varepsilon = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s^0 dB_s + \int_0^t \sigma_s^1 dW_s + \int_0^t \beta_s (K^\varepsilon)'_s ds$$

and  $\mu_t^\varepsilon = [X_t^\varepsilon | W]$ . By (15) and the fact that  $K$  is  $\mathcal{F}^W$  adapted, we have

$$\mathbb{E} \left[ \int_0^T |\beta_s (K^\varepsilon)'_s|^2 ds \right] \leq \|(\phi^\varepsilon)'\|_\infty \mathbb{E} \left[ \sup_{0 \leq s \leq T} \mathbb{E}^0 [|\beta_s|^2] |K_T|^2 \right] < +\infty.$$

So, by Itô's formula (13), we obtain

$$\begin{aligned} H(\mu_t^\varepsilon) &= H(\mu_0) + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot b_s] ds + \int_0^t \mathbb{E}^0 [(\sigma_s^1)^* D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon)] \cdot dW_s \\ &\quad + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s] (K^\varepsilon)'_s ds + \frac{1}{2} \int_0^t \mathbb{E}^0 [\text{Tr}(a_s \partial_y D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon))] ds \\ &\quad + \frac{1}{2} \int_0^t \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr}(D_{\mu\mu}^2 H(\mu_s^\varepsilon)(X_s^\varepsilon, \tilde{X}_s^\varepsilon) \sigma_s^1 (\tilde{\sigma}_s^1)^*)] ds. \end{aligned}$$

In order to let  $\varepsilon \rightarrow 0$ , we first prove the uniform convergence of  $X^\varepsilon$  to  $X$ . We note that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^\varepsilon - X_t|^2 \right] \leq \mathbb{E} \left[ \sup_{0 \leq t \leq T} \left| \int_0^t (\beta_s - \int_s^t \beta_u \phi^\varepsilon(u-s) du) dK_s \right|^2 \right],$$

where

$$\sup_{0 \leq t \leq T} \left| \int_0^t (\beta_s - \int_s^t \beta_u \phi^\varepsilon(u-s) du) dK_s \right|^2 \leq 2 \sup_{0 \leq s \leq T} |\beta_s|^2 K_T^2,$$

the right-hand side being integrable by (15). Fix  $\delta > 0$  small. By the continuity of  $\beta$ , we have, uniformly in  $s \in (0, t - \delta)$ ,

$$\lim_{\varepsilon \rightarrow 0} \int_s^t \beta_u \phi^\varepsilon(u-s) du = \lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}} \beta_u \phi^\varepsilon(u-s) du = \beta_s.$$

So, a.s.,

$$\lim_{\varepsilon \rightarrow 0} \sup_{\delta \leq t \leq T} \left| \int_0^{t-\delta} (\beta_s - \int_s^t \beta_u \phi^\varepsilon(u-s) du) dK_s \right|^2 = 0.$$

On the other hand,

$$\left| \int_{(t-\delta) \vee 0}^t (\beta_s - \int_s^t \beta_u \phi^\varepsilon(u-s) du) dK_s \right| \leq 2 \sup_{0 \leq u \leq T} |\beta_u|^2 (K_t - K_{(t-\delta) \vee 0}).$$

So, by the continuity of  $K$ ,

$$\lim_{\delta \rightarrow 0} \limsup_{\varepsilon \rightarrow 0} \sup_{0 \leq t \leq T} \left| \int_{(t-\delta) \vee 0}^t (\beta_s - \int_s^t \beta_u \phi^\varepsilon(u-s) du) dK_s \right|^2 = 0.$$

We can then infer by dominate convergence that

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^\varepsilon - X_t|^2 \right] = 0 \quad \text{and thus} \quad \lim_{\varepsilon \rightarrow 0} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_2^2(\mu_t^\varepsilon, \mu_t) \right] = 0. \quad (17)$$

By our assumptions on the data, one easily derive from this the convergence in  $L_2(\Omega)$  of

$$\begin{aligned} H(\mu_t^\varepsilon) - H(\mu_0) - \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot b_s] ds - \int_0^t \mathbb{E}^0 [(\sigma_s^1)^* D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon)] \cdot dW_s \\ - \frac{1}{2} \int_0^t \mathbb{E}^0 [\text{Tr}(a_s \partial_y D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon))] ds - \frac{1}{2} \int_0^t \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr}(D_{\mu\mu}^2 H(\mu_s^\varepsilon)(X_s^\varepsilon, \tilde{X}_s^\varepsilon) \sigma_s^1 (\tilde{\sigma}_s^1)^*)] ds \end{aligned}$$

to

$$\begin{aligned} H(\mu_t) - H(\mu_0) + \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s)(X_s) \cdot b_s] ds - \int_0^t \mathbb{E}^0 [(\sigma_s^1)^* D_\mu H(\mu_s)(X_s)] \cdot dW_s \\ - \frac{1}{2} \int_0^t \mathbb{E}^0 [\text{Tr}(a_s \partial_y D_\mu H(\mu_s)(X_s))] ds - \frac{1}{2} \int_0^t \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr}(D_{\mu\mu}^2 H(\mu_s)(X_s, \tilde{X}_s) \sigma_s^1 (\tilde{\sigma}_s^1)^*)] ds. \end{aligned}$$

To complete the proof we just need to check that, in  $L_2(\Omega)$ ,

$$\lim_{\varepsilon \rightarrow 0} \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s] (K^\varepsilon)'_s ds = \int_0^t \mathbb{E}^0 [D_\mu H(\mu_s)(X_s) \cdot \beta_s] dK_s. \quad (18)$$

Indeed

$$\int_0^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s] (K^\varepsilon)'_s ds = \int_0^t \left( \int_r^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s] \phi_\varepsilon(s-r) ds \right) dK_r,$$

where, by Cauchy-Schwartz and our assumption on  $D_\mu H$ ,

$$\begin{aligned} \left| \int_0^t \left( \int_r^t \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s] \phi_\varepsilon(s-r) ds \right) dK_r \right| \\ \leq \int_0^t \left( \int_r^t (\mathbb{E}^0 [ |D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon)|^2 ])^{1/2} (\mathbb{E}^0 [ |\beta_s|^2 ])^{1/2} \phi_\varepsilon(s-r) ds \right) dK_r \leq C \left( \mathbb{E}^0 \left[ \sup_{0 \leq s \leq T} |\beta_s|^2 \right] \right)^{1/2} K_T. \end{aligned}$$

The right-hand side of the above inequality is in  $L^2(\Omega)$  thanks to (15). Using the uniform convergence of  $s \rightarrow \mathbb{E}^0 [D_\mu H(\mu_s^\varepsilon)(X_s^\varepsilon) \cdot \beta_s]$  implied by (17), we can conclude exactly as in the proof of (17) that (18) holds.  $\square$

**On uniform convergence of conditionally i.i.d. sequence of processes.** The aim of this part is to recast in the conditionally i.i.d. setting the result of [29] (see Theorem 10.2.7 therein) taking into account the recent results obtained in [17] and in Chapter 5 of [12].

**Lemma 1.** *Let  $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$  and  $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1)$  be two different probability spaces. Let  $\{X^i\}_{1 \leq i \leq N}$  be a sequence of processes defined on  $(\Omega^0 \times \Omega^1, \mathcal{F}^0 \otimes \mathcal{F}^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$  satisfying that the  $X^i$  are, conditionally to  $\mathcal{F}^1$ , i.i.d. processes. Assume that there exists constants  $C_0 > 0$ ,  $q > 4$  and  $p > 2$  such that*

$$\mathbb{E}^1 \left[ \sup_{t \in [0,1]} \mathbb{E}^0 [|X_t|^q]^{2p/q} \right] \leq C_0$$

and

$$\begin{aligned} \mathbb{E}[|X_s - X_r|^p | X_s - X_t|^p] &\leq C_0 |t-r|^2, \quad \text{for } 0 \leq r < s < t \leq 1; \\ \mathbb{E}[|X_t - X_s|^p] &\leq C_0 |t-s|, \quad \text{for } 0 \leq s \leq t \leq 1; \end{aligned} \quad (19)$$

$$\forall 0 \leq s \leq t \leq 1, \quad \mathbb{E} \left[ \sup_{u \in [s,t]} |\mathbb{E}^0 [|X_u - X_s|^2]|^p \right] \leq C_0 |t-s|^{p/2}.$$

Then, setting  $\mu_s = [X_s | \mathcal{F}^1]$  and  $\bar{\mu}_s^N = N^{-1} \sum_{i=1}^N \delta_{X_s^i}$ , there exists a constant  $C > 0$ , depending on  $n, p, q$  and  $C_0$  only, such that

$$\mathbb{E} \left[ \sup_{s \leq T} W_2^2(\bar{\mu}_s^N, \mu_s) \right] \leq C \epsilon_N := C \times \begin{cases} N^{-1/2+1/p} & \text{if } d < 4, \\ N^{-1/2+1/p} (\log(1+N))^{1-2/p} & \text{if } d = 4, \\ N^{-2(1-2/p)/d} & \text{if } d > 4. \end{cases} \quad (20)$$

*Proof.* Using a time discretization  $(t_k)_{0 \leq k \leq m}$  with constant time step  $1/m$  of the interval  $[0; 1]$  we obtain that

$$\sup_{s \leq 1} W_2^2(\bar{\mu}_s^N, \mu_s) \leq 3 \left\{ \max_k \chi_k + \max_k W_2^2(\bar{\mu}_{t_k}^N, \mu_{t_k}) + \max_k \sup_{t_k \leq t \leq t_{k+1}} W_2^2(\mu_{t_k}, \mu_t) \right\},$$

where

$$\chi_k = \sup_{t_k \leq t \leq t_{k+1}} \left\{ W_2^2(\bar{\mu}_t^N, \bar{\mu}_{t_k}^N) \wedge W_2^2(\bar{\mu}_t^N, \bar{\mu}_{t_{k+1}}^N) \right\}.$$

Now we have

$$\begin{aligned} \mathbb{E} \left[ \max_k \sup_{t_k \leq t \leq t_{k+1}} W_2^2(\mu_{t_k}, \mu_t) \right] &\leq \mathbb{E} \left[ \max_k \sup_{t_k \leq t \leq t_{k+1}} \mathbb{E}^0 [ |X_t - X_{t_k}|^2 ] \right] \\ &\leq \mathbb{E}^{1/p} \left[ \max_k \sup_{t_k \leq t \leq t_{k+1}} |\mathbb{E}^0 [ |X_t - X_{t_k}|^2 ]|^p \right] \\ &\leq \mathbb{E}^{1/p} \left[ \sum_k \sup_{t_k \leq t \leq t_{k+1}} |\mathbb{E}^0 [ |X_t - X_{t_k}|^2 ]|^p \right] \end{aligned}$$

Therefore, using third line in Assumption (19) we obtain

$$\mathbb{E} \left[ \max_k \sup_{t_k \leq t \leq t_{k+1}} W_2^2(\mu_{t_k}, \mu_t) \right] \leq C m^{-1/2+1/p}. \quad (21)$$

Following now the proof of Theorem 10.2.7 of [29] we have

$$\mathbb{E} \left[ \max_k \chi_k \right] \leq \frac{C}{\sqrt{m}}, \quad (22)$$

which will lead to a negligible contribution. Concerning the last term we write

$$\mathbb{E} \left[ \max_k W_2^2(\bar{\mu}_{t_k}^N, \mu_{t_k}) \right] \leq m^{1/p} \left( \max_k \mathbb{E} \left[ W_2^{2p}(\bar{\mu}_{t_k}^N, \mu_{t_k}) \right] \right)^{1/p}. \quad (23)$$

Since  $(\mathbb{E}^0)^{2p/q} [|X_t|^q] < +\infty$   $\mathbb{P}^1$ -a.s. for some  $q > 4$ , we have, from Theorem 2 case (3) of [17] (see also Theorem 5.8 and Remark 5.9 in [12]) that there exists a deterministic constant  $c > 0$  such that  $\mathbb{P}^1$ -a.s.

$$W_2^{2p}(\bar{\mu}_{t_k}^N, \mu_{t_k}) \leq c (\mathbb{E}^0)^{2p/q} [|X_{t_k}|^q] \begin{cases} N^{-p/2} & \text{if } d \leq 3 \\ N^{-p/2} (\log(1+N))^p & \text{if } d = 4 \\ N^{-2p/d} & \text{if } d > 4 \end{cases}$$

Hence,

$$\left( \mathbb{E} \left[ \max_k W_2^{2p}(\bar{\mu}_{t_k}^N, \mu_{t_k}) \right] \right)^{1/p} \leq c \mathbb{E} \left[ \sup_{0 \leq t \leq 1} (\mathbb{E}^0)^{2p/q} [|X_t|^q] \right] \begin{cases} N^{-1/2} & \text{if } d \leq 3 \\ N^{-1/2} (\log(1+N)) & \text{if } d = 4 \\ N^{-2/d} & \text{if } d > 4 \end{cases} \quad (24)$$

We conclude by optimizing over  $m$ .  $\square$

*Remark 2.* We emphasize that, when working with i.i.d. sequence of processes, the above result holds assuming only that

$$\mathbb{E} [|X_s - X_t|^2] \leq |t - s|, \quad \text{for all } 0 \leq s \leq t \leq 1,$$

instead of the third line in (19). In such case the control in (21) is direct and is  $m^{-1}$  (this follows essentially from the fact that in that case we have  $\mathbb{E} = \mathbb{E}^0$  so that  $p$  could be, formally,  $\infty$ ). The resulting rate of convergence are then given by the  $\epsilon_N$  defined above with  $p = \infty$  therein up to the ad hoc constant.

## 2. SDEs with normal reflexion in law.

In this section, we are concerned with the existence and uniqueness for the solution of (3). To avoid unnecessary repetitions, we postpone the discussion of the particle system and of the Feynman-Kac formula to the more general setting of the SDE with normal reflexion in conditional law. As a warm up we start by studying

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s + \int_0^t \nabla h(X_s) dK_s, & t \geq 0, \\ \mathbb{E}[h(X_t)] \geq 0, \quad \int_0^t \mathbb{E}[h(X_s)] dK_s = 0, & t \geq 0, \end{cases} \quad (25)$$

which is the *normally reflected* version of (1).

**2.1. Warm-up.** In this subsection, we study the case where  $H(\mu)$  is given by  $\int h d\mu$  in the spirit of the works [4], [3]. Roughly speaking, we extend the results of the previous papers to the multidimensional setting.

In the following, we say that assumptions **(Hwu)** hold when the following assumptions are in force:

**(Hc):** The functions  $b : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$  are measurable with respect to  $\mathcal{E} \otimes \mathcal{B}(\mathbb{R}^n)$  and

(i) For all  $T > 0$ , there exists  $L_T$  such that,  $\mathbb{P}$ -a.s., for each  $t \in [0, T]$ ,

$$\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n, \quad |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L_T |x - y|;$$

(ii) For all  $T > 0$ ,

$$\sup_{t \leq T} \mathbb{E} [|b(t, 0)|^2 + |\sigma(t, 0)|^2] < +\infty.$$

**(Hh):** The function  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  is  $\mathcal{C}^2$  with  $|\nabla^2 h|_\infty < \infty$  and

(i)  $h$  is concave;

(ii) For all  $x \in \mathbb{R}^n$ ,  $|\nabla h(x)|^2 > 0$ ;

(iii) There exists  $x_0 \in \mathbb{R}^n$  such that  $h(x_0) \geq 0$ .

**(H0):** The initial condition is  $\mathcal{F}_0$ -measurable, square integrable and  $H([X_0]) \geq 0$ .

Before going further, let us recall that, in this subsection,  $H([X_0]) = \mathbb{E}[h(X_0)]$ .

*Remark 3.* Let us observe that the assumption **(Hh)**-(iii) is needed to ensure that the set

$$\left\{ \mu \in \mathcal{P}^2(\mathbb{R}^n) : H(\mu) = \int h d\mu \geq 0 \right\}$$

is not empty. For example, if  $h(x) = -e^{-x}$  this set is empty. Moreover, when **(Hh)** holds true the set

$$\left\{ \mu \in \mathcal{P}^2(\mathbb{R}^n) : H(\mu) = \int h d\mu > 0 \right\}$$

is also non empty since it contains, in particular, a Dirac mass. Indeed, let  $(x_t)_{t \geq 0}$  be the solution to

$$x'(t) = \nabla h(x(t)) dt, \quad t \geq 0, \quad x(0) = x_0.$$

Then, for  $t > 0$ ,

$$h(x(t)) = h(x_0) + \int_0^t |\nabla h(x(s))|^2 ds \geq \int_0^t |\nabla h(x(s))|^2 ds > 0.$$

*Remark 4.* Let us point out that the concavity of  $h$  allows to weaken the assumption on  $\nabla h$  used in [4] and [3] namely: there exists  $\beta > 0$  such that

$$\forall x \in \mathbb{R}^n, \quad |\nabla h(x)|^2 \geq \beta.$$

We only assume here that  $|\nabla h|^2$  does not vanish instead of being bounded from below by some positive constant. Moreover, the gradient of  $h$  is only supposed to have a linear growth instead of being bounded as assumed in the two papers mentioned above.

To conclude the comparison with [4, 3], let us mention that, in the one dimensional case, one can switch from the framework of the two quoted papers, namely

$$\begin{cases} P_t = X_0 + \int_0^t d(s, P_s) ds + \int_0^t S(s, P_s) dB_s + K_t, & t \geq 0, \\ \mathbb{E}[l(P_t)] \geq 0, \quad \int_0^t \mathbb{E}[l(P_s)] dK_s = 0, & t \geq 0, \end{cases}$$

with  $K$  deterministic to the setting of (25) by the transformation  $P_t = F(X_t)$  where

$$F(x) = \int_0^x \frac{dx}{h'(x)}, \quad x \in \mathbb{R}.$$

**Definition 5.** By a solution to (25), we mean a couple of adapted and continuous processes  $(X, K)$ ,  $K$  being deterministic and non decreasing with  $K_0 = 0$ , such that (25) holds.

**Theorem 6.** *Under assumptions (Hwu), the SDE (25) has a unique square integrable solution with  $K$  deterministic.*

*Proof.* Let us start by uniqueness. Suppose that  $(X, K^X)$  and  $(Y, K^Y)$  are two solutions of (25). Then, we deduce from Itô's formula that

$$\begin{aligned} \mathbb{E}[|X_t - Y_t|^2] &= 2\mathbb{E}\left[\int_0^t (X_s - Y_s) \cdot (b(s, X_s) - b(s, Y_s)) ds\right] + \mathbb{E}\left[\int_0^t |\sigma(s, X_s) - \sigma(s, Y_s)|^2 ds\right] \\ &\quad + 2\mathbb{E}\left[\int_0^t (X_s - Y_s) \cdot (\nabla h(X_s) dK_s^X - \nabla h(Y_s) dK_s^Y)\right]. \end{aligned}$$

Since,  $h$  is concave, for all  $x \in \mathbb{R}^n$ ,  $y \in \mathbb{R}^n$ ,  $h(y) \leq h(x) + \nabla h(x) \cdot (y - x)$ . This last inequality rewrites

$$\forall x \in \mathbb{R}^n, \quad \forall y \in \mathbb{R}^n, \quad (x - y) \cdot \nabla h(x) \leq h(x) - h(y). \quad (26)$$

Thus, using the Skorokhod condition together with the constraint, we have

$$\int_0^t \mathbb{E}[(X_s - Y_s) \cdot \nabla h(X_s)] dK_s^X \leq \int_0^t (\mathbb{E}[h(X_s)] - \mathbb{E}[h(Y_s)]) dK_s^X \leq 0,$$

and, similarly,

$$-\mathbb{E}\left[\int_0^t (X_s - Y_s) \cdot \nabla h(Y_s) dK_s^Y\right] \leq 0.$$

Using the fact that  $b$  and  $\sigma$  are Lipschitz, we get, for all  $t \leq T$ ,

$$\mathbb{E}[|X_t - Y_t|^2] \leq (2L_T + L_T^2) \int_0^t \mathbb{E}[|X_s - Y_s|^2] ds,$$

and Gronwall's lemma implies that  $X$  and  $Y$  are equal on  $[0, T]$  for each  $T > 0$ .

Since  $X = Y$ , we get from Itô's formula, that for any  $t$ ,

$$\begin{aligned} A_t^X &:= \int_0^t \mathbb{E}[|\nabla h(X_u)|^2] dK_u^X \\ &= \mathbb{E}[h(X_t)] - \mathbb{E}[h(X_0)] - \int_0^t \mathbb{E}[\nabla h(X_u) b(u, X_u)] du - \frac{1}{2} \int_0^t \mathbb{E}[\text{Tr}(\nabla^2 h(X_u) \sigma \sigma^*(u, X_u))] du \\ &= \int_0^t \mathbb{E}[|\nabla h(Y_u)|^2] dK_u^Y = \int_0^t \mathbb{E}[|\nabla h(X_u)|^2] dK_u^Y =: A_t^Y. \end{aligned}$$

Thus the nondecreasing continuous functions  $A^X$  and  $A^Y$  are equal. For any step function  $\varphi$  on  $[0, T]$ ,

$$\int_0^T \varphi(u) dA_u^X = \int_0^T \varphi(u) \mathbb{E}[|\nabla h(X_u)|^2] dK_u^X = \int_0^T \varphi(u) \mathbb{E}[|\nabla h(X_u)|^2] dK_u^Y = \int_0^T \varphi(u) dA_u^Y. \quad (27)$$

The previous equality hold true also for any continuous function on  $[0, T]$  as uniform limit of step functions. The map  $u \mapsto \mathbb{E}[|\nabla h(X_u)|^2]$  is continuous and does not vanish since  $|\nabla h(x)|^2 > 0$  for all  $x \in \mathbb{R}^n$ . As a byproduct,

$$u \mapsto \varphi(u) = \mathbb{E}[|\nabla h(X_u)|^2]^{-1}$$

is continuous and we get, plugging this  $\varphi$  in (27),  $K_T^X = K_T^Y$ .

Let us fix  $T > 0$  and let us now construct a solution to (25) on  $[0, T]$ . For this, let us consider, for  $k \geq 1$ ,  $X^k$  solution to the following McKean-Vlasov SDE:

$$X_t^k = X_0 + \int_0^t b(s, X_s^k) ds + \int_0^t \sigma(s, X_s^k) dB_s + \int_0^t \nabla h(X_s^k) \psi_k(\mathbb{E}[h(X_s^k)]) ds, \quad t \geq 0, \quad (28)$$

where the function  $\psi_k$  is defined as

$$\psi_k(x) = r \text{ if } x \leq -1/k, \quad \psi_k(x) = -krx, \text{ if } -1/k \leq x \leq 0, \quad \psi_k(x) = 0, \text{ if } x \geq 0.$$

The function  $\psi_k$  depends on the constant  $r > 0$  which will be chosen later.

Existence and uniqueness of solutions to (28) under the assumptions **(HWu)** follows from straightforward generalizations of results from [10]. We set

$$K_t^k = \int_0^t \psi_k(\mathbb{E}[h(X_s^k)]) ds,$$

and rewrite (28) as

$$X_t^k = X_0 + \int_0^t b(s, X_s^k) ds + \int_0^t \sigma(s, X_s^k) dB_s + \int_0^t \nabla h(X_s^k) dK_s^k, \quad t \geq 0.$$

Let us start with some  $L^2$  bounds. Let us compute  $|X_t^k - X_0|^2$  with the help of Itô's formula. We have, for  $0 \leq t \leq T$ ,

$$\begin{aligned} |X_t^k - X_0|^2 &= 2 \int_0^t (X_s^k - X_0) \cdot b(s, X_s^k) ds + \int_0^t |\sigma(s, X_s^k)|^2 ds \\ &\quad + 2 \int_0^t (X_s^k - X_0) \cdot \sigma(s, X_s^k) dB_s + 2 \int_0^t (X_s^k - X_0) \cdot \nabla h(X_s^k) dK_s^k. \end{aligned}$$

Since  $b$  and  $\sigma$  are Lipschitz in space, uniformly in time, we deduce from the above equality that, for

$$\begin{aligned} |X_t^k - X_0|^2 &\leq (2L_T + 2L_T^2) \int_0^t |X_s^k - X_0|^2 ds + 2 \int_0^t (|b(s, X_0)|^2 + |\sigma(s, X_0)|^2) ds \\ &\quad + 2 \int_0^t (X_s^k - X_0) \cdot \sigma(s, X_s^k) dB_s + 2 \int_0^t (X_s^k - X_0) \cdot \nabla h(X_s^k) dK_s^k. \end{aligned} \quad (29)$$

Since  $h$  is concave and  $K^k$  is deterministic, we obtain, using (26),

$$\begin{aligned} \mathbb{E} \left[ \int_0^t (X_s^k - X_0) \cdot \nabla h(X_s^k) dK_s^k \right] &\leq \int_0^t (\mathbb{E}[h(X_s^k)] - \mathbb{E}[h(X_0)]) dK_s^k \\ &= \int_0^t (\mathbb{E}[h(X_s^k)] - \mathbb{E}[h(X_0)]) \psi_k(\mathbb{E}[h(X_s^k)]) ds \leq 0, \end{aligned}$$

the last inequality coming from the fact that  $x\psi_k(x) \leq 0$  and  $\mathbb{E}[h(X_0)] \geq 0$ . It follows from Gronwall's inequality, that,

$$\sup_{k \geq 1} \sup_{t \leq T} \mathbb{E} \left[ |X_t^k|^2 \right] \leq C(T, L_T, \mathbb{E}[|X_0|^2]) =: \gamma. \quad (30)$$

Coming back to (29), since  $\psi_k$  is bounded by  $r$  and  $\nabla h$  has a linear growth, BDG's inequality gives

$$\sup_{k \geq 1} \mathbb{E} \left[ \sup_{t \leq T} |X_t^k|^2 \right] \leq C(T, L_T, \mathbb{E}[|X_0|^2], |\nabla^2 h|_\infty, r).$$

Let us apply Itô's formula to compute the expected value of  $h(X_t^k)$ . We get, for  $0 \leq s \leq t \leq T$ ,

$$\begin{aligned} \mathbb{E}[h(X_t^k)] &= \mathbb{E}[h(X_s^k)] + \int_s^t \mathbb{E} \left[ |\nabla h(X_u^k)|^2 \right] \psi_k(\mathbb{E}[h(X_u^k)]) du \\ &\quad + \int_s^t \mathbb{E} [\nabla h(X_u^k) \cdot b(u, X_u^k)] du + \frac{1}{2} \int_s^t \mathbb{E} [\text{Tr}(\nabla^2 h(X_u^k) \sigma \sigma^*(u, X_u^k))] du. \end{aligned}$$

Since  $\nabla^2 h$  is bounded,  $b$  and  $\sigma$  are Lipschitz, we deduce, taking into account (30) and assumption **(Hc)**-**(ii)** that, for some constant  $C \geq 0$  independent of  $k$  and  $r$ ,

$$\mathbb{E}[h(X_t^k)] \geq \mathbb{E}[h(X_s^k)] + \int_s^t \mathbb{E} \left[ |\nabla h(X_u^k)|^2 \right] \psi_k(\mathbb{E}[h(X_u^k)]) du - C(t-s).$$

On the other hand, for any  $k \geq 1$ ,  $0 \leq u \leq T$  and  $a > 0$ , we have using Markov's inequality,

$$\begin{aligned} \mathbb{E} \left[ |\nabla h(X_u^k)|^2 \right] &\geq \mathbb{E} \left[ |\nabla h(X_u^k)|^2 \mathbf{1}_{|X_u^k| \leq a} \right] \geq \inf \{ |\nabla h|^2(x) : |x| \leq a \} \mathbb{P}(|X_u^k| \leq a), \\ &\geq \inf \{ |\nabla h|^2(x) : |x| \leq a \} \left( 1 - a^{-2} \mathbb{E} \left[ |X_u^k|^2 \right] \right). \end{aligned}$$

Having in mind the bound (30), we choose  $a = \sqrt{2\gamma}$ , to get, since  $|\nabla h|^2$  is continuous, for any  $k \geq 1$ ,  $0 \leq u \leq T$ ,

$$\mathbb{E} \left[ |\nabla h(X_u^k)|^2 \right] \geq \min \{ |\nabla h|^2(x) : |x| \leq \sqrt{2\gamma} \} / 2 = m^2 > 0.$$

It follows that, for any  $k \geq 1$ ,  $0 \leq s \leq t \leq T$ ,

$$\mathbb{E} [h(X_t^k)] \geq \mathbb{E} [h(X_s^k)] + m^2 \int_s^t \psi_k(\mathbb{E} [h(X_u^k)]) du - C(t-s). \quad (31)$$

We choose  $r$  such that  $rm^2 > C$ . As a byproduct, we have

$$\forall k \geq 1, \quad \forall 0 \leq t \leq T, \quad \mathbb{E} [h(X_t^k)] \geq -\frac{1}{k}. \quad (32)$$

Indeed, if  $\mathbb{E} [h(X_t^k)] < -1/k$  for some  $k \geq 1$  and  $t > 0$ , let  $s = \sup\{u \leq t : \mathbb{E} [h(X_u^k)] \geq -1/k\}$ . Since  $\mathbb{E} [h(X_0)] \geq 0$ , we have  $0 < s < t$  and  $\mathbb{E} [h(X_u^k)] \leq -1/k$  for  $s \leq u \leq t$ . Thus, by definition of  $\psi_k$ ,  $\psi_k(\mathbb{E} [h(X_u^k)]) \geq r$  for  $u \in [s, t]$  and it follows from (31) that

$$\mathbb{E} [h(X_t^k)] \geq \mathbb{E} [h(X_s^k)] + (m^2 r - C)(t-s) > -\frac{1}{k}.$$

This is a contradiction.

Let  $k, m \in \mathbb{N}^*$ , by Itô's formula we have, for  $t \geq 0$ ,

$$\begin{aligned} |X_t^k - X_t^m|^2 &= 2 \int_0^t (X_s^k - X_s^m) \cdot (b(s, X_s^k) - b(s, X_s^m)) ds + \int_0^t |\sigma(s, X_s^k) - \sigma(s, X_s^m)|^2 ds \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (\sigma(s, X_s^k) - \sigma(s, X_s^m)) dB_s \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (\nabla h(X_s^k) dK_s^k - \nabla h(X_s^m) dK_s^m). \end{aligned}$$

Using the fact that  $b$  and  $\sigma$  are Lipschitz continuous, we get, for  $0 \leq t \leq T$ ,

$$\begin{aligned} |X_t^k - X_t^m|^2 &\leq (2L_T + L_T^2) \int_0^t |X_s^k - X_s^m|^2 ds + 2 \int_0^t (X_s^k - X_s^m) \cdot (\sigma(s, X_s^k) - \sigma(s, X_s^m)) dB_s \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (\nabla h(X_s^k) dK_s^k - \nabla h(X_s^m) dK_s^m). \end{aligned} \quad (33)$$

Arguing as in the proof of uniqueness, taking the expectation, we have, since  $h$  is concave,

$$\begin{aligned} \mathbb{E} \left[ \int_0^t (X_s^k - X_s^m) \cdot (\nabla h(X_s^k) dK_s^k - \nabla h(X_s^m) dK_s^m) \right] \\ \leq \mathbb{E} \left[ \int_0^t (h(X_s^k) - h(X_s^m)) (dK_s^k - dK_s^m) \right] \\ = \int_0^t (\mathbb{E} [h(X_s^k)] - \mathbb{E} [h(X_s^m)]) (\psi_k(\mathbb{E} [h(X_s^k)]) - \psi_m(\mathbb{E} [h(X_s^m)])) ds, \end{aligned}$$

and from (32), since, for any  $k \geq 1$ ,  $x\psi_k(x) \leq 0$ ,  $\mathbb{E} [h(X_s^k)] \geq -1/k$  and  $0 \leq \psi_k \leq r$ , we obtain that

$$\mathbb{E} \left[ \int_0^t (X_s^k - X_s^m) \cdot (\nabla h(X_s^k) dK_s^k - \nabla h(X_s^m) dK_s^m) \right] \leq \frac{r}{k} + \frac{r}{m}.$$

We conclude using Gronwall's lemma that there exists  $C \geq 0$  depending on  $T$  and  $L_T$  such that for any  $k, m \in \mathbb{N}^*$  we have:

$$\sup_{0 \leq t \leq T} \mathbb{E} [|X_t^k - X_t^m|^2] \leq C r \left( \frac{1}{k} + \frac{1}{m} \right). \quad (34)$$



Coming back to (33), since  $\sup_{k \geq 1} |\psi_k|_\infty \leq r$  and  $\nabla h$  has at most a linear growth, taking into account (30), we derive, from BDG and Cauchy-Schwarz inequalities, the estimate

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^k - X_t^m|^2 \right] \leq C \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{m}} \right),$$

where the constant  $C$  does not depend on  $k$  and  $m$ .

Thus,  $(X^k)_{k \geq 1}$  is a Cauchy sequence in  $\mathcal{S}^2$ . Let us denote by  $X$  its limit. Finally, since  $\psi_k$  is bounded by  $r$  for all  $k$ ,  $K^k$  is Lipschitz with  $|K^k|_{\text{Lip}} \leq r$ . By Arzelà-Ascoli's theorem, up to a subsequence,  $(K^k)_{k \geq 1}$  converges, uniformly on  $[0, T]$ , towards a non decreasing, Lipschitz continuous function  $K$ .

It is straightforward to check that  $(X, K)$  solves (25). Indeed, since  $\nabla h$  has a linear growth, we have

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathbb{E} [|h(X_t^k) - h(X_t)|] &\leq C \sup_{0 \leq t \leq T} \mathbb{E} [|X_t^k - X_t| (1 + |X_t^k| + |X_t|)] \\ &\leq C \sup_{0 \leq t \leq T} \mathbb{E}^{1/2} [|X_t^k - X_t|^2] \left( 1 + \sup_{k \geq 1} \mathbb{E}^{1/2} [|X_t^k|^2] \right) \end{aligned}$$

and thus, for  $0 \leq t \leq T$ ,

- $\mathbb{E} [h(X_t)] = \lim_{k \rightarrow \infty} \mathbb{E} [h(X_t^k)] \geq 0$  by (32),
- the Skorokhod condition is also satisfied : since  $x\psi_k(x) \leq 0$ ,

$$0 \leq \int_0^t \mathbb{E} [h(X_s)] dK_s = \lim_{k \rightarrow \infty} \int_0^t \mathbb{E} [h(X_s^k)] dK_s^k = \lim_{k \rightarrow \infty} \int_0^t \mathbb{E} [h(X_s^k)] \psi_k (\mathbb{E} [h(X_t^k)]) ds \leq 0.$$

The proof of this result is complete.  $\square$

**2.2. Existence and uniqueness of the solution.** In this section, we are concerned with the existence and the uniqueness of a solution to (3) when  $H(\mu)$  is not necessarily of the form  $\int h d\mu$  and is not necessarily concave in  $\mu$ . However, in this case, we have to assume that the volatility  $\sigma$  is bounded. More precisely we assume that the following conditions **(H)** hold:

**(Hc):** The functions  $b : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$  are measurable with respect to  $\mathcal{E} \otimes \mathcal{B}(\mathbb{R}^n)$  and, for all  $T > 0$ , there exists  $L_T$  such that,  $\mathbb{P}$ -a.s.,

(i) For each  $t \in [0, T]$ ,

$$\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n, \quad |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L_T |x - y|;$$

(ii) For each  $t \in [0, T]$ ,

$$\mathbb{E} [|b(t, 0)|^2] + \sup_{x \in \mathbb{R}^n} |\sigma(t, x)| \leq L_T.$$

**(H0):** The initial condition is  $\mathcal{F}_0$ -measurable, square integrable and  $H([X_0]) \geq 0$  ;

**(HH):** The function  $H : \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  is partially  $\mathcal{C}^2$  and

(i) there exists  $M > 0$  such that

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx) \leq M^2, \quad (35)$$

(ii) there exist  $\beta > 0$  and  $\eta > 0$  such that

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n) \text{ with } -\eta \leq H(\mu) \leq 0, \quad \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx) \geq \beta^2, \quad (36)$$

(iii) there exists  $C \geq 0$  such that

$$\forall X, Y \in L^2(\Omega), \quad \mathbb{E} [|D_\mu H([X])(X) - D_\mu H([Y])(Y)|^2] \leq C \mathbb{E} [|X - Y|^2]. \quad (37)$$

**Examples.** We now illustrate our assumptions through two examples. In the first one we consider the case where  $H$  depends on the first order moment of the measure: Let  $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  be of class  $\mathcal{C}^2$  with bounded first and second derivatives, and let

$$H_1(\mu) = f_1 \left( \int_{\mathbb{R}^n} y \mu(dy) \right), \quad \forall \mu \in \mathcal{P}^2(\mathbb{R}^n).$$

We also assume that the set  $\{f_1 = 0\}$  is compact and that 0 is a noncritical value of  $f_1$ :  $\nabla f_1(x) \neq 0$  if  $f_1(x) = 0$ . Note that this implies the existence of a constant  $C > 0$  such that  $C^{-1} \leq |\nabla f_1(x)| \leq C$  for any  $x \in \mathbb{R}^n$  with  $f_1(x) = 0$ . As

$$D_\mu H_1(\mu)(x) = \nabla f_1 \left( \int_{\mathbb{R}^n} y \mu(dy) \right)$$

is independent of  $x$ , it is clear that (35), (36) and (37) hold.

In the second example, we assume that  $H_2$  depends on the second order moments of the measure. Let  $S^n$  be the set of  $n \times n$  symmetric matrices, endowed with the usual euclidean distance:  $|A| = (\text{Tr}(A^2))^{\frac{1}{2}}$ , and let  $S_+^n$  be the subset of nonnegative matrices. Let  $f_2 : S^n \rightarrow \mathbb{R}$  be of class  $C^2$ . We assume that the set  $\{f_2 \geq -\eta_0\} \cap S_+^n$  is compact for some  $\eta_0 > 0$ , that 0 is a noncritical value of  $f_2$  and that  $f_2(0) \neq 0$ . Then there exists  $C > 0$  with  $C^{-1} \leq |\nabla f_2| \leq C$  on  $\{|f_2| \leq \eta\} \cap S_+^n$ , for  $\eta \in (0, \eta_0)$  small enough, where  $\nabla f_2$  is the gradient of  $f_2$  in  $S^n$ . Moreover, as  $f_2(0) \neq 0$  and  $\{f_2 \geq -\eta\} \cap S_+^n$  is compact, reducing  $\eta > 0$  if necessary, there exists  $\delta > 0$  such that  $\text{Tr}(A) \geq \delta$  if  $|f_2(A)| \leq \eta$  and  $A \in S_+^n$ . Finally, we can assume, without changing the level-set  $\{f_2 = 0\} \cap S_+^n$ , that the restriction of  $f_2$  to  $S_+^n$  is bounded and that its derivatives have compact support. Let

$$H_2(\mu) = f_2 \left( \int_{\mathbb{R}^n} xx^* \mu(dx) \right) \quad \forall \mu \in \mathcal{P}^2(\mathbb{R}^n).$$

Then

$$D_\mu H_2(\mu)(x) = \nabla f_2 \left( \int_{\mathbb{R}^n} xx^* \mu(dx) \right) x.$$

Hence

$$\int_{\mathbb{R}^n} |D_\mu H_2(\mu)(x)|^2 \mu(dx) = \left| \nabla f_2 \left( \int_{\mathbb{R}^n} xx^* \mu(dx) \right) \right|^2 \int_{\mathbb{R}^n} |x|^2 \mu(dx).$$

Let us note that condition (35) holds because, for any  $\mu \in \mathcal{P}^2(\mathbb{R}^n)$ ,  $\int_{\mathbb{R}^n} xx^* \mu(dx)$  belongs to  $S_+^n$  and because the restriction  $\nabla f_2$  to  $S_+^n$  has a compact support. We now check the bound (36). Let  $\mu$  be such that  $-\eta \leq H_2(\mu) \leq 0$  and let us set  $A := \int_{\mathbb{R}^n} xx^* \mu(dx)$ . Then  $A \in S_+^n$  and, as  $|f_2(A)| \leq \eta$ , we have  $\text{Tr}(A) \geq \delta$  and  $|\nabla f_2(A)| \geq C^{-1}$ . So

$$\int_{\mathbb{R}^n} |D_\mu H_2(\mu)(x)|^2 \mu(dx) = |\nabla f_2(A)|^2 \text{Tr}(A) \geq C^{-2} \delta,$$

which proves (36). Finally, we check (37). Let us recall that the restrictions to  $S_+^n$  of  $\nabla f_2$  and  $\nabla^2 f_2$  have a compact support (say in a ball of radius  $R$ ). Let  $C_0$  be a bound on  $\nabla f_2$  and  $\nabla^2 f_2$  on  $S_+^n$ . Let also  $X, Y \in L_{\mathbb{F}}^2$ ,  $\mu = [X]$ ,  $\nu = [Y]$  and  $A := \int_{\mathbb{R}^n} xx^* \mu(dx)$ ,  $B := \int_{\mathbb{R}^n} xx^* \nu(dx)$ . We assume that  $|A|, |B| \leq R$ , since otherwise the result is obvious. Then

$$\begin{aligned} \mathbb{E} [ |D_\mu H_2([X])(X) - D_\mu H_2([Y])(Y)|^2 ] &= \mathbb{E} [ |\nabla f_2(A)X - \nabla f_2(B)Y|^2 ] \\ &\leq C_0 (|A - B|^2 \mathbb{E}[|X|^2] + \mathbb{E}[|X - Y|^2]) \leq C \mathbb{E}[|X - Y|^2], \end{aligned}$$

where the constant  $C$  depends also on  $R$ . This shows (36).

**Comments on the assumptions.** Let us recall that, as explained in the introduction, Assumption (35) implies that  $H$  is globally Lipschitz continuous in  $\mathcal{P}^2(\mathbb{R}^n)$ . Since (37) holds,  $H$  is semiconcave and semiconvex, in the sense, if  $\mu = [X]$  and  $\nu = [Y]$ , then

$$|H(\nu) - H(\mu) - \mathbb{E}[D_\mu H([X])(X) \cdot (Y - X)]| \leq C \mathbb{E}[|X - Y|^2]. \quad (38)$$

Indeed, we have

$$\begin{aligned}
& |H(\nu) - H(\mu) - \mathbb{E} [D_\mu H([X])(X) \cdot (Y - X)]| \\
&= \left| \int_0^1 \mathbb{E} [\{D_\mu H([(1-t)X + tY])((1-t)X + tY) - D_\mu H([X])(X)\} \cdot (Y - X)] dt \right| \\
&\leq \int_0^1 \mathbb{E}^{1/2} [|D_\mu H([(1-t)X + tY])((1-t)X + tY) - D_\mu H([X])(X)|^2] \mathbb{E}^{1/2} [|Y - X|^2]^{1/2} dt \\
&\leq C \mathbb{E} [|Y - X|^2],
\end{aligned}$$

where we used (37) in the last inequality. Moreover, under (37), for each  $\mu \in \mathcal{P}^2(\mathbb{R}^n)$ , there exists a Lipschitz continuous version of

$$D_\mu H(\mu)(\cdot) : \mathbb{R}^n \longrightarrow \mathbb{R}^n,$$

with a Lipschitz constant independent of  $\mu$  and such that  $(\mu, x) \mapsto D_\mu H(\mu)(x)$  is measurable and continuous at each point  $(\mu, x)$  such that  $x \in \text{Supp}(\mu)$  (Corollary 5.38 in [12]). Moreover,

$$\mathbb{E} [|D_\mu H(\mu)(X) - D_\mu H(\nu)(X)|^2] \leq C \mathbb{E} [|X - Y|^2].$$

As a byproduct (inequality (5.49) in [12]), for  $[X] = \mu$

$$\mathbb{E} [|D_\mu H(\mu)(X) - D_\mu H(\nu)(X)|^2] \leq C W_2^2(\mu, \nu). \quad (39)$$

**Definition 7.** By a solution to (3), we mean a couple of adapted and continuous processes  $(X, K)$ ,  $K$  being deterministic and non decreasing, with  $K_0 = 0$ .

**Theorem 8.** Under assumptions **(H)**, the SDE (3) has a unique square integrable solution.

*Remark 9.* The result can be easily generalized to the case where  $b$  and  $\sigma$  depend also on the law. In the proof we actually show that the process  $K$  is locally Lipschitz continuous.

*Proof.* Let us start by uniqueness. Suppose that  $(X, K^X)$  and  $(Y, K^Y)$  are two solutions of (3). Then, we deduce from Itô's formula that

$$\begin{aligned}
\mathbb{E} [|X_t - Y_t|^2] &= 2\mathbb{E} \left[ \int_0^t (X_s - Y_s) \cdot (b(s, X_s) - b(s, Y_s)) ds \right] + \mathbb{E} \left[ \int_0^t |\sigma(s, X_s) - \sigma(s, Y_s)|^2 ds \right] \\
&\quad + 2\mathbb{E} \left[ \int_0^t (X_s - Y_s) \cdot (D_\mu H([X_s])(X_s) dK_s^X - D_\mu H([Y_s])(Y_s) dK_s^Y) \right].
\end{aligned}$$

Since  $H$  satisfies (38), using the Skorokhod condition together with the constraint, we have

$$\begin{aligned}
\mathbb{E} \left[ \int_0^t (X_s - Y_s) \cdot D_\mu H([X_s])(X_s) dK_s^X \right] &\leq \int_0^t ((H([X_s]) - H([Y_s])) + C\mathbb{E} [|X_s - Y_s|^2]) dK_s^X \\
&= \int_0^t H([X_s]) dK_s^X - \int_0^t H([Y_s]) dK_s^X + \int_0^t C\mathbb{E} [|X_s - Y_s|^2] dK_s^X \\
&\leq \int_0^t C\mathbb{E} [|X_s - Y_s|^2] dK_s^X.
\end{aligned}$$

Similarly,

$$-\mathbb{E} \left[ \int_0^t (X_s - Y_s) \cdot D_\mu H([Y_s])(Y_s) dK_s^Y \right] \leq \int_0^t C\mathbb{E} [|X_s - Y_s|^2] dK_s^Y.$$

Using the fact that  $b$  and  $\sigma$  are Lipschitz continuous, we get, for some constant  $C > 0$ ,

$$\mathbb{E} [|X_t - Y_t|^2] \leq C \int_0^t \mathbb{E} [|X_s - Y_s|^2] (ds + dK_s^X + dK_s^Y),$$

and equality  $X = Y$  follows from Gronwall's lemma applied to the continuous maps  $s \mapsto K_s^X$ ,  $s \mapsto K_s^Y$  (see Lemma 4 in [21] or Theorem 17.1 in [2]) and  $s \rightarrow \mathbb{E} [|X_s - Y_s|^2]$ .

Next we show the equality  $K^X = K^Y$ . We get from Itô's formula (14) and for any  $t \geq 0$

$$\begin{aligned} A_t^X &:= \int_0^t \mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2] dK_u^X = H([X_t]) - H([X_0]) - \int_0^t \mathbb{E} [\mathrm{D}_\mu H([X_u])(X_u) \cdot b(u, X_u)] du \\ &\quad - \frac{1}{2} \int_0^t \mathbb{E} [\mathrm{Tr}(\partial_x \mathrm{D}_\mu H([X_u])(X_u) \sigma \sigma^*(u, X_u))] du \\ &= A_t^Y := \int_0^t \mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2] dK_u^Y. \end{aligned}$$

Thus the nondecreasing continuous functions  $A^X$  and  $A^Y$  are equal. For any step function  $\varphi$  on  $[0, T]$ ,

$$\begin{aligned} \int_0^T \varphi(u) dA_u^X &= \int_0^T \varphi(u) \mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2] dK_u^X \\ &= \int_0^T \varphi(u) \mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2] dK_u^Y = \int_0^T \varphi(u) dA_u^Y. \end{aligned} \quad (40)$$

The previous equality hold true also for any continuous function on  $[0, T]$  as uniform limit of step functions. Since  $H([X_u]) = 0$  for  $dK^X + dK^Y$ -a.e.  $u \in [0, T]$ , we have

$$\mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2] \geq \beta^2 > 0 \quad \text{for } dK^X + dK^Y \text{-a.e. } u \in [0, T].$$

As a byproduct, we can extend the map

$$u \mapsto \varphi(u) = \mathbb{E} [|\mathrm{D}_\mu H([X_u])(X_u)|^2]^{-1}$$

from the support of  $dK^X + dK^Y$  into a continuous map on  $[0, T]$  and we get, plugging this  $\varphi$  in (40),  $K_T^X = K_T^Y$ . This completes the proof of the uniqueness of the solution:  $(X, K^X) = (Y, K^Y)$ .

Let us now construct a solution to (3) on a time interval  $[0, T]$  (for  $T > 0$  arbitrary). For this, let us consider, for  $k > 1/\eta$ , the solution  $X^k$  to the following McKean-Vlasov SDE:

$$X_t^k = X_0 + \int_0^t b(s, X_s^k) ds + \int_0^t \sigma(s, X_s^k) dB_s + \int_0^t \mathrm{D}_\mu H([X_s^k])(X_s^k) \psi_k(s, H([X_s^k])) ds, \quad t \geq 0, \quad (41)$$

where the function  $\psi_k : \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  is defined as

$$\psi_k(t, x) = r(t) \text{ if } x \leq -1/k, \quad \psi_k(t, x) = -kr(t)x, \text{ if } -1/k \leq x \leq 0, \quad \psi_k(t, x) = 0, \text{ if } x \geq 0$$

and  $t \rightarrow r(t)$  is a continuous, positive and increasing map to be chosen later. We set

$$K_t^k = \int_0^t \psi_k(s, H([X_s^k])) ds.$$

Let us start with some  $L^2$  bounds. In order to use condition (36), we introduce the deterministic time

$$T_k = \inf\{t \geq 0, H([X_t^k]) \leq -\eta\},$$

with the convention that  $T_k = T$  if the right-hand side is empty. Note that  $T_k > 0$  since  $H([X_0]) \geq 0$ . Let us compute  $|X_t^k|^2$  for  $t \in [0, T_k]$  with the help of Itô's formula. We have

$$\begin{aligned} |X_t^k|^2 &= |X_0|^2 + 2 \int_0^t X_s^k \cdot b(s, X_s^k) ds + \int_0^t \mathrm{Tr}(\sigma \sigma^*(s, X_s^k)) ds \\ &\quad + 2 \int_0^t X_s^k \cdot \sigma(s, X_s^k) dB_s + 2 \int_0^t X_s^k \cdot \mathrm{D}_\mu H([X_s^k])(X_s^k) dK_s^k. \end{aligned} \quad (42)$$

Since  $b$  and  $\sigma$  are Lipschitz continuous in space, uniformly in time, we deduce from this equality that there exists a constant  $C = C(L_T)$  such that

$$|X_t^k|^2 \leq |X_0|^2 + C \int_0^t (1 + |b(s, 0)|^2 + |X_s^k|^2) ds + 2 \int_0^t X_s^k \cdot \sigma(s, X_s^k) dB_s + 2 \int_0^t X_s^k \cdot \mathrm{D}_\mu H([X_s^k])(X_s^k) dK_s^k. \quad (43)$$

Using Cauchy-Schwarz's inequality, Assumption (35) and then Young's inequality, we obtain,

$$\begin{aligned} & \mathbb{E} \left[ \int_0^t X_s^k \cdot D_\mu H([X_s^k])(X_s^k) dK_s^k \right] \\ & \leq \int_0^t (\mathbb{E} [|X_s^k|^2])^{1/2} \left( \mathbb{E} [|D_\mu H([X_s^k])(X_s^k)|^2] \right)^{1/2} \psi_k(s, H([X_s^k])) ds \\ & \leq \int_0^t (\mathbb{E} [|X_s^k|^2])^{1/2} M r(s) ds \leq \int_0^t (\mathbb{E} [|X_s^k|^2] + M^2 r^2(s)) ds. \end{aligned}$$

Hence

$$\mathbb{E} [|X_t^k|^2] \leq \mathbb{E} [|X_0|^2] + C \int_0^t (1 + M^2 r^2(s) + \mathbb{E} [|X_s^k|^2]) ds.$$

It follows from Gronwall's inequality that there exists  $C_0 > 0$ , and, for any  $t \in [0, T]$ ,  $C_1(T) > 0$ , depending on  $T$ ,  $\mathcal{M}_2([X_0])$ ,  $L_T$  and  $\sup_{t \leq T} \mathbb{E} [|b(t, 0)|^2]$  only, such that

$$\sup_{s \leq t} \mathbb{E} [|X_s^k|^2] \leq C_1(T) + M^2 \int_0^t e^{C_0(t-s)} r^2(s) ds. \quad (44)$$

Coming back to (43), since  $\psi_k$  is bounded by  $r$ , the upper bound of Assumption (35) together with BDG's inequality give

$$\mathbb{E} \left[ \sup_{t \leq T} |X_t^k|^2 \right] \leq C \left( T, L_T, \mathcal{M}_2([X_0]), \sup_{t \leq T} \mathbb{E} [|b(t, 0)|^2], M, r(\cdot) \right).$$

Let us apply Itô's formula (12) to compute  $H([X_t^k])$ . We get, for  $0 \leq s \leq t$ , such that  $-\eta \leq H([X_u^k]) \leq -1/k$  on  $[s, t]$ ,

$$\begin{aligned} H([X_t^k]) &= H([X_s^k]) + \int_s^t \mathbb{E} [|D_\mu H([X_u^k])(X_u^k)|^2] \psi_k(u, [X_u^k]) du \\ & \quad + \int_s^t \mathbb{E} [D_\mu H([X_u^k])(X_u^k) \cdot b(u, X_u^k)] du + \frac{1}{2} \int_s^t \mathbb{E} [\text{Tr} (\partial_x D_\mu H([X_u^k])(X_u^k) \sigma \sigma^*(u, X_u^k))] du \\ & \geq H([X_s^k]) + \beta^2 \int_s^t r(u) du - C \int_s^t \left( 1 + (\mathbb{E} [|X_u^k|^2])^{1/2} \right) du, \end{aligned} \quad (45)$$

for some constant  $C \geq 0$  independent of  $k$  and  $r$ . In the above inequality, we used the fact that  $x \mapsto D_\mu H(\mu)(x)$  is Lipschitz uniformly in  $\mu$ , that  $\sigma$  is bounded, that  $\beta^2 \leq \mathbb{E} [|D_\mu H([X])(X)|^2] \leq M^2$  for any  $X$  such that  $-\eta \leq H([X]) \leq 0$  and that  $b$  has a linear growth while  $\sigma$  is bounded. We deduce from (44) that,

$$\begin{aligned} H([X_t^k]) &\geq H([X_s^k]) + \int_s^t \left( \beta^2 r(u) - C \left( 1 + \left( C_1(T) + M^2 \int_0^u e^{C_0(u-v)} r^2(v) dv \right)^{1/2} \right) \right) du \\ &\geq H([X_s^k]) + \int_s^t \left( \beta^2 r(u) - C \left( 1 + C_1(T) + M^2 e^{C_0 u} \int_0^u r^2(v) dv \right)^{1/2} \right) du. \end{aligned} \quad (46)$$

We can then choose the map  $t \rightarrow r(t)$ , independent of  $k$ , such that  $\rho(t) := \int_0^t r^2(v) dv$  satisfies the ODE

$$\rho'(u) = C^2 \beta^{-4} (1 + C_1(T) + M^2 e^{C_0 u} \rho(u)), \quad \forall u \geq 0,$$

so that

$$r^2(u) = C^2 \beta^{-4} \left( 1 + C_1(T) + M^2 e^{C_0 u} \int_0^u r^2(v) dv \right), \quad \forall u \geq 0. \quad (47)$$

With this choice of  $r$ , we claim that

$$\forall k \geq 1, \quad \forall t \leq T_k, \quad H([X_t^k]) \geq -\frac{1}{k}. \quad (48)$$

Indeed, if  $H([X_{t_0}^k]) < -1/k$  for some  $t_0 > 0$ , let

$$s = \sup \{ u \leq t_0 : H([X_u^k]) \geq -1/k \} \text{ and } t = \inf \{ u \geq s : H([X_u^k]) \leq H([X_{t_0}^k]) \vee (-\eta) \}.$$

Since  $H([X_0]) \geq 0$ , we have  $0 < s < t \leq t_0$  and  $-\eta \leq H([X_u^k]) \leq -1/k$  for  $s \leq u \leq t$ . Hence, by (46) and (47), we get

$$H([X_{t_0}^k]) \vee (-\eta) = H([X_t^k]) \geq H([X_s^k]) = -\frac{1}{k}.$$

This is a contradiction and (48) holds. Note that, with this choice of  $r(\cdot)$ , we have  $T_k = T$  and

$$\sup_k K_t^k \leq C(t) \quad \forall t \in [0, T]. \quad (49)$$

Let  $k, m \in \mathbb{N}^*$ , with  $k, m > 1/\eta$ . By Itô's formula we have, for  $t \geq 0$ ,

$$\begin{aligned} |X_t^k - X_t^m|^2 &= 2 \int_0^t (X_s^k - X_s^m) \cdot (b(s, X_s^k) - b(s, X_s^m)) ds \\ &\quad + 2 \int_0^t \text{Tr}((\sigma(s, X_s^k) - \sigma(s, X_s^m))(\sigma(s, X_s^k) - \sigma(s, X_s^m))^*) ds \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (\sigma(s, X_s^k) - \sigma(s, X_s^m)) dB_s \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (D_\mu H([X_s^k])(X_s^k) dK_s^k - D_\mu H([X_s^m])(X_s^m) dK_s^m). \end{aligned}$$

Using the fact that  $b$  and  $\sigma$  are  $L_T$ -Lipschitz continuous, we get

$$\begin{aligned} |X_t^k - X_t^m|^2 &\leq (2L_T + L_T^2) \int_0^t |X_s^k - X_s^m|^2 ds + 2 \int_0^t (X_s^k - X_s^m) \cdot (\sigma(s, X_s^k) - \sigma(s, X_s^m)) dB_s \\ &\quad + 2 \int_0^t (X_s^k - X_s^m) \cdot (D_\mu H([X_s^k])(X_s^k) dK_s^k - D_\mu H([X_s^m])(X_s^m) dK_s^m). \end{aligned} \quad (50)$$

Arguing as in the proof of uniqueness, taking the expectation, we have, since  $H$  satisfies (38),

$$\begin{aligned} &\mathbb{E} \left[ \int_0^t (X_s^k - X_s^m) \cdot (D_\mu H([X_s^k])(X_s^k) dK_s^k - D_\mu H([X_s^m])(X_s^m) dK_s^m) \right] \\ &\leq \int_0^t (H([X_s^k]) - H([X_s^m])) (dK_s^k - dK_s^m) + C \int_0^t \mathbb{E} [|X_s^k - X_s^m|^2] (dK_s^k + dK_s^m) \\ &= \int_0^t (H([X_s^k]) - H([X_s^m])) (\psi_k(s, H([X_s^k])) - \psi_m(s, H([X_s^m]))) ds + C \int_0^t \mathbb{E} [|X_s^k - X_s^m|^2] (dK_s^k + dK_s^m). \end{aligned}$$

From (48), since  $x\psi_k(t, x) \leq 0$ ,  $H([X_t^k]) \geq -1/k$  and  $\psi_k(s, H([X_s^k]))$  and  $\psi_m(s, H([X_s^m]))$  are in  $[0, r(t)]$  for  $s \in [0, t]$ , we obtain that

$$\begin{aligned} &\mathbb{E} \left[ \int_0^t (X_s^k - X_s^m) \cdot (D_\mu H([X_s^k])(X_s^k) dK_s^k - D_\mu H([X_s^m])(X_s^m) dK_s^m) \right] \\ &\leq t \left( \frac{r(t)}{k} + \frac{r(t)}{m} \right) + C \int_0^t \mathbb{E} [|X_s^k - X_s^m|^2] (dK_s^k + dK_s^m). \end{aligned}$$

Therefore

$$\mathbb{E} [|X_t^k - X_t^m|^2] \leq C \int_0^t \mathbb{E} [|X_s^k - X_s^m|^2] (ds + dK_s^k + dK_s^m) + t \left( \frac{r(t)}{k} + \frac{r(t)}{m} \right).$$

We conclude using Gronwall's lemma and the bound on  $(K^k)$  in (49) that, for any  $T > 0$ ,

$$\sup_{t \in [0, T]} \mathbb{E} [|X_t^k - X_t^m|^2] \leq T \left( \frac{r(T)}{k} + \frac{r(T)}{m} \right) \exp\{C(T + K_T^k + K_T^m)\} \leq C \left( \frac{1}{k} + \frac{1}{m} \right).$$

Coming back to (50), we derive, from (34), (49), the BDG and Cauchy-Schwarz inequalities, the estimate

$$\mathbb{E} \left[ \sup_{t \leq T} |X_t^k - X_t^m|^2 \right] \leq C \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{m}} \right),$$

where the constant  $C$  does not depend on  $k$  and  $m$ . Thus,  $\{X^k\}_{k \geq 1}$  is a Cauchy sequence in  $\mathcal{S}^2$ . Let us denote by  $X$  its limit. Since  $\psi_k$  is bounded on  $[0, T]$  by  $r(T)$  for all  $k$ ,  $K^k$  is Lipschitz continuous with

a Lipschitz constant bounded by  $r(T)$  on  $[0, T]$ . By Ascoli-Arzelà theorem, up to a subsequence,  $(K^k)$  converges locally uniformly to a non decreasing, locally Lipschitz continuous function  $K$ .

It is straightforward to check that  $(X, K)$  solves (3). Indeed,

- $H([X_t]) = \lim_{k \rightarrow \infty} H([X_t^k]) \geq 0$  by (48),
- the Skorokhod condition is also satisfied : since  $x\psi_k(t, x) \leq 0$ ,

$$0 \leq \int_0^t H([X_s]) dK_s = \lim_{k \rightarrow \infty} \int_0^t H([X_s^k]) \psi_k(s, H([X_s^k])) ds \leq 0.$$

□

### 3. SDEs with normal reflexion in conditional law.

In this section, we consider the reflected SDE:

$$\begin{cases} X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma_0(s, X_s) dB_s + \int_0^t \sigma_1(s, X_s) dW_s + \int_0^t D_\mu H([X_s|W])(X_s) dK_s, \\ H([X_t|W]) \geq 0, \quad \int_0^t H([X_s|W]) dK_s = 0, \quad t \geq 0, \end{cases} \quad (51)$$

where now  $B$  and  $W$  are independent Brownian motions and  $(K_s)$  is a continuous nondecreasing process adapted to the filtration  $\mathcal{F}^W$  associated with  $W$ . As before,  $H$  is a map from  $\mathcal{P}^2(\mathbb{R}^n)$  to  $\mathbb{R}$ . The notation  $[X_t|W]$  stands for the conditional probability of  $X_t$  given  $W$ . We assume that the initial condition of the process,  $X_0$ , is independent of  $B$  and of  $W$  and in  $L^2$ . As explained below, this reflected process is the limit of a reflected particle system with a common noise.

We assume in this part that the following conditions **(Hcl)** hold:

- (HΩ)**: The probability space is  $(\Omega, \mathbb{P}) = (\Omega^0 \times \Omega^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$ , where  $\Omega^0$  supports the  $X_0$  and  $B$ , while  $\Omega^1$  supports  $W$  with associated filtration  $\mathcal{F}^W = \mathcal{F}^1$ .
- (Hc)**: The functions  $b : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\sigma_0, \sigma_1 : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$  are measurable with respect to  $\mathcal{E} \otimes \mathcal{B}(\mathbb{R}^n)$  and
- For all  $T > 0$ , there exists  $L_T$  such that,  $\mathbb{P}$ -a.s., for each  $t \in [0, T]$ ,
 
$$\forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n, \quad |b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq L_T |x - y|;$$
  - $b, \sigma_0, \sigma_1$  are globally bounded: for all  $T > 0$ , there exists  $C_T$  such that,  $\mathbb{P}$ -a.s.,

$$\sup_{t \leq T, x \in \mathbb{R}^n} \{|b(t, x)| + |\sigma_0(t, x)| + |\sigma_1(t, x)|\} \leq C_T.$$

**(H0)**: The initial condition is  $X_0$  is independent of  $B$  and  $W$ , in  $L^2(\Omega^0)$  and with  $H([X_0]) \geq 0$ ;

**(HH)**: The function  $H : \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  is globally  $\mathcal{C}^2$  and

- there exists  $M > 0$  such that:  $\forall \mu \in \mathcal{P}^2(\mathbb{R}^n)$ ,

$$\begin{aligned} |H(\mu)| + \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx) + \int_{\mathbb{R}^n} |\partial_y D_\mu H(\mu)(y)| \mu(dy) \\ + \int_{\mathbb{R}^n \times \mathbb{R}^n} |D_{\mu\mu}^2 H(\mu)(x, y)| \mu(dx) \mu(dy) \leq M^2, \end{aligned} \quad (52)$$

- there exist  $\beta > 0$  and  $\eta > 0$  such that

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n) \text{ with } -\eta \leq H(\mu) \leq 0, \quad \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx) \geq \beta^2, \quad (53)$$

- there exists  $C_1 \geq 0$  such that

$$\forall X, Y \in L_{\mathbb{P}}^2, \quad \mathbb{E} [ |D_\mu H([X])(X) - D_\mu H([Y])(Y)|^2 ] \leq C_1 \mathbb{E} [|X - Y|^2]. \quad (54)$$

**Examples.** They are the same as in the previous section. For the first example, let  $f_1 : \mathbb{R}^n \rightarrow \mathbb{R}$  be of class  $\mathcal{C}^2$  and such that the set  $\{f_1 = 0\}$  is compact and 0 is a noncritical value of  $f_1$ . We also assume, without loss of generality, that  $f_1$  is bounded and that  $\nabla f_1$  and  $\nabla^2 f_1$  have compact support. We set

$$H_1(\mu) = f_1 \left( \int_{\mathbb{R}^n} y \mu(dy) \right) \quad \forall \mu \in \mathcal{P}^2(\mathbb{R}^n).$$

Then

$$D_\mu H_1(\mu)(x) = \nabla f_1 \left( \int_{\mathbb{R}^n} y \mu(dy) \right) \text{ and } D_{\mu\mu}^2 H_1(\mu)(x, x') = \nabla^2 f_1 \left( \int_{\mathbb{R}^n} y \mu(dy) \right).$$

So conditions **(Hcl)** hold as in the previous section.

For the second example, we assume that  $H_2$  is given by

$$H_2(\mu) = f_2 \left( \int_{\mathbb{R}^n} x x^* \mu(dx) \right) \quad \forall \mu \in \mathcal{P}^2(\mathbb{R}^n),$$

where  $f_2$  is as in the previous section. Then one can show with the same argument as in the previous section that  $H_2$  is bounded and that its first and second order derivatives have bounded support and that (53) holds. This easily implies that  $H_2$  satisfies conditions **(Hcl)**.

### 3.1. Existence and uniqueness of the solution.

**Definition 10.** We say that  $(X, K)$  is a solution to (51) if  $(X, K)$  is a continuous and progressively measurable process such that  $X_t$  is square integrable for any  $t \geq 0$  and such that  $K$  is nondecreasing and adapted to the filtration  $\mathcal{F}^W$  with  $K_0 = 0$ .

Note that, under the above assumptions, the equation has a meaning since (denoting by  $\mathbb{E}^0$  the conditional expectation with respect to  $W$ ) we have, for any  $t \geq 0$  and  $\mathbb{P}^1$ -a.s.,

$$\begin{aligned} \mathbb{E}^0 \left[ \int_0^t |D_\mu H([X_s|W])(X_s)| dK_s \right] &= \int_0^t \mathbb{E}^0 [|D_\mu H([X_s|W])(X_s)|] dK_s \\ &\leq \left( \sup_{\mu \in \mathcal{P}^2} \int_{\mathbb{R}^d} |D_\mu H(\mu)(x)|^2 \mu(dx) \right)^{1/2} K_t < +\infty. \end{aligned}$$

**Theorem 11.** Under assumptions **(Hcl)**, the SDE (51) has a unique square integrable solution. In addition, we have, for any  $T > 0$  and for any  $\theta > 0$ ,

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_s|^2 \right] \leq C(T), \quad \mathbb{E} \left[ \exp\{\theta K_T\} \right] \leq C_\theta(T), \quad (55)$$

for some constants  $C(T)$  and  $C_\theta(T)$  depending on the data,  $T$  and, for  $C_\theta(T)$ ,  $\theta$ .

*Remark 12.* We emphasize that weak uniqueness also holds for (3) and for (51). We explain why at the end of the proof of Theorem 11.

*Remark 13.* In contrast with the previous section (without ‘‘common noise’’) the process  $K$  is neither deterministic nor Lipschitz continuous in general. For this reason, the assumptions under which we work here are much stronger than the ones in Section 2.

From the proof of the Theorem (see estimate (66) below), we also deduce the following result.

**Corollary 14.** Let  $p \geq 2$  and  $T > 0$ . Assume that in addition to **(Hcl)** we have  $M_p([X_0]) + \mathcal{H}_p < +\infty$ . Then, there exists a positive constant  $C_p(T) := C((p, T, M_p([X_0]), \mathcal{H}_p)) > 0$  such that

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} |\mathbb{E}^0 [|X_s|^2]|^p \right] \leq C_p(T). \quad (56)$$

One can also show that the process depends in a continuous way of the initial condition:

**Proposition 15.** Let  $X_0$  and  $Y_0$  be two initial conditions with  $H([X_0]) \geq 0$  and  $H([Y_0]) \geq 0$ . Let  $(X, K^X)$  and  $(Y, K^Y)$  be solutions of (51) with initial conditions  $X_0$  and  $Y_0$ , respectively. Then, for any  $T > 0$ ,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathbb{E}^0 [|X_t - Y_t|^2] \right] \leq C \mathbb{E} [|X_0 - Y_0|^2]$$

while

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t - Y_t|^2 \right] \leq C \left( \mathbb{E} [|X_0 - Y_0|^2] + \mathbb{E} [|X_0 - Y_0|^2]^{1/2} \right).$$



Eventually, we have the following regularity estimate on the continuity of the path  $K$  and  $X$ .

**Proposition 16.** *Let  $(X, K)$  be the unique solution of SDE (51). Assume in addition to **(Hcl)** that there exists  $p \geq 2$  such that  $\mathcal{M}_p([X_0]) + \mathcal{H}_p < +\infty$ .*

*Then, there exists a constant  $C_p(T) := C((p, T, \mathcal{M}_p([X_0]), \mathcal{H}_p) > 0$  such that*

$$\mathbb{E}[|K_t - K_s|^p] \leq C_{p,T}|t - s|^{p/4}, \quad (57)$$

and if moreover  $p$  is in  $\mathbb{N}$

$$\mathbb{E}[|X_t - X_s|^p] + \mathbb{E}\left[\sup_{u \in [s, t]} |\mathbb{E}^0[|X_u - X_s|^2]|^{p/2}\right] \leq C_{p,T}|t - s|^{p/4}. \quad (58)$$

*Proof of Theorem 11: Existence under an additional moment condition.* The structure of proof is roughly the same as for Theorem 8. However, because the process  $K$  is no longer Lipschitz continuous nor deterministic, we have to pay extra attention in the various estimates. Here we address the existence of the solution under the additional condition that  $X_0$  has a moment of order 4:  $\mathbb{E}[|X_0|^4] < +\infty$ . This extra condition is removed after the proof of Proposition 15. The uniqueness is a straightforward consequence of Proposition 15.

**Step 1: Approximate solutions.** We build the solution on the time interval  $[0, T]$  (for an arbitrary  $T > 0$ ). All the constants below depend on this horizon  $T$ . As before we argue by penalization. We consider, for  $k \geq 1/\eta$ ,  $X^k$  solution to the following McKean-Vlasov SDE:

$$\begin{aligned} X_t^k &= X_0 + \int_0^t b(s, X_s^k) ds + \int_0^t \sigma_0(s, X_s^k) dB_s + \int_0^t \sigma_1(s, X_s^k) dW_s \\ &\quad + \int_0^t D_\mu H([X_s^k|W])(X_s^k) \psi_k(H([X_s^k|W])) ds, \quad t \geq 0, \end{aligned} \quad (59)$$

where the function  $\psi_k$  is defined as

$$\psi_k(x) = k^2 \text{ if } x \leq -1/k, \quad \psi_k(x) = -k^3 x, \text{ if } -1/k \leq x \leq 0, \quad \psi_k(x) = 0, \text{ if } x \geq 0.$$

We set  $\mu_t^k = [X_t^k|W]$  and  $K_t^k = \int_0^t \psi_k(H([X_s^k|W])) ds = \int_0^t \psi_k(H(\mu_s^k)) ds$  and define the  $\mathcal{F}^W$ -stopping time  $\tau^k$  by

$$\tau^k := T \wedge \inf\{t \geq 0, H(\mu_t^k) \leq -\eta\}.$$

We note for a later use that  $\mu_{t \wedge \tau^k}^k = [X_{t \wedge \tau^k}^k|W]$  since  $\tau^k$  is an  $\mathcal{F}^W$ -stopping time. We also stress that  $H(\mu_s^k) \in [-1/k, 0]$  for  $dK^k$ -a.e.  $s \in [0, \tau^k]$ , so that, by (53) and for any nonnegative process  $\varphi_s$  adapted to  $\mathcal{F}^W$ ,

$$\int_0^{t \wedge \tau^k} \varphi_s \mathbb{E}^0 \left[ |D_\mu H(\mu_s^k)(X_s^k)|^2 \right] dK_s^k \geq \beta^2 \int_0^{t \wedge \tau^k} \varphi_s dK_{s \wedge \tau^k}^k. \quad (60)$$

We finally note that (59) can be rewritten as

$$\begin{aligned} X_t^k &= X_0 + \int_0^t b(s, X_s^k) ds + \int_0^t \sigma_0(s, X_s^k) dB_s + \int_0^t \sigma_1(s, X_s^k) dW_s \\ &\quad + \int_0^t D_\mu H([X_s^k|W])(X_s^k) dK_s^k, \quad t \geq 0. \end{aligned}$$

**Step 2: Estimate on  $K^k$ .** For  $0 \leq s \leq t \leq T$  and by Itô's formula (13), we have

$$\begin{aligned}
 H(\mu_{t \wedge \tau^k}^k) &= H(\mu_{s \wedge \tau^k}^k) + \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [D_\mu H(\mu_u^k)(X_u^k) \cdot b(u, X_u^k)] du \\
 &\quad + \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [ |D_\mu H(\mu_u^k)(X_u^k)|^2 ] dK_u^k \\
 &\quad + \frac{1}{2} \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [ \text{Tr} ( \partial_x D_\mu H(\mu_u^k)(X_u^k) (a_1(u, X_u^k) + a_2(u, X_u^k)) ) ] du \\
 &\quad + \frac{1}{2} \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 \tilde{\mathbb{E}}^0 [ \text{Tr} ( D_{\mu\mu}^2 H(\mu_u^k)(X_u^k, \tilde{X}_u^k) \sigma_1(u, X_u^k) \sigma_1^*(u, \tilde{X}_u^k) ) ] du \\
 &\quad + \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [ \sigma_1^*(u, X_u^k) D_\mu H(\mu_u^k)(X_u^k) ] \cdot dW_u. \tag{61}
 \end{aligned}$$

We use (60) to bound below the term  $\int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [ |D_\mu H(\mu_u^k)(X_u^k)|^2 ] dK_u^k$ . We also use our bounds on  $H$  in (52) and the  $L^\infty$  bounds on  $b$  and  $\sigma_i$  to estimate the other terms and obtain:

$$\begin{aligned}
 H(\mu_{t \wedge \tau^k}^k) &\geq H(\mu_{s \wedge \tau^k}^k) - C(t \wedge \tau^k - s \wedge \tau^k) + \beta^2 (K_{t \wedge \tau^k}^k - K_{s \wedge \tau^k}^k) \\
 &\quad + \int_{s \wedge \tau^k}^{t \wedge \tau^k} \mathbb{E}^0 [ \sigma_1^*(u, X_u^k) D_\mu H(\mu_u^k)(X_u^k) ] \cdot dW_u. \tag{62}
 \end{aligned}$$

Taking  $s = 0$  and  $t = T$  and using once again the fact that  $\sigma_1$  is bounded and (52), we infer that the  $K^k$  have uniformly bounded exponential moments:

$$\sup_k \mathbb{E} [ \exp\{\theta K_{\tau^k}^k\} ] \leq C(\theta), \quad \forall \theta > 0. \tag{63}$$

**Step 3: Some estimates on the  $(p)$ -moments of  $X_t^k$ .** We claim that:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau^k} |X_t^k|^2 \right] \leq C_{2,T} (\mathbb{E} [|X_0|^2] + 1); \tag{64}$$

and, assuming in addition to the current assumption that there exists  $p \geq 2$  such that  $\mathcal{M}_p([X_0]) + \mathcal{H}_p < +\infty$ , it holds:

$$\mathbb{E} \left[ \sup_{0 \leq t \leq \tau^k} \mathbb{E}^0 [ |X_t^k|^p ] \right] \leq C_{p,T}, \tag{65}$$

$$\mathbb{E} \left[ (\mathbb{E}^0)^{p/2} \left[ \sup_{0 \leq t \leq \tau^k} |X_t^k|^2 \right] \right] \leq C_{p,T}, \tag{66}$$

where  $C_{p,T} := C_{p,T}(\mathcal{M}_p([X_0]) + \mathcal{H}_p) > 0$ .

Let us set  $\gamma_t^k := \exp\{-\alpha(\delta t + H(\mu_t^k))\}$  for  $\alpha, \delta \geq 1$  to be chosen below. We note for later use that, because of our boundedness assumption on  $H$ ,  $\gamma_t^k$  is bounded above and below by positive constants. We

have, by Itô's formula (13),

$$\begin{aligned}
\gamma_{t \wedge \tau^k}^k |X_{t \wedge \tau^k}^k|^p &= \gamma_0^k |X_0|^p + \int_0^{t \wedge \tau^k} \gamma_s^k p |X_s^k|^{p-2} X_s^k \cdot b(s, X_s^k) ds \\
&+ \frac{1}{2} \int_0^{t \wedge \tau^k} \gamma_s^k \text{Tr}[a(s, X_s^k)(p |X_s^k|^{p-2} I_d + p(p-2) |X_s^k|^{p-4} X_s^k \otimes X_s^k)] ds \\
&+ \int_0^{t \wedge \tau^k} p \gamma_s^k \sigma_0^*(s, X_s^k) |X_s^k|^{p-2} X_s^k \cdot dB_s + \int_0^{t \wedge \tau^k} p \gamma_s^k \sigma_1^*(s, X_s^k) |X_s^k|^{p-2} X_s^k \cdot dW_s \\
&+ \int_0^{t \wedge \tau^k} p \gamma_s^k D_\mu H(\mu_s^k)(X_s^k) \cdot |X_s^k|^{p-2} X_s^k dK_s^k - \alpha \delta \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p ds \\
&- \alpha \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \mathbb{E}^0 [D_\mu H(\mu_s^k)(X_s^k) \cdot b(s, X_s^k)] ds \\
&- \alpha \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] \cdot dW_s \\
&- \frac{\alpha}{2} \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \mathbb{E}^0 [\text{Tr}(a_0(s, X_s^k) \partial_y D_\mu H(\mu_s^k)(X_s^k))] ds \\
&- \frac{\alpha}{2} \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \mathbb{E}^0 \tilde{\mathbb{E}}^0 [\text{Tr}(D_{\mu\mu}^2 H(\mu_s^k)(X_s^k, \tilde{X}_s^k) \sigma_1(s, X_s^k) \sigma_1^*(s, \tilde{X}_s^k))] ds \\
&- p \alpha \int_0^{t \wedge \tau^k} \gamma_s^k \text{Tr}(\sigma_1^*(s, X_s^k) |X_s^k|^{p-2} X_s^k (\mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)]^*)) ds \\
&+ \frac{\alpha^2}{2} \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \text{Tr}(\mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] (\mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)]^*)) ds \\
&- \alpha \int_0^{t \wedge \tau^k} \gamma_s^k |X_s^k|^p \mathbb{E}^0 [ |D_\mu H(\mu_s^k)(X_s^k)|^2 ] dK_s^k. \tag{67}
\end{aligned}$$

From Young's inequality, we have

$$\begin{aligned}
&\mathbb{E}^0 \left[ \int_0^{t \wedge \tau^k} p \gamma_s^k D_\mu H(\mu_s^k)(X_s^k) \cdot |X_s^k|^{p-2} X_s^k dK_s^k \right] \\
&\leq \int_0^{t \wedge \tau^k} p \gamma_s^k \left( \mathbb{E}^0 \left[ \frac{p-1}{p} |X_s^k|^p \right] + \mathbb{E}^0 \left[ \frac{1}{p} |D_\mu H(\mu_s^k)(X_s^k)|^p \right] \right) dK_s^k. \tag{68}
\end{aligned}$$

By (52), the terms  $\mathbb{E}^0 [ |D_\mu H(\mu_s^k)(X_s^k)| ]$ ,  $\mathbb{E}^0 [ |\partial_y D_\mu H(\mu_s^k)(X_s^k)| ]$  and  $\mathbb{E}^0 \tilde{\mathbb{E}}^0 [ |D_{\mu\mu}^2 H(\mu_s^k)(X_s^k, \tilde{X}_s^k)| ]$  are a.s. bounded. So, by (60) applied to the last term and for  $\alpha$  large enough (to absorb partially the first term in  $dK_s^k$ ) and then  $\delta$  large enough (to absorb partially all the terms in  $ds$ ), we have

$$\begin{aligned}
\mathbb{E}^0 [\gamma_{t \wedge \tau^k}^k |X_{t \wedge \tau^k}^k|^p] &\leq \mathbb{E}^0 [\gamma_0^k |X_0|^p] + C_{\alpha, \delta} T + C \mathcal{H}_p K_T^k \\
&+ \int_0^{t \wedge \tau^k} p \gamma_s^k \mathbb{E}^0 [\sigma_1^*(s, X_s^k) |X_s^k|^{p-2} X_s^k] \cdot dW_s \\
&- \alpha \int_0^{t \wedge \tau^k} \mathbb{E}^0 [\gamma_s^k |X_s^k|^p] \mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] \cdot dW_s. \tag{69}
\end{aligned}$$

Taking expectation and using the bound on  $K^k$  in (63) we find

$$\sup_{t \in [0, T]} \mathbb{E} [ |X_{t \wedge \tau^k}^k|^p ] \leq C(\mathbb{E} [|X_0|^p] + \mathcal{H}_p) + C(T). \tag{70}$$

We come back to (69) and use the boundedness of  $K^k$  in (63) and BDG inequality to find

$$\begin{aligned}
 & \mathbb{E} \left[ \sup_{s \leq T} \mathbb{E}^0 [\gamma_{t \wedge \tau^k}^k |X_{t \wedge \tau^k}^k|^p] \right] \\
 & \leq C \mathbb{E} [|X_0|^p] + C_{\alpha, \delta}(T, \mathcal{H}_p) + C \mathbb{E} \left[ \left( \int_0^T (\gamma_{s \wedge \tau^k}^k)^2 (\mathbb{E}^0)^2 [|X_{s \wedge \tau^k}^k|^p] ds \right)^{1/2} \right] \\
 & \leq C \mathbb{E} [|X_0|^p] + C_{\alpha, \delta}(T, \mathcal{H}_p) + C \mathbb{E} \left[ \left( \sup_{s \leq T} \mathbb{E}^0 [\gamma_{t \wedge \tau^k}^k |X_{s \wedge \tau^k}^k|^p] \right)^{1/2} \left( \int_0^T \gamma_{t \wedge \tau^k}^k \mathbb{E}^0 [|X_{s \wedge \tau^k}^k|^p] ds \right)^{1/2} \right] \\
 & \leq C \mathbb{E} [|X_0|^p] + C_{\alpha, \delta}(T, \mathcal{H}_p) + \frac{1}{2} \mathbb{E} \left[ \sup_{s \leq T} \mathbb{E}^0 [\gamma_{t \wedge \tau^k}^k |X_{s \wedge \tau^k}^k|^p] \right] + C \mathbb{E} \left[ \int_0^T \mathbb{E}^0 [\gamma_{t \wedge \tau^k}^k |X_{s \wedge \tau^k}^k|^p] ds \right]. \quad (71)
 \end{aligned}$$

Using (70), this proves (65).

Now, from Itô's formula on  $|X_{t \wedge \tau^k}^k|^2$  we obtain an expression similar to (67) (with  $p = 2$  therein and  $\alpha = \delta = 0$ ) without the last eight terms. Taking first the supremum, then the expectation  $\mathbb{E}$  and using the fact that

$$\begin{aligned}
 & \mathbb{E}^1 \otimes \mathbb{E}^0 \left[ \sup_{s \leq t} \int_0^{s \wedge \tau^k} |D_\mu H(\mu_s^k)(X_s^k)| |X_s^k| dK_s^k \right] \\
 & \leq \mathbb{E}^1 \left[ \int_0^{T \wedge \tau^k} (\mathbb{E}^0)^{1/2} [|D_\mu H(\mu_s^k)(X_s^k)|^2] (\mathbb{E}^0)^{1/2} [|X_s^k|^2] dK_s^k \right] \\
 & \leq C(M) (\mathbb{E}^1)^{1/2} \left[ \sup_{0 \leq t \leq T \wedge \tau^k} \{ \mathbb{E}^0 [|X_t^k|^2] \} \right] (\mathbb{E}^1)^{1/2} [(K_T^k)^2],
 \end{aligned}$$

thanks to Cauchy-Schwarz inequality we deduce, from (65) with  $p = 2$ , estimates on the moment of  $K$  and BDG's inequality that (64) holds. Also, still working with the Itô's expansion of  $|X_{t \wedge \tau^k}^k|^2$ , we obtain, taking first the supremum of this expression, then the expectation  $\mathbb{E}^0$ , taking the  $p$  power of this expression, taking next the expectation  $\mathbb{E}$ , using BDG inequality, and Gronwall's lemma that (66) holds.

**Step 4: Estimate of the difference  $X^k - X^m$ , first part.** Let  $\tau^{k,m} := \tau^k \wedge \tau^m$  and let us set  $\gamma_t := \exp\{-\alpha(\delta t + H(\mu_t^k) + H(\mu_t^m))\}$  for  $\alpha, \delta \geq 1$  to be chosen below. By Itô's formula (13), by using the Lipschitz continuity of  $b$  and  $\sigma_i$  and absorbing the terms in  $dt$  as in Step 3, using the bounds on the derivatives of  $H$  and on  $\sigma_1$  as well as (60), we get, for  $\delta$  large enough:

$$\begin{aligned}
 & \gamma_{t \wedge \tau^{k,m}} |X_{t \wedge \tau^{k,m}}^k - X_{t \wedge \tau^{k,m}}^m|^2 \\
 & \leq \int_0^{t \wedge \tau^{k,m}} 2\gamma_s (X_s^k - X_s^m) \cdot D_\mu H(\mu_s^k)(X_s^k) dK_s^k - \int_0^{t \wedge \tau^{k,m}} 2\gamma_s (X_s^k - X_s^m) \cdot D_\mu H(\mu_s^m)(X_s^m) dK_s^m \\
 & + \int_0^{t \wedge \tau^{k,m}} 2\gamma_s (X_s^k - X_s^m) \cdot (\sigma_0(s, X_s^k) - \sigma_0(s, X_s^m)) dB_s \\
 & + \int_0^{t \wedge \tau^{k,m}} 2\gamma_s (X_s^k - X_s^m) \cdot (\sigma_1(s, X_s^k) - \sigma_1(s, X_s^m)) dW_s - \frac{\alpha\delta}{2} \int_0^{t \wedge \tau^{k,m}} \gamma_s |X_s^k - X_s^m|^2 ds \\
 & - \alpha \int_0^{t \wedge \tau^{k,m}} \gamma_s |X_s^k - X_s^m|^2 (\mathbb{E}^0 [\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] + \mathbb{E}^0 [\sigma_1^*(s, X_s^m) D_\mu H(\mu_s^m)(X_s^m)]) \cdot dW_s \\
 & - \alpha\beta^2 \int_0^{t \wedge \tau^{k,m}} \gamma_s |X_s^k - X_s^m|^2 (dK_s^k + dK_s^m). \quad (72)
 \end{aligned}$$

Let us now estimate the first two terms in the right-hand side: by (38), we have, for  $0 \leq s \leq t \wedge \tau^{k,m}$  and  $\mathbb{P}^1$ -a.s.,

$$\begin{aligned} & \mathbb{E}^0 \left[ (X_s^k - X_s^m) \cdot D_\mu H(\mu_s^k)(X_s^k) \psi_k(H(\mu_s^k)) \right] \\ & \leq (H(\mu_s^k) - H(\mu_s^m) + C\mathbb{E}^0[|X_s^k - X_s^m|^2]) \psi_k(H(\mu_s^k)) \\ & \leq \left( -H(\mu_s^m)(\mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) > -2/m\}} + \mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) \leq -2/m\}}) \right. \\ & \quad \left. + C\mathbb{E}^0[|X_s^k - X_s^m|^2] \right) \psi_k(H(\mu_s^k)) \\ & \leq \left( \frac{2}{m} + C\mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) \leq -2/m\}} + C\mathbb{E}^0[|X_s^k - X_s^m|^2] \right) \psi_k(H(\mu_s^k)), \end{aligned}$$

where we used that  $z\psi_k(z) \leq 0$  and that  $H$  is bounded. The symmetric inequality holds when exchanging the roles of  $k$  and  $m$ . Therefore, taking the conditional expectation with respect to  $\mathbb{E}^0$  in (72), we obtain:

$$\begin{aligned} & \gamma_{t \wedge \tau^{k,m}} \mathbb{E}^0 \left[ |X_{t \wedge \tau^{k,m}}^k - X_{t \wedge \tau^{k,m}}^m|^2 \right] \\ & \leq C \int_0^{t \wedge \tau^{k,m}} \gamma_s \left( (1/m) + \mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) \leq -2/m\}} + \mathbb{E}^0[|X_s^k - X_s^m|^2] \right) dK_s^k \\ & + C \int_0^{t \wedge \tau^{k,m}} \gamma_s \left( (1/k) + \mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^k) \leq -2/k\}} + \mathbb{E}^0[|X_s^k - X_s^m|^2] \right) dK_s^m \\ & + \int_0^{t \wedge \tau^{k,m}} 2\gamma_s \mathbb{E}^0 \left[ (X_s^k - X_s^m) \cdot (\sigma_1(s, X_s^k) - \sigma_1(s, X_s^m)) \right] dW_s - \frac{\alpha\delta}{2} \int_0^{t \wedge \tau^{k,m}} \gamma_s \mathbb{E}^0[|X_s^k - X_s^m|^2] ds \\ & - \alpha \int_0^{t \wedge \tau^{k,m}} \gamma_s \mathbb{E}^0[|X_s^k - X_s^m|^2] \\ & \quad \times (\mathbb{E}^0[\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] + \mathbb{E}^0[\sigma_1^*(s, X_s^m) D_\mu H(\mu_s^m)(X_s^m)]) \cdot dW_s \\ & - \alpha\beta^2 \int_0^{t \wedge \tau^{k,m}} \gamma_s \mathbb{E}^0[|X_s^k - X_s^m|^2] \{dK_s^k + dK_s^m\}. \end{aligned}$$

So, for  $\alpha$  large enough, we find (recalling that  $\gamma_t$  is positive and bounded),

$$\begin{aligned} & \gamma_{t \wedge \tau^{k,m}} \mathbb{E}^0 \left[ |X_{t \wedge \tau^{k,m}}^k - X_{t \wedge \tau^{k,m}}^m|^2 \right] \\ & \leq C \left( (1/k) + (1/m) + \mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) \leq -2/m\}} + \mathbf{1}_{\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^k) \leq -2/k\}} \right) (K_{t \wedge \tau^{k,m}}^k + K_{t \wedge \tau^{k,m}}^m) \\ & + \int_0^{t \wedge \tau^{k,m}} 2\gamma_s \mathbb{E}^0 \left[ (X_s^k - X_s^m) \cdot (\sigma_1(s, X_s^k) - \sigma_1(s, X_s^m)) \right] dW_s - \frac{\alpha\delta}{2} \int_0^{t \wedge \tau^{k,m}} \gamma_s \mathbb{E}^0[|X_s^k - X_s^m|^2] ds \\ & - \alpha \int_0^{t \wedge \tau^{k,m}} \gamma_s \mathbb{E}^0[|X_s^k - X_s^m|^2] \left( \mathbb{E}^0[\sigma_1^*(s, X_s^k) D_\mu H(\mu_s^k)(X_s^k)] \right. \\ & \quad \left. + \mathbb{E}^0[\sigma_1^*(s, X_s^m) D_\mu H(\mu_s^m)(X_s^m)] \right) \cdot dW_s. \end{aligned} \tag{73}$$

In order to proceed, we need to control the probability of the events  $\{\inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^k) \leq -2/k\}$ . This is the aim of the next step.

**Step 5: Estimate of the probability**  $\mathbb{P}[\{\inf_{u \leq T} H(\mu_u^k) \leq -2/k\}]$ . As  $H([X_0]) \geq 0$  and the map  $u \rightarrow \mu_u^k$  is continuous on  $\mathcal{P}^2(\mathbb{R}^n)$ , on the event  $\{\inf_{u \leq T} H(\mu_u^k) \leq -2/k\}$ , there exist  $0 \leq s \leq t' \leq \tau^k$  such that  $H([X_s^k|W]) = -1/k$ ,  $H(\mu_u^k) \leq -1/k$  on  $[s, t']$  and  $H([X_{t'}^k|W]) \leq -2/k$ . Then (62) (applied between  $s$  and  $t'$ ) implies that

$$-(2/k) \geq -(1/k) + \beta^2 k^2 (t' - s) + \int_s^{t'} \int_{\mathbb{R}^n} \sigma_1^T(u, x) D_\mu H^k(\mu_u^k)(x) \mu_u^k(dx) \cdot dW_u,$$

from which we derive that, for any  $\gamma \in (1/3, 1/2)$ ,

$$\begin{aligned} & \sup_{0 \leq s < t' \leq T} \left| \frac{1}{(t' - s)^\gamma} \int_s^{t'} \int_{\mathbb{R}^n} \sigma_1^T(u, x) D_\mu H^k(\mu_u^k)(x) \mu_u^k(dx) \cdot dW_u \right| \\ & \geq \inf_{0 \leq s < t' \leq T} \left( \frac{1}{k(t' - s)^\gamma} + \beta^2 k^2 (t' - s)^{1-\gamma} \right) \geq C^{-1} k^{3\gamma-1}. \end{aligned}$$

If we set

$$M_t^k := \int_0^t \int_{\mathbb{R}^n} \sigma_1^T(u, x) D_\mu H^k(\mu_u^k)(x) \mu_u^k(dx) \cdot dW_u,$$

then, by Dubin-Schwarz, there is a standard 1-dimensional Brownian motion  $\tilde{W}^k$  such that  $M_t^k = \tilde{W}_{\langle M^k \rangle_t}^k$ . Hence

$$\begin{aligned} \sup_{0 \leq s < t \leq T} \left| \frac{1}{(t-s)^\gamma} (M_t^k - M_s^k) \right| &= \sup_{0 \leq s < t \leq T} \left| \frac{1}{(t-s)^\gamma} (\tilde{W}_{\langle M^k \rangle_t}^k - \tilde{W}_{\langle M^k \rangle_s}^k) \right| \\ &\leq C \sup_{0 \leq s < t \leq CT} \left| \frac{1}{(t-s)^\gamma} (\tilde{W}_t^k - \tilde{W}_s^k) \right|, \end{aligned}$$

because

$$|\langle M^k \rangle_t - \langle M^k \rangle_s| \leq C(t-s),$$

since  $\sigma_1$  is bounded and (52) holds. This proves that

$$\mathbb{P}[\tau^k < T] \leq \mathbb{P}\left[\left\{\inf_{u \leq T} H(\mu_u^k) \leq -2/k\right\}\right] \leq \mathbb{P}\left[\sup_{0 \leq s < t \leq CT} \left| \frac{1}{(t-s)^\gamma} (\tilde{W}_t^k - \tilde{W}_s^k) \right| \geq C^{-1} k^{3\gamma-1}\right] =: \varepsilon_k, \quad (74)$$

where  $\varepsilon_k \rightarrow 0$  since  $\gamma \in (1/3, 1/2)$  by standard properties of the Brownian motion.

**Step 6: Estimate of the difference  $X^k - X^m$ , second part.** We come back to (73): taking expectation we obtain, by (63) and (74),

$$\begin{aligned} & \mathbb{E}[\gamma_{t \wedge \tau^{k,m}} |X_{t \wedge \tau^{k,m}}^k - X_{t \wedge \tau^{k,m}}^m|^2] \\ & \leq C \left( \frac{1}{k} + \frac{1}{m} \right) + C \left( \mathbb{P}^{1/2} \left[ \left\{ \inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^m) \leq -2/m \right\} \right] \right. \\ & \quad \left. + \mathbb{P}^{1/2} \left[ \left\{ \inf_{u \leq t \wedge \tau^{k,m}} H(\mu_u^k) \leq -2/k \right\} \right] \right) \times \mathbb{E}^{1/2} [(K_{t \wedge \tau^{k,m}}^k + K_{t \wedge \tau^{k,m}}^m)^2] \\ & \leq C \left( (1/k) + (1/m) + \varepsilon_k^{1/2} + \varepsilon_m^{1/2} \right). \end{aligned}$$

Therefore

$$\sup_{t \in [0, T]} \mathbb{E} [|X_{t \wedge \tau^{k,m}}^k - X_{t \wedge \tau^{k,m}}^m|^2] \leq C \left( (1/k) + (1/m) + \varepsilon_k^{1/2} + \varepsilon_m^{1/2} \right). \quad (75)$$

Coming back once again to (73), we have by the BDG inequality, (63) and (74),

$$\begin{aligned} & \mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_{s \wedge \tau^{k,m}} \mathbb{E}^0 [|X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2] \right] \\ & \leq C \left( (1/k) + (1/m) + \varepsilon_k^{1/2} + \varepsilon_m^{1/2} \right) + C(\alpha + 1) \mathbb{E} \left[ \left( \int_0^{t \wedge \tau^{k,m}} |\mathbb{E}^0 [|X_s^k - X_s^m|^2]|^2 ds \right)^{1/2} \right]. \end{aligned}$$

Hence by Jensen's inequality

$$\begin{aligned}
& \mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_{s \wedge \tau^{k,m}} \mathbb{E}^0 [ |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 ] \right] \\
& \leq C \left( (1/k) + (1/m) + \varepsilon_m^{1/2} + \varepsilon_k^{1/2} \right) \\
& \quad + C_\alpha \mathbb{E} \left[ \sup_{0 \leq s \leq t} (\gamma_{s \wedge \tau^{k,m}} \mathbb{E}^0 [ |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 ])^{1/2} \left( \int_0^{t \wedge \tau^{k,m}} \mathbb{E}^0 [ |X_s^k - X_s^m|^2 ] ds \right)^{1/2} \right] \\
& \leq C \left( (1/k) + (1/m) + \varepsilon_m^{1/2} + \varepsilon_k^{1/2} \right) + \frac{1}{2} \mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_{s \wedge \tau^{k,m}} \mathbb{E}^0 [ |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 ] \right] \\
& \quad + C \mathbb{E} \left[ \int_0^t |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 ds \right],
\end{aligned}$$

from which we derive, by (75),

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} \mathbb{E}^0 [ |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 ] \right] \leq C \left( (1/k) + (1/m) + \varepsilon_m^{1/2} + \varepsilon_k^{1/2} \right). \quad (76)$$

We finally need to remove the  $\mathbb{E}^0$  in the left-hand side. For this we come back to (72). Taking the supremum in time and the expectation therein we find, by the BDG inequality, since  $\gamma$  is bounded below and above by positive constants,

$$\begin{aligned}
& \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 \right] \\
& \leq C \mathbb{E} \left[ \int_0^{\tau^{k,m}} \mathbb{E}^0 [ |X_s^k - X_s^m|^2 ]^{1/2} \mathbb{E}^0 [ |D_\mu H([X_s^k|W])(X_s^k)|^2 ]^{1/2} dK_s^k \right] \\
& \quad + C \mathbb{E} \left[ \int_0^{\tau^{k,m}} \mathbb{E}^0 [ |X_s^k - X_s^m|^2 ]^{1/2} \mathbb{E}^0 [ |D_\mu H([X_s^m|W])(X_s^m)|^2 ]^{1/2} dK_s^m \right] \\
& \quad + C \mathbb{E} \left[ \left( \int_0^{\tau^{k,m}} |X_s^k - X_s^m|^4 ds \right)^{1/2} \right] \\
& \leq C \mathbb{E} \left[ \sup_{0 \leq s \leq \tau^{k,m}} \mathbb{E}^0 [ |X_s^k - X_s^m|^2 ]^{1/2} (K_{\tau^{k,m}}^k + K_{\tau^{k,m}}^m) \right] \\
& \quad + C \mathbb{E} \left[ \sup_{0 \leq s \leq \tau^{k,m}} |X_s^k - X_s^m| \left( \int_0^{\tau^{k,m}} |X_s^k - X_s^m|^2 ds \right)^{1/2} \right] \\
& \leq C \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq \tau^{k,m}} \mathbb{E}^0 [ |X_s^k - X_s^m|^2 ] \right] \mathbb{E}^{1/2} [ (K_{\tau^{k,m}}^k + K_{\tau^{k,m}}^m)^2 ] \\
& \quad + \frac{1}{2} \mathbb{E} \left[ \sup_{0 \leq s \leq \tau^{k,m}} |X_s^k - X_s^m|^2 \right] + C \mathbb{E} \left[ \int_0^{\tau^{k,m}} |X_s^k - X_s^m|^2 ds \right],
\end{aligned}$$

from which we obtain, in view of our bounds on  $K^k$  and  $K^m$  in (63) and by (75) and (76),

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 \right] \leq C \left( (1/k^{1/2}) + (1/m^{1/2}) + \varepsilon_m^{1/4} + \varepsilon_k^{1/4} \right).$$

We finally need to replace the stopping time  $\tau^{k,m}$  by the stopping times  $\tau^k$  and  $\tau^m$  in the above inequality. We have

$$\begin{aligned} \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^k}^k - X_{s \wedge \tau^m}^m|^2 \right] &\leq \mathbb{E} \left[ \mathbb{E}^0 \left[ \sup_{0 \leq s \leq T} \gamma_s |X_{s \wedge \tau^{k,m}}^k - X_{s \wedge \tau^{k,m}}^m|^2 \right] \mathbf{1}_{\{\tau^{k,m}=T\}} \right] \\ &\quad + 2\mathbb{E} \left[ \left( \mathbb{E}^0 \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^k}^k|^2 \right] + \mathbb{E}^0 \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^m}^m|^2 \right] \right) (\mathbf{1}_{\{\tau^k < T\}} + \mathbf{1}_{\{\tau^m < T\}}) \right] \\ &\leq C \left( (1/k^{1/2}) + (1/m^{1/2}) + \varepsilon_m^{1/4} + \varepsilon_k^{1/4} \right) \\ &\quad + C\mathbb{E}^{1/2} \left[ (\mathbb{E}^0)^2 \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^k}^k|^2 \right] + (\mathbb{E}^0)^2 \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^m}^m|^2 \right] \right] \left( \mathbb{P}^{1/2} [\{\tau^k < T\}] + \mathbb{P}^{1/2} [\{\tau^m < T\}] \right). \end{aligned}$$

By the estimate (66) and the estimate of  $\mathbb{P} [\{\tau^k < T\}]$  in (74) we conclude that

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^k}^k - X_{s \wedge \tau^m}^m|^2 \right] \leq C \left( (1/k^{1/2}) + (1/m^{1/2}) + \varepsilon_m^{1/4} + \varepsilon_k^{1/4} \right).$$

**Step 7: Analysis of the limit process.** We infer from the above inequality that  $(X_{s \wedge \tau^k}^k)$  is a Cauchy sequence in  $\mathcal{S}^2$ . We denote by  $X$  the limit of this sequence and set  $\mu_s = [X_s | W]$ . We note that

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} W_2^2(\mu_{s \wedge \tau^k}^k, \mu_s) \right] \leq \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge \tau^k}^k - X_s|^2 \right] \rightarrow 0.$$

Next we remove the stopping time in the above inequality. By (74),

$$\mathbb{P} [\tau^k < T] \leq \mathbb{P} \left[ \left\{ \inf_{u \leq T} H(\mu_u^k) \leq -2/k \right\} \right] \leq \varepsilon_k,$$

where  $(\varepsilon_k)$  tends to zero as  $k \rightarrow +\infty$ . We can choose a subsequence  $(k')$ , such that  $\sup_{0 \leq s \leq T} W_2^2(\mu_{s \wedge \tau^{k'}}^{k'}, \mu_s)$  tends to 0 a.s. and, by the Borel-Cantelli Lemma, such that, a.s. and for  $k'$  large enough,  $H([X_t^{k'} | W]) \geq -2/k'$  on  $[0, T]$ . In particular, this implies that, a.s. and for  $k'$  large enough,  $\tau^{k'} = T$  and

$$\sup_{0 \leq s \leq T} W_2^2(\mu_s^{k'}, \mu_s) \rightarrow 0.$$

Thus  $H(\mu_t) \geq 0$  a.s. on  $[0, T]$ . Moreover, by (61) and the convergence of  $X^k$ , there exists a continuous process  $(L_t)$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \int_{\mathbb{R}^n} |D_\mu H([X_s^{k'} | W])(x)|^2 \psi_k(H([X_s^{k'} | W])) d\mu_s^{k'}(dx) ds - L_t \right|^2 \right] \rightarrow 0.$$

Note that  $L$  is nondecreasing and adapted to  $\mathcal{F}^W$ . Let us set

$$h(\mu) := \left[ \beta^2 \vee \int_{\mathbb{R}^n} |D_\mu H(\mu)(x)|^2 \mu(dx) \right]^{-1}, \quad \forall \mu \in \mathcal{P}^2(\mathbb{R}^n)$$

and

$$L_t^k = \int_0^t \int_{\mathbb{R}^n} |D_\mu H(\mu_s^k)(x)|^2 d\mu_s^k(dx) dK_s^k.$$

Then  $h$  is continuous and bounded on  $\mathcal{P}^2(\mathbb{R}^n)$  and, as (53) holds and  $-(1/k') \leq H(\mu_s^{k'}) \leq 0$   $dK_s^{k'}$ -a.s. for  $k'$  large enough,

$$\begin{aligned} \lim_{k'} \int_0^t D_\mu H([X_s^{k'} | W])(X_s^{k'}) dK_s^{k'} &= \lim_{k'} \int_0^t D_\mu H(\mu_s^{k'})(X_s^{k'}) h(\mu_s^{k'}) dL_s^{k'} \\ &= \int_0^t D_\mu H([X_s | W])(X_s) h([X_s | W]) dL_s. \end{aligned}$$

Let us set

$$K_t = \int_0^t h([X_s | W]) dL_s.$$



Then  $K$  is nondecreasing and  $\mathcal{F}^W$  adapted and we have, passing to the limit in (59),

$$\begin{aligned} X_t &= X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma_0(s, X_s) dB_s + \int_0^t \sigma_1(s, X_s) dW_s \\ &\quad + \int_0^t D_\mu H([X_s|W])(X_s) dK_s, \quad t \geq 0. \end{aligned}$$

Moreover, for any continuous and bounded map  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\phi(z) = 0$  on  $[0, +\infty)$ , one has

$$\begin{aligned} 0 &= \lim_{k'} \int_0^T \phi(H([X_t^{k'}|W])) \psi_{k'}([X_t^{k'}|W]) dt = \lim_{k'} \int_0^T \phi(H([X_t^{k'}|W])) h([X_t^{k'}|W]) dL_t^{k'} \\ &= \int_0^T \phi(H([X_t|W])) h([X_t|W]) dL_t = \int_0^T \phi(H([X_t|W])) dK_t. \end{aligned}$$

Therefore  $\int_0^T \mathbf{1}_{\{H([X_t|W]) > 0\}} dK_t = 0$ . Note finally that (55) can be derived from (63) and (64).  $\square$

*Proof of Proposition 15.* The proof goes along the same lines as the estimate of the difference  $X^k - X^m$  in the proof of Theorem 11 (Step 4). Let  $(X, K^X)$  and  $(Y, K^Y)$  be two solutions, let us set  $\mu_t^X = [X_t|W]$ ,  $\mu_t^Y = [Y_t|W]$  and  $\gamma_t := \exp\{-\alpha(\delta t + H(\mu_t^X) + H(\mu_t^Y))\}$  for  $\alpha, \delta \geq 1$  to be chosen below. By Itô's formula (16) (which holds because (15) is satisfied thanks to assumption (52)), we have, using the Lipschitz continuity of  $b$  and  $\sigma_i$  and for  $\delta > 0$  large enough:

$$\begin{aligned} \gamma_t |X_t - Y_t|^2 &\leq |X_0 - Y_0|^2 \\ &\quad + \int_0^t 2\gamma_s (X_s - Y_s) \cdot (D_\mu H(\mu_s^X)(X_s) dK_s^X - D_\mu H(\mu_s^Y)(Y_s) dK_s^Y) \\ &\quad + \int_0^t 2\gamma_s (X_s - Y_s) \cdot (\sigma_0(s, X_s) - \sigma_0(s, Y_s)) dB_s \\ &\quad + \int_0^t 2\gamma_s (X_s - Y_s) \cdot (\sigma_1(s, X_s) - \sigma_1(s, Y_s)) dW_s - \frac{\alpha\delta}{2} \int_0^t \gamma_s |X_s - Y_s|^2 ds \\ &\quad - \alpha \int_0^t \gamma_s |X_s - Y_s|^2 (\mathbb{E}^0 [(\sigma_1(s, X_s))^* D_\mu H(\mu_s^X)(X_s)] + \mathbb{E}^0 [(\sigma_1(s, Y_s))^* D_\mu H(\mu_s^Y)(Y_s)]) \cdot dW_s \\ &\quad - \alpha \int_0^t \gamma_s |X_s - Y_s|^2 \{ \mathbb{E}^0 [|D_\mu H(\mu_s^X)(X_s)|^2] dK_s^X + \mathbb{E}^0 [|D_\mu H(\mu_s^Y)(Y_s)|^2] dK_s^Y \}. \end{aligned} \quad (77)$$

As  $H$  satisfies (38), we have, for  $dK^X$ -a.e.  $s$ ,

$$\mathbb{E}^0 [(X_s - Y_s) \cdot D_\mu H(\mu_s^X)(X_s)] \leq H(\mu_s^X) - H(\mu_s^Y) + C\mathbb{E}^0 [|X_s - Y_s|^2] \leq C\mathbb{E}^0 [|X_s - Y_s|^2].$$

The symmetric inequality holds when exchanging the roles of  $X$  and  $Y$ . Therefore, taking the conditional expectation with respect to  $\mathbb{E}^0$  in (77), we obtain, for  $\alpha$  and  $\delta$  large enough and by (53) (recalling that  $\gamma_t$  is positive and bounded):

$$\begin{aligned} \gamma_t \mathbb{E}^0 [|X_t - Y_t|^2] &\leq \mathbb{E}^0 [|X_0 - Y_0|^2] \\ &\quad + \int_0^t 2\gamma_s \mathbb{E}^0 [(X_s - Y_s) \cdot (\sigma_1(s, X_s) - \sigma_1(s, Y_s))] dW_s - \frac{\alpha\delta}{4} \int_0^t \gamma_s \mathbb{E}^0 [|X_s - Y_s|^2] ds \\ &\quad - \alpha \int_0^t \gamma_s \mathbb{E}^0 [|X_s - Y_s|^2] \left( \mathbb{E}^0 [(\sigma_1(s, X_s))^* D_\mu H(\mu_s^X)(X_s)] + \mathbb{E}^0 [(\sigma_1(s, Y_s))^* D_\mu H(\mu_s^Y)(Y_s)] \right) \cdot dW_s. \end{aligned} \quad (78)$$

Taking expectation, we obtain:

$$\sup_{0 \leq t \leq T} \mathbb{E} [|X_t - Y_t|^2] \leq \mathbb{E} [|X_0 - Y_0|^2]. \quad (79)$$

Coming back to (78), we have by the BDG inequality

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_s \mathbb{E}^0 [|X_s - Y_s|^2] \right] \leq \mathbb{E} [|X_0 - Y_0|^2] + C(\alpha + 1) \mathbb{E} \left[ \left( \int_0^t |\mathbb{E}^0 [|X_s - Y_s|^2]|^2 ds \right)^{1/2} \right],$$

from which we easily conclude using (79) that

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} \mathbb{E}^0 [ |X_s - Y_s|^2 ] \right] \leq C_\alpha \mathbb{E} [ |X_0 - Y_0|^2 ]. \quad (80)$$

We finally remove the  $\mathbb{E}^0$  in the left-hand side. For this we come back to (77), take the sup in time and the expectation to find, again for  $\delta$  large enough,

$$\begin{aligned} & \mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_s |X_s - Y_s|^2 \right] \leq \mathbb{E} [ |X_0 - Y_0|^2 ] \\ & + C \mathbb{E} \left[ \int_0^t (\mathbb{E}^0 [ |X_s - Y_s|^2 ])^{1/2} (\mathbb{E}^0 [ |D_\mu H(\mu_s^X)(X_s)|^2 ])^{1/2} dK_s^X \right] \\ & + C \mathbb{E} \left[ \int_0^t (\mathbb{E}^0 [ |X_s - Y_s|^2 ])^{1/2} (\mathbb{E}^0 [ |D_\mu H(\mu_s^Y)(Y_s)|^2 ])^{1/2} dK_s^Y \right] \\ & + C(\alpha + 1) \mathbb{E} \left[ \left( \int_0^t \gamma_s^2 |X_s - Y_s|^4 ds \right)^{1/2} \right] \\ & \leq \mathbb{E} [ |X_0 - Y_0|^2 ] + C \mathbb{E} \left[ \sup_{0 \leq s \leq t} (\mathbb{E}^0 [ |X_s - Y_s|^2 ])^{1/2} (K_t^X + K_t^Y) \right] \\ & + C(\alpha + 1) \mathbb{E} \left[ \sup_{0 \leq s \leq t} (\gamma_s^{1/2} |X_s - Y_s|) \left( \int_0^t \gamma_s |X_s - Y_s|^2 ds \right)^{1/2} \right] \\ & \leq \mathbb{E} [ |X_0 - Y_0|^2 ] + C \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq t} \mathbb{E}^0 [ |X_s - Y_s|^2 ] \right] \mathbb{E}^{1/2} [(K_t^X + K_t^Y)^2] \\ & + \frac{1}{2} \mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_s |X_s - Y_s|^2 \right] + C(\alpha + 1) \mathbb{E} \left[ \int_0^t |X_s - Y_s|^2 ds \right]. \end{aligned}$$

In conclusion, by our bounds on  $K^X$  and  $K^Y$ , (79) and (80), we have

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} |X_s - Y_s|^2 \right] \leq C \left( \mathbb{E}^{1/2} [ |X_0 - Y_0|^2 ] + \mathbb{E} [ |X_0 - Y_0|^2 ] \right).$$

□

We are now ready to show the existence of a solution without the extra moment assumption on the initial condition:

*Proof of Theorem 11: existence in the general case.* Let  $X_0^k := X_0 \mathbf{1}_{\{|X_0| \leq k\}}$ . Then  $(X_0^k)$  has a moment of order 4 and converges in  $L_2(\Omega)$  to  $X_0$ . Let  $(X^k, K^k)$  be the solution to (51) associated with the initial condition  $X_0^k$ . Proposition 15 states that  $(X^k)$  is a Cauchy sequence in  $\mathcal{S}^2$ . Then we can complete the proof exactly as before. □

*Proof of Theorem 11: uniqueness.* Uniqueness of the path  $X$  follows from Proposition 15 so that uniqueness of the process  $K$  can be deduced by reproducing the proof done in Theorem 8.

We finally explain why the weak uniqueness also holds for (51). Let  $(X, K, B, W)$  and  $(X', K', B', W')$  be two weak solutions of (51). Let us now build a pair of processes  $(\tilde{X}, \tilde{K})$  from the data  $(B', W')$  as we did for  $(X, K)$ . It is hence clear that  $(X, K, B, W)$  and  $(\tilde{X}, \tilde{K}, B', W')$  have the same law. But from pathwise uniqueness we have  $(X', K') = (\tilde{X}, \tilde{K})$  so that weak uniqueness holds. □

*Proof of Proposition 16.* On the one hand we apply Itô's formula to  $H(\mu_t)$  and then proceed as in (61) to obtain similarly to (62) that

$$\beta^{-2} \left\{ |H(\mu_t) - H(\mu_s)| + C(t-s) - \int_s^t \mathbb{E}^0 [ \sigma_1^*(u, X_u) D_\mu H(\mu_u)(X_u) ] \cdot dW_u \right\} \geq K_t - K_s,$$

since  $u \mapsto K_u$  is non-decreasing. Also, from boundedness of  $D_\mu H$  we have thanks to Cauchy-Schwartz's inequality that  $\mathbb{P}^1$ -a.s.

$$|H(\mu_t) - H(\mu_s)| \leq M(\mathbb{E}^0)^{1/2} [ |X_t - X_s|^2 ].$$

On the other hand, applying Itô's formula on  $\gamma(t)|X_t - X_s|^2$  where  $\gamma(t) = \exp\{-\alpha H(\mu_t)\}$  for some  $\alpha > 0$ , then taking expectation  $\mathbb{E}^0$  and using the boundedness of the coefficients and estimate (55) we obtain that

$$\begin{aligned} \mathbb{E}^0 [\gamma(t)|X_t - X_s|^2] &\leq C(t-s) + 2 \int_s^t \gamma(u) \mathbb{E}^0 [(X_u - X_s) \cdot D_\mu H(\mu_u)(X_u)] dK_u \\ &\quad - \alpha \int_s^t \gamma(u) \mathbb{E}^0 [|X_u - X_s|^2] \mathbb{E}^0 [|D_\mu H(\mu_u)(X_u)|^2] dK_u \\ &\quad + 2 \int_s^t \gamma_u \mathbb{E}^0 [\sigma_1^*(u, X_u)(X_u - X_s)] \cdot dW_u \\ &\quad - \alpha \int_s^t \mathbb{E}^0 [\gamma_u |X_u - X_s|^2] \mathbb{E}^0 [\sigma_1^*(u, X_u) D_\mu H(\mu_u)(X_u)] \cdot dW_u, \end{aligned}$$

provided that  $\alpha$  is large enough. Since  $H$  is semi-convex we have

$$\begin{aligned} &\int_s^t \gamma(u) \mathbb{E}^0 [(X_u - X_s) \cdot D_\mu H(\mu_u)(X_u)] dK_u \\ &\leq \int_s^t \gamma(u) \left( H(\mu_u) - H(\mu_s) + \mathbb{E}^0 [|X_t - X_s|^2] \right) dK_u \leq \int_s^t \gamma(u) \mathbb{E}^0 [|X_t - X_s|^2] dK_u, \end{aligned}$$

according to the Skorokhod condition. Hence we have, for  $\alpha$  large enough,

$$\begin{aligned} \mathbb{E}^0 [\gamma(t)|X_t - X_s|^2] &\leq C(t-s) + 2 \int_s^t \gamma_u \mathbb{E}^0 [\sigma_1^*(u, X_u)(X_u - X_s)] \cdot dW_u \\ &\quad - \alpha \int_s^t \mathbb{E}^0 [\gamma_u |X_u - X_s|^2] \mathbb{E}^0 [\sigma_1^*(u, X_u) D_\mu H(\mu_u)(X_u)] \cdot dW_u. \end{aligned} \quad (81)$$

Recalling that  $\gamma$  is bounded from below away from 0, then using the above estimate we have

$$\begin{aligned} K_t - K_s &\leq C \left\{ (t-s) - \int_s^t \mathbb{E}^0 [\sigma_1^*(u, X_u^k) D_\mu H(\mu_u^k)(X_u^k)] \cdot dW_u \right. \\ &\quad + \left( C(t-s) - \alpha \int_s^t \mathbb{E}^0 [\gamma_u |X_u - X_s|^2] \mathbb{E}^0 [\sigma_1^*(u, X_u) D_\mu H(\mu_u)(X_u)] \cdot dW_u \right. \\ &\quad \left. \left. + \int_s^t \gamma_u \mathbb{E}^0 [\sigma_1^*(u, X_u)(X_u - X_s)] \cdot dW_u \right)^{1/2} \right\}. \end{aligned}$$

Taking the expectation  $\mathbb{E}$ , using boundedness of the coefficients, (66) and BDG's inequality we deduce (57).

To obtain estimate on the first term in the left hand side of (58) one just has to use standard computations and the following estimate:

$$\begin{aligned} \mathbb{E}^0 \left[ \left( \int_s^t D_\mu H(\mu_u)(X_u) dK_u \right)^p \right] &= \int_t^s \dots \int_t^s \mathbb{E}^0 [D_\mu H(\mu_{u_1})(X_{u_1}) \dots D_\mu H(\mu_{u_p})(X_{u_p})] dK_{u_1} \dots dK_{u_p} \\ &\leq \int_t^s \dots \int_t^s \mathbb{E}^0 \left[ p^{-1} \sum_{\ell=1}^p |D_\mu H(\mu_{u_\ell})(X_{u_\ell})|^p \right] dK_{u_1} \dots dK_{u_p} \\ &\leq \mathcal{H}_p |K_t - K_s|^p. \end{aligned}$$

To conclude, let us precise that the second term in the left hand side of (58) is controlled starting from the difference of the processes  $X_t - X_s$ . Then, standards computations give

$$\begin{aligned} &\mathbb{E} \left[ \sup_{s \leq r \leq t} |\mathbb{E}^0 [|X_r - X_s|^{2p/2}] \right] \\ &\leq C_{T,p} \left\{ 1 + \mathbb{E} \left[ \left( \sup_{s \leq r \leq t} \mathbb{E}^0 \left[ \left| \int_s^r D_\mu H(\mu_u)(X_u) dK_u \right|^2 \right] \right)^{p/2} \right] + \mathbb{E} \left[ \sup_{s \leq r \leq t} \left( \mathbb{E}^0 \left[ \left| \int_s^r \sigma^1(u, X_u) dW_u \right|^2 \right] \right)^{p/2} \right] \right\} \end{aligned}$$

The second term in the above r.h.s is dealt using similar arguments as above while the third one can be estimated by using the boundedness of  $\sigma^1$  and writing, provided  $p/2$  is an integer,

$$\begin{aligned} \left( \mathbb{E}^0 \left[ \left| \int_s^r \sigma^1(u, X_u) dW_u \right|^2 \right] \right)^{p/2} &= \left( \left| \int_s^r \int_s^u \mathbb{E}^0 [\sigma^1(u', X_{u'}) (\sigma^1)^*(u, X_u)] dW_{u'} dW_u \right. \right. \\ &\quad \left. \left. + \int_s^r \int_s^{u'} \mathbb{E}^0 [\sigma^1(u', X_{u'}) (\sigma^1)^*(u, X_u)] dW_u dW_{u'} \right| \right)^{p/2}. \end{aligned}$$

□

**3.2. The mean field limit.** We consider the particle system (for  $i = 1, \dots, N$ )

$$\begin{cases} X_t^i = X_0^i + \int_0^t b(s, X_s^i) ds + \int_0^t \sigma_0(s, X_s^i) dB_s^i + \int_0^t \sigma_1(s, X_s^i) dW_s + \int_0^t D_\mu H(\mu_s^N)(X_s^i) dK_s^N, \\ \mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}, \quad H(\mu_t^N) \geq 0, \quad \int_0^t H(\mu_s^N) dK_s^N = 0, \quad t \geq 0. \end{cases} \quad (82)$$

Here the data  $b$ ,  $\sigma$  and  $H$  satisfy conditions **(Hcl)** at the beginning of the section and, in addition,  $b$ ,  $\sigma_0$  and  $\sigma_1$  are deterministic. The  $B^i$  and  $W$  are independent Brownian motions. The initial conditions of the particles  $X_0^i$  are i.i.d. random variable with law  $\mu_0$  and are independent of the  $B^i$  and of  $W$ .  $K^N$  is a continuous, nondecreasing process adapted to the filtration  $\mathcal{F}^N$  generated by the  $X_0^i$ , the  $B^i$  and  $W$ . We still assume that the probability space is  $(\Omega, \mathbb{P}) = (\Omega^0 \times \Omega^1, \mathbb{P}^0 \otimes \mathbb{P}^1)$ , where  $\Omega^0$  supports the  $X_0^i$  and the  $B^i$ , while  $\Omega^1$  supports  $W$  with associated filtration  $\mathcal{F}^W = \mathcal{F}^1$ . We work on the set

$$\Omega_N := \left\{ H \left( \frac{1}{N} \sum_{i=1}^N \delta_{X_0^i} \right) \geq 0 \right\}.$$

Let us remark that (82) reads as a classical reflected SDE in  $(\mathbb{R}^n)^N$ , with normal reflexion in the constraint

$$\mathcal{O}_N = \left\{ (x_1, \dots, x_N) \in (\mathbb{R}^n)^N, H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) > 0 \right\}.$$

Indeed, let us recall that, if

$$g(x_1, \dots, x_N) = H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right),$$

then (see [11])

$$\partial_{x_i} g(x_1, \dots, x_N) = \frac{1}{N} D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_i).$$

Therefore the vector  $-(D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_1), \dots, D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_N))$  is an outward normal to the set  $\mathcal{O}_N$  at the point  $(x_1, \dots, x_N) \in \partial \mathcal{O}_N$ . Existence and uniqueness of the solution are therefore immediate under our standing assumptions (cf. [25]).

Our aim is to show the convergence, on  $\Omega_N$ , of the  $X^i$  to the solution  $\bar{X}^i$  to

$$\begin{cases} \bar{X}_t^i = X_0^i + \int_0^t b(s, \bar{X}_s^i) ds + \int_0^t \sigma_0(s, \bar{X}_s^i) dB_s^i + \int_0^t \sigma_1(s, \bar{X}_s^i) dW_s + \int_0^t D_\mu H(\bar{\mu}_s)(\bar{X}_s^i) dK_s, \\ \bar{\mu}_t = [X_t^i | W], \quad H(\bar{\mu}_t) \geq 0, \quad \int_0^t H(\bar{\mu}_s) dK_s = 0, \quad t \geq 0. \end{cases}$$

Note that the processes  $\bar{\mu}$  and  $K$  do not depend on  $i$  and  $N$ : this is obvious by the uniqueness in law, see Remark 12. The  $\bar{X}^i$  are no longer i.i.d., but they are i.i.d. given  $W$ . In particular, if  $\bar{\mu}_s^N$  stands for the empirical measure of  $(\bar{X}_s^i)_{1 \leq i \leq N}$ , then we will show later that

$$\lim_{N \rightarrow +\infty} \mathbb{E} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right] = 0.$$

As the initial conditions  $X_0^i$  of the particles are i.i.d. and distributed according to  $\mu_0$ ,  $\Omega_N$  (the set of events for which the initial condition belongs to  $\mathcal{O}_N$  and for which for which system (82) has a meaning) has not a full probability in general. There are several solutions to overcome this issue. We illustrate two of them in this paper. One can, for instance, modify the initial position and “project” it onto  $\mathcal{O}^N$ . We shall follow this approach in Subsection 4.2 when we analyze the particle system associated with backward SDEs with normal reflexion in law. Here we simply concentrate on the set  $\Omega_N$ . If one assumes, for instance, that  $H(\mu_0)$  is positive, this event occurs with a probability which tends to 1 as  $N \rightarrow +\infty$ .

**Theorem 17.** *For  $T \geq 0$ , there exists a constant  $C(T)$  independent of  $N$  such that*

$$\sup_{i=1,\dots,N} \sup_{t \in [0,T]} \mathbb{E} [ |X_t^i - \bar{X}_t^i|^2 \mathbf{1}_{\Omega_N} ] \leq C(T) \mathbb{E}^{1/2} \left[ \sup_{0 \leq t \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right]$$

and, for  $N$  large enough,

$$\sup_{i=1,\dots,N} \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t^i - \bar{X}_t^i|^2 \mathbf{1}_{\Omega_N} \right] \leq C(T) \mathbb{E}^{1/4} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right].$$

**Lemma 18.** *Assume that the initial data satisfies the following moment condition: there exists an integer  $p \geq 8$  such that  $\mathcal{M}_p([X_0]) + \mathcal{H}_p$  is finite. Then, there exists  $C_T := C(T, (\mathbf{Hcl}), \mathcal{M}_p([\xi]), \mathcal{H}_p) > 0$  such that*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right] \leq C_T \epsilon_N,$$

where  $\epsilon_N$  is given in Lemma 1.

*Proof.* Since  $p \geq 8$  we have, thanks to Proposition 16 and Corollary 14, that the assumptions of Lemma 1 are satisfied.  $\square$

Before proving Theorem 17, let us point out some consequences of (38) and (39). Let  $(x_i)_{1 \leq i \leq N}$  and  $(y_i)_{1 \leq i \leq N}$  be two sequences of points of  $\mathbb{R}^n$ . We set

$$\mu_y^N = \frac{1}{N} \sum_{i=1}^N \delta_{y_i}, \quad \mu_x^N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}.$$

Inequality (38) becomes in this case

$$H(\mu_y^N) - H(\mu_x^N) \leq \mathbb{E} [D_\mu H(\mu_x^N)(X) \cdot (Y - X)] + C \mathbb{E} [|X - Y|^2],$$

where  $[X] = \mu_x^N$  and  $[Y] = \mu_y^N$ . Of course, this inequality is true for any pair  $(X, Y)$  with  $[X] = \mu_x^N$  and  $[Y] = \mu_y^N$  so that we can choose a specific law for  $(X, Y)$ . If all  $(x_i)$  are distinct, let us pick  $Y = f(X)$  with  $f(x_i) = y_i$ . We get

$$\mathbb{E} [D_\mu H(\mu_x^N)(X) \cdot (f(X) - X)] + C \mathbb{E} [|X - f(X)|^2] = \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_x^N)(x_i) \cdot (y_i - x_i) + \frac{C}{N} \sum_{i=1}^N |y_i - x_i|^2.$$

Therefore,

$$\frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_x^N)(x_i) \cdot (x_i - y_i) \leq H(\mu_x^N) - H(\mu_y^N) + \frac{C}{N} \sum_{i=1}^N |y_i - x_i|^2. \quad (83)$$

This inequality also holds when the  $(x_i)$  are not necessarily distinct by approximation. Moreover, inequality (39) becomes, with  $[X] = \mu_x^N$  and for any  $\nu \in \mathcal{P}_2(\mathbb{R}^n)$ ,

$$\mathbb{E} [ |D_\mu H(\mu_x^N)(X) - D_\mu H(\nu)(X)|^2 ] = \frac{1}{N} \sum_{i=1}^N |D_\mu H(\mu_x^N)(x_i) - D_\mu H(\nu)(x_i)|^2 \leq C W_2^2(\mu_x^N, \nu). \quad (84)$$

We will also need below a few points on the L-derivative. For  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{R}^n)^N$ , we set  $H^N(\mathbf{x}) = H(\frac{1}{N} \sum_{i=1}^N \delta_{x_i})$ . We know from [11] that

$$\begin{aligned} D_{x_i} H^N(\mathbf{x}) &= \frac{1}{N} D_\mu H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right)(x_i), \\ D_{x_i x_i}^2 H^N(\mathbf{x}) &= \frac{1}{N} \partial_y D_\mu H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right)(x_i) + \frac{1}{N^2} D_{\mu\mu}^2 H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right)(x_i, x_i), \\ D_{x_i x_j}^2 H^N(\mathbf{x}) &= \frac{1}{N^2} D_{\mu\mu}^2 H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right)(x_i, x_j) \quad i \neq j. \end{aligned}$$

So, if  $(X^i)$  and  $\mu^N$  satisfies the reflected SDE (82), we have, by Itô's formula,

$$\begin{aligned} dH(\mu_s^N) &= \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_s^N)(X_s^i) \cdot b(s, X_s^i) ds \\ &\quad + \frac{1}{2N} \sum_{i=1}^N \text{Tr}((\sigma_0 \sigma_0^*(s, X_s^i) + \sigma_1 \sigma_1^*(s, X_s^i)) \partial_y D_\mu H(\mu_s^N)(X_s^i)) ds \\ &\quad + \frac{1}{2N^2} \sum_{i=1}^N \text{Tr}(\sigma_0 \sigma_0^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^i, X_s^i)) ds \\ &\quad + \frac{1}{2N^2} \sum_{i,j=1}^N \text{Tr}(\sigma_1(s, X_s^j) \sigma_1^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^j, X_s^i)) ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_0(s, X_s^i) dB_s^i \\ &\quad + \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_1(s, X_s^i) dW_s + \frac{1}{N} \sum_{i=1}^N |D_\mu H(\mu_s^N)(X_s^i)|^2 dK_s^N. \end{aligned} \tag{85}$$

We also note that, by Assumption (52),

$$\frac{1}{N} \sum_{i=1}^N |D_\mu H(\mu_s^N)(X_s^i)| = \int_{\mathbb{R}^n} |D_\mu H(\mu_s^N)(x)| \mu_s^N(dx) \leq M, \tag{86}$$

and

$$\frac{1}{N^2} \sum_{i,j=1}^N |D_{\mu\mu}^2 H(\mu_s^N)(X_s^j, X_s^i)| = \int_{\mathbb{R}^{2n}} |D_{\mu\mu}^2 H(\mu_s^N)(x, y)| \mu_s^N(dx) \mu_s^N(dy) \leq M. \tag{87}$$

*Proof of Theorem 17.* We work on  $\Omega_N$ . Let us set  $\Delta X_t^i = X_t^i - \bar{X}_t^i$ . For  $\alpha, \delta \geq 1$  to be chosen below, we set  $\gamma_t = \exp\{-\alpha(\delta t + H(\mu_s^N) + H(\bar{\mu}_s))\}$ . We note that  $\gamma_t$  is bounded above and below by positive constant on bounded time intervals. Choosing  $\alpha$  and  $\delta$  large enough, we have, by Itô's formula (16) and (85) and by the Lipschitz continuity of  $b$ ,  $\sigma_0$  and  $\sigma_1$  as well as the bounds on the coefficients (see also

(86) and (87):

$$\begin{aligned}
\gamma_t |\Delta X_t^i|^2 &\leq 2 \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\mu_s^N)(X_s^i) dK_s^N - 2 \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\bar{\mu}_s)(\bar{X}_s^i) dK_s \\
&\quad - \frac{\alpha \delta}{2} \int_0^t \gamma_s |\Delta X_s^i|^2 ds + 2 \int_0^t \gamma_s \Delta X_s^i \cdot (\sigma_0(s, X_s^i) - \sigma_0(s, \bar{X}_s^i)) dB_s^i \\
&\quad + 2 \int_0^t \gamma_s \Delta X_s^i \cdot (\sigma_1(s, X_s^i) - \sigma_1(s, \bar{X}_s^i)) dW_s - \frac{\alpha}{N} \sum_{j=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 D_\mu H(\mu_s^N)(X_s^j) \cdot \sigma_0(s, X_s^j) dB_s^j \\
&\quad - \alpha \int_0^t \gamma_s |\Delta X_s^i|^2 \left( \frac{1}{N} \sum_{j=1}^N \sigma_1^*(s, X_s^j) D_\mu H(\mu_s^N)(X_s^j) + \mathbb{E}^0 [\sigma_1^*(s, \bar{X}_s^i) D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)] \right) \cdot dW_s \\
&\quad - \alpha \int_0^t \gamma_s |\Delta X_s^i|^2 \left( \frac{1}{N} \sum_{j=1}^N |D_\mu H(\mu_s^N)(X_s^j)|^2 \right) dK_s^N - \alpha \int_0^t \gamma_s |\Delta X_s^i|^2 \mathbb{E}^0 [ |D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)|^2 ] dK_s.
\end{aligned} \tag{88}$$

Using (83), we have, since  $H(\mu_s^N) = 0$   $dK$ -a.e. and  $H[\bar{\mu}_s] = H([\bar{X}_s^i|W]) \geq 0$ ,

$$\begin{aligned}
&\frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\mu_s^N)(X_s^i) dK_s^N \\
&\quad \leq 2 \int_0^t \gamma_s (H(\mu_s^N) - H(\bar{\mu}_s^N)) dK_s^N + \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s^N \\
&\quad \leq 2 \int_0^t \gamma_s (H(\bar{\mu}_s) - H(\bar{\mu}_s^N)) dK_s^N + \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s^N.
\end{aligned}$$

Since  $H$  is  $M$ -Lipschitz for  $W_2$ , we get

$$\begin{aligned}
&\frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\mu_s^N)(X_s^i) dK_s^N \\
&\quad \leq 2M \sup_{0 \leq s \leq t} W_2(\bar{\mu}_s^N, \bar{\mu}_s) K_t^N + \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s^N.
\end{aligned} \tag{89}$$

We split the term  $-\frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\bar{\mu}_s)(\bar{X}_s^i) dK_s$  into two parts:

$$\begin{aligned}
A &:= -\frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot D_\mu H(\bar{\mu}_s^N)(\bar{X}_s^i) dK_s, \\
B &:= -\frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot \{D_\mu H(\bar{\mu}_s^N)(\bar{X}_s^i) - D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)\} dK_s.
\end{aligned}$$

We use (83) for the first one: since  $H(\mu_s^N) \geq 0$  and  $H(\bar{\mu}_s) = 0$   $dK_s$ -a.e.

$$\begin{aligned}
A &= \frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s (\bar{X}_s^i - X_s^i) \cdot D_\mu H(\bar{\mu}_s^N)(\bar{X}_s^i) dK_s \\
&\quad \leq 2 \int_0^t \gamma_s (H(\bar{\mu}_s^N) - H(\bar{\mu}_s)) dK_s + \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s,
\end{aligned}$$

and using the Lipschitz continuity of  $H$ , we get

$$A \leq 2M \sup_{0 \leq s \leq t} W_2(\bar{\mu}_s^N, \bar{\mu}_s) K_t + \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s. \tag{90}$$

Finally, we have, using Cauchy-Schwarz inequality, the global bound on  $\int_{\mathbb{R}^n} |D_\mu H(\mu)| d\mu$  (given by assumption (52)) and (84),

$$\begin{aligned} B &\leq 2 \int_0^t \gamma_s \left( \frac{1}{N} \sum_{i=1}^N |\Delta X_s^i|^2 \right)^{1/2} \left( \frac{1}{N} \sum_{i=1}^N |D_\mu H(\bar{\mu}_s^N)(\bar{X}_s^i) - D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)|^2 \right)^{1/2} dK_s \\ &\leq C \int_0^t \gamma_s \left( \frac{1}{N} \sum_{i=1}^N |\Delta X_s^i|^2 \right)^{1/2} \left( \frac{1}{N} \sum_{i=1}^N |D_\mu H(\bar{\mu}_s^N)(\bar{X}_s^i) - D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)|^2 \right)^{1/4} dK_s \\ &\leq C \int_0^t \gamma_s \left( \frac{1}{N} \sum_{i=1}^N |\Delta X_s^i|^2 \right)^{1/2} W_2^{1/2}(\bar{\mu}_s^N, \bar{\mu}_s) dK_s, \end{aligned}$$

from which we deduce

$$B \leq \frac{C}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 dK_s + C \sup_{0 \leq s \leq t} W_2(\bar{\mu}_s^N, \bar{\mu}_s) K_t. \quad (91)$$

Inserting Assumption (53), estimates (89), (90) and (91) into (88), we get, for  $\alpha$  and  $\delta$  large enough,

$$\begin{aligned} \frac{\gamma_t}{N} \sum_{i=1}^N |\Delta X_t^i|^2 &\leq C (K_t^N + K_t) \sup_{0 \leq s \leq t} W_2(\bar{\mu}_s^N, \bar{\mu}_s) + \frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot (\sigma_0(X_s^i) - \sigma_0(\bar{X}_s^i)) dB_s^i \\ &+ \frac{2}{N} \sum_{i=1}^N \int_0^t \gamma_s \Delta X_s^i \cdot (\sigma_1(s, X_s^i) - \sigma_1(s, \bar{X}_s^i)) dW_s - \frac{\alpha}{N^2} \sum_{i,j=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 D_\mu H(\mu_s^N)(X_s^j) \cdot \sigma_0(s, X_s^j) dB_s^j \\ &- \frac{\alpha}{N} \sum_{i=1}^N \int_0^t \gamma_s |\Delta X_s^i|^2 \left( \frac{1}{N} \sum_{j=1}^N \sigma_1^*(s, X_s^j) D_\mu H(\mu_s^N)(X_s^j) + \mathbb{E}^0 [\sigma_1^*(s, \bar{X}_s^i) D_\mu H(\bar{\mu}_s)(\bar{X}_s^i)] \right) \cdot dW_s. \quad (92) \end{aligned}$$

We prove below the following exponential moment estimate on  $K^N$ :

$$\sup_{N \geq 1} \mathbb{E} [\exp\{\theta K_t^N\} \mathbf{1}_{\Omega_N}] \leq C_\theta(t), \quad \forall \theta > 0. \quad (93)$$

Then, taking expectation in (92) and using the Cauchy-Schwarz inequality, (93) and the fact that  $(K_t)$  has also exponential moments (Theorem 11), we get, as  $\Omega_N$  is independent of  $B^i$  and  $W$ ,

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\Delta X_t^i|^2 \mathbf{1}_{\Omega_N} \right] &\leq C (\mathbb{E}^{1/2} [(K_t^N)^2 \mathbf{1}_{\Omega_N}] + \mathbb{E}^{1/2} [K_t^2]) \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq t} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right] \\ &\leq C(t) \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq t} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right]. \quad (94) \end{aligned}$$

Note that, by the exchangeability of the  $(X^i)$ , this implies that, for any  $T > 0$ ,

$$\sup_{t \in [0, T]} \sup_{i=1, \dots, N} \mathbb{E} [|\Delta X_t^i|^2 \mathbf{1}_{\Omega_N}] \leq C(T) \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right]. \quad (95)$$

In order to improve this inequality and have the sup in time into the expectation, we come back to (92) and obtain, by the BDG inequality,

$$\begin{aligned} &\mathbb{E} \left[ \sup_{0 \leq t \leq T} \frac{1}{N} \sum_{i=1}^N \mathbb{E}^0 [ |X_t^i - \bar{X}_t^i|^2 ] \mathbf{1}_{\Omega_N} \right] \\ &\leq C \mathbb{E} \left[ (K_T^N \mathbf{1}_{\Omega_N} + K_T) \sup_{0 \leq s \leq T} W_2(\bar{\mu}_s^N, \bar{\mu}_s) \right] + C \mathbb{E} \left[ \left( \int_0^T \frac{1}{N} \sum_{i=1}^N (\mathbb{E}^0 [ |\Delta X_s^i|^2 ]) ds \right)^{1/2} \mathbf{1}_{\Omega_N} \right], \end{aligned}$$

from which we infer by the estimate on  $K$  in Theorem 11 and on  $K^N$  in (93), by (94) and by the usual argument, that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \frac{1}{N} \sum_{i=1}^N \mathbb{E}^0 [ |X_t^i - \bar{X}_t^i|^2 ] \mathbf{1}_{\Omega_N} \right] \leq C(T) \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right]. \quad (96)$$



Then we come back to the corresponding version of (88) with  $\gamma_t = \exp(-\delta t)$  therein, take the sup in time and use the BDG inequality, the bound on  $D_\mu H$  (through (86) in particular) to get

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} \gamma_t |\Delta X_s^i|^2 \mathbf{1}_{\Omega_N} \right] \leq C \mathbb{E} \left[ \int_0^t (\mathbb{E}^0 [|\Delta X_s^i|^2])^{1/2} \{dK_s^N + dK_s\} \mathbf{1}_{\Omega_N} \right] + C \mathbb{E} \left[ \left( \int_0^t \gamma_s |\Delta X_s^i|^2 ds \right)^{1/2} \mathbf{1}_{\Omega_N} \right].$$

By exchangeability, the first term in the right-hand side does not depend on  $i$ . Thus

$$\begin{aligned} \mathbb{E} \left[ \int_0^t (\mathbb{E}^0 [|\Delta X_s^i|^2])^{1/2} \{dK_s^N + dK_s\} \mathbf{1}_{\Omega_N} \right] &= \mathbb{E} \left[ \int_0^t \frac{1}{N} \sum_{i=1}^N (\mathbb{E}^0 [|\Delta X_s^i|^2])^{1/2} \{dK_s^N + dK_s\} \mathbf{1}_{\Omega_N} \right] \\ &\leq \mathbb{E} \left[ \int_0^t \left( \frac{1}{N} \sum_{i=1}^N \mathbb{E}^0 [|\Delta X_s^i|^2] \right)^{1/2} \{dK_s^N + dK_s\} \mathbf{1}_{\Omega_N} \right] \\ &\leq \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq t} \frac{1}{N} \sum_{i=1}^N \mathbb{E}^0 [|\Delta X_s^i|^2] \mathbf{1}_{\Omega_N} \right] \mathbb{E}^{1/2} [(K_t^N \mathbf{1}_{\Omega_N} + K_t)^2]. \end{aligned}$$

By (95), (96) and the bounds on  $K^N$  and  $K$ , this finally implies that

$$\mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Delta X_s^i|^2 \mathbf{1}_{\Omega_N} \right] \leq C(T) \left( \mathbb{E}^{1/4} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right] + \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq T} W_2^2(\bar{\mu}_s^N, \bar{\mu}_s) \right] \right).$$

To complete the proof, it remains to show that (93) holds. We have, by Itô's formula (85),

$$\begin{aligned} H(\mu_T^N) &= H(\mu_0^N) + \frac{1}{N} \sum_{i=1}^N \int_0^T D_\mu H(\mu_s^N)(X_s^i) \cdot b(s, X_s^i) ds \\ &\quad + \frac{1}{2N} \sum_{i=1}^N \int_0^T \text{Tr} ((\sigma_0 \sigma_0^*(s, X_s^i) + \sigma_1 \sigma_1^*(s, X_s^i)) \partial_y D_\mu H(\mu_s^N)(X_s^i)) ds \\ &\quad + \frac{1}{2N^2} \sum_{i=1}^N \int_0^T \text{Tr} (\sigma_0 \sigma_0^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^i, X_s^i)) ds \\ &\quad + \frac{1}{2N^2} \sum_{i,j=1}^N \int_0^T \text{Tr} (\sigma_1(s, X_s^j) \sigma_1^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^j, X_s^i)) ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^T D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_0(s, X_s^i) dB_s^i \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^T D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_1(s, X_s^i) dW_s + \frac{1}{N} \sum_{i=1}^N \int_0^T |D_\mu H(\mu_s^N)(X_s^i)|^2 dK_s^N. \end{aligned}$$

Note that, by (53) and on  $\{H(\mu_s^N) = 0\}$ , one has

$$\frac{1}{N} \sum_{i=1}^N |D_\mu H(\mu_s^N)(X_s^i)|^2 = \int_{\mathbb{R}^n} |D_\mu H(\mu_s^N)(x)|^2 \mu_s^N(dx) \geq \beta^2.$$

On the other hand, by (52) and the  $L^\infty$  bound on  $b$ ,

$$\left| \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_s^N)(X_s^i) \cdot b(s, X_s^i) \right| \leq \frac{C}{N} \sum_{i=1}^N |D_\mu H(\mu_s^N)(X_s^i)| \leq C.$$

As  $\sigma_0$  and  $\sigma_1$  are bounded and (52) holds, we obtain in the same way

$$\begin{aligned} & \left| \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_s^N)(X_s^i) \cdot b(s, X_s^i) \right| \\ & + \left| \frac{1}{2N} \sum_{i=1}^N \text{Tr} (\sigma_0 \sigma_0^*(s, X_s^i) + \sigma_1 \sigma_1^*(s, X_s^i)) \partial_y D_\mu H(\mu_s^N)(X_s^i) \right| \\ & + \left| \frac{1}{2N^2} \sum_{i=1}^N \text{Tr} (\sigma_0 \sigma_0^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^i, X_s^i)) \right| \\ & + \left| \frac{1}{2N^2} \sum_{i,j=1}^N \int_0^T \text{Tr} (\sigma_1(s, X_s^j) \sigma_1^*(s, X_s^i) D_{\mu\mu}^2 H(\mu_s^N)(X_s^j, X_s^i)) ds \right| \leq C. \end{aligned}$$

As, in addition,  $H$  is bounded, we finally have

$$\beta^2 K_T^N \leq C - \frac{1}{N} \sum_{i=1}^N \int_0^T D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_0(s, X_s^i) dB_s^i - \frac{1}{N} \sum_{i=1}^N \int_0^T D_\mu H(\mu_s^N)(X_s^i) \cdot \sigma_1(s, X_s^i) dW_s,$$

which yield (93).  $\square$

**3.3. Neumann Problem on Wasserstein space and Feynman-Kac formula.** We connect here our reflected process with a PDE on the Wasserstein space. For this we assume in addition to **(Hcl)** that  $b$  and  $\sigma$  are deterministic, continuous in time and let  $\mathcal{O} := \{\mu \in \mathcal{P}^2(\mathbb{R}^n), H(\mu) > 0\}$ . Given a bounded, continuous map  $G : \overline{\mathcal{O}} \rightarrow \mathbb{R}$  we consider the map  $u : [0, T] \times \overline{\mathcal{O}} \rightarrow \mathbb{R}$  defined by

$$u(t_0, \mu_0) = \mathbb{E} \left[ G([X_T^{t_0, X_0} | W]) \right], \quad (97)$$

where  $(X_s^{t_0, X_0}, K_s^{t_0, X_0})$  solves the reflected SDE (51) on  $[t_0, T]$  with initial condition  $X_{t_0}^{t_0, X_0} = X_0$  and  $[X_0] = \mu_0$ . Thanks to Remark 12, the uniqueness in law holds for (51) so that (97) is defined without ambiguity. By the semi-group property, we have, for any  $\mathcal{F}^W$ -stopping time  $\tau \geq t_0$ ,

$$u(t_0, \mu_0) = \mathbb{E} [u(\tau, [X_\tau^{t_0, X_0} | W])]. \quad (98)$$

So we can expect that  $u$  is, in a suitable sense, a solution to the following Neumann problem on the Wasserstein space:

$$\begin{cases} (i) & (\partial_t + \mathcal{A})u(t, \mu) = 0 & \text{in } (0, T) \times \mathcal{O}, \\ (ii) & \int_{\mathbb{R}^n} D_\mu u(t, \mu)(y) \cdot D_\mu H(\mu)(y) \mu(dy) = 0, & \text{in } (0, T) \times \partial\mathcal{O}, \\ (iii) & u(T, \mu) = G(\mu), & \text{in } \mathcal{O}, \end{cases} \quad (99)$$

where the operator  $\mathcal{A}$  is given, for any smooth function  $\phi : [0, T] \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$ , by

$$\begin{aligned} \mathcal{A}\phi(t, \mu) := & \int_{\mathbb{R}^n} b(t, z) \cdot D_\mu \phi(t, \mu)(z) d\mu(z) + \frac{1}{2} \int_{\mathbb{R}^n} \text{Tr} [(a_0(t, z) + a_1(t, z)) \partial_z D_\mu \phi(t, \mu)(z)] d\mu(z) \\ & + \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} \text{Tr} (D_{\mu\mu}^2 u(t, \mu)(z, z') \sigma_1(t, z) \sigma_1^*(t, z')) \mu(dz) \mu(dz'), \end{aligned} \quad (100)$$

with  $a_i = \sigma_i \sigma_i^*$  ( $i = 0, 1$ ).

Our results are the following : we first show that the map  $u$  given by (97) satisfies (99) in the viscosity sense. Then we provide a Feynman-Kac formula showing that any classical solution of (99), if it exists, is given by (97). Note that the question of the uniqueness of the viscosity solution of (99) is not considered in this work.

In order to show that the map  $u$  given by (97) satisfies (99) in a viscosity sense, let us introduce the following definition (inspired from Definition 11.17 of [13]).

**Definition 19.** We say that a continuous function  $u$  is a viscosity solution of (99) if

- (i)  $u$  is continuous on  $[0, T] \times \overline{\mathcal{O}}$ ;

(ii) for any  $(t, \mu)$  in  $(0, T) \times \mathcal{P}^2(\mathbb{R}^n)$ , for any test function  $\varphi : [0, T] \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  (see Definition 11.17 of [13] for the class of test functions) such that  $u - \varphi$  has a local minimum (resp. max) in  $(t, \mu)$  we have

$$\begin{cases} (\partial_t + \mathcal{A})\varphi(t, \mu) \leq 0, \text{ (resp. } \geq 0) \text{ in } \mathcal{O}, \\ \min \left\{ (\partial_t + \mathcal{A})\varphi(t, \mu) ; \int_{\mathbb{R}^n} D_\mu H(\mu)(z) \cdot D_\mu \varphi(t, \mu)(z) d\mu(z) \right\} \leq 0, \text{ (resp. } \geq 0) \text{ in } \partial\mathcal{O}; \end{cases} \quad (101)$$

(iii)  $u(T, \mu) = G(\mu)$  in  $\mathcal{O}$ .

**Theorem 20.** *If (Hcl) holds and if, in addition, the coefficients  $b$  and  $\sigma$  are continuous in time and deterministic, then the map  $u$  defined by (97) is a viscosity solution of (99) in the sense of Definition 19.*

*Proof.* Let us first check that  $u$  is continuous. Let  $\mu_0, \nu_0 \in \mathcal{P}_2(\mathbb{R}^n)$  and let  $[X_0] = \mu_0$  and  $[Y_0] = \nu_0$ . By Proposition 15, there exists a  $C > 0$  such that:

$$\mathbb{E}[|X_t^{t_0, X_0} - X_t^{t_0, Y_0}|^2] \leq C\mathbb{E}[|X_0 - Y_0|^2].$$

Combining this inequality with the dynamic programming principle in (98), it is easy to see that the map  $u$  is a continuous function.

Let us now prove that  $u$  solves (99) in the viscosity sense. Suppose that  $\varphi : [0, T] \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  is a test function such that  $u - \varphi$  have a local minimum in  $(t, \mu)$ . We can assume, without loss of generality, that:  $\varphi(t, \mu) = u(t, \mu)$  (which can be done by translating  $\varphi$ ). Let  $r > 0$  be such that  $u(s, \nu) \geq \varphi(s, \nu)$  for  $s \in [t, t+r]$  and  $W_2(\nu, \mu) \leq r$ . Let us set  $\mu_s := [X_s^{t, \mu} | W]$  and let  $\tau$  be the  $\mathcal{F}^W$ -stopping time

$$\tau = \inf\{s \geq t, W_2(\mu_s, \mu) \leq r\}.$$

Finally, for  $h \in (0, r]$ , we set  $\tau_h = \tau \wedge (t+h)$ . Note that, in view of the exponential estimate in Theorem 11 and our assumptions on the coefficients of the reflected SDE, it is not difficult to check that  $\mathbb{E}[\tau_h - t]/h \rightarrow 1$ . On the one hand we have from (98):

$$\varphi(t, \mu) = u(t, \mu) = \mathbb{E}[u(\tau_h, \mu_{\tau_h})] \geq \mathbb{E}[\varphi(\tau_h, \mu_{\tau_h})]. \quad (102)$$

On the other hand, by applying Itô's formula (16), we get:

$$\begin{aligned} \varphi(\tau_h, \mu_{\tau_h}) &= \varphi(t, \mu) + \int_t^{\tau_h} (\partial_t + \mathcal{A})\varphi(r, \mu_r) dr \\ &\quad + \int_t^{\tau_h} \int_{\mathbb{R}^n} D_\mu H(\mu_r)(z) \cdot D_\mu \varphi(r, \mu_r)(z) \mu_r(dz) dK_r \\ &\quad + \int_{t_0}^{\tau_h} \int_{\mathbb{R}^n} \sigma_1^T(s, z) D_\mu \varphi(r, \mu_r)(z) \mu_r(dz) \cdot dW_r. \end{aligned}$$

Plugging this equality into (102) and taking expectation we find:

$$0 \geq \mathbb{E}\left[\int_t^{\tau_h} (\partial_t + \mathcal{A})\varphi(r, \mu_r) dr + \int_t^{\tau_h} \int_{\mathbb{R}^n} D_\mu H(\mu_r)(z) \cdot D_\mu \varphi(r, \mu_r)(z) \mu_r(dz) dK_r\right].$$

Assume now that  $\mu$  is such that  $H(\mu) > 0$ . Then choosing  $r > 0$  small enough we have  $H(\mu_s) > 0$  a.s. in  $[t, \tau]$ , so that  $dK([t, \tau_h]) = 0$ . Hence

$$\mathbb{E}\left[\int_t^{\tau_h} (\partial_t + \mathcal{A})\varphi(r, \mu_r) dr\right] \leq 0.$$

Dividing by  $h > 0$ , letting  $h \rightarrow 0$  and using that  $\mathbb{E}[\tau_h - t]/h \rightarrow 1$ , we eventually have

$$(\partial_t + \mathcal{A})\varphi(t, \mu) \leq 0. \quad (103)$$

Assume now that  $\mu$  is such that  $H(\mu) = 0$  and that

$$\min \left\{ (\partial_t + \mathcal{A})\varphi(t, \mu) ; \int D_\mu \varphi(t, \mu)(z) d\mu(z) \right\} > 0.$$

Changing  $r > 0$  if necessary, there exists  $\alpha > 0$  such that, for any  $s \in [t, t+r]$  and  $W_2(\nu, \mu) \leq r$ , we have

$$\min \left\{ (\partial_t + \mathcal{A})\varphi(s, \nu) ; \int D_\mu \varphi(s, \nu)(z) d\nu(z) \right\} \geq \alpha.$$

Then

$$0 \geq \mathbb{E} \left[ \int_t^{\tau_h} (\partial_t + \mathcal{A})\varphi(r, \mu_r) dr + \int_t^{\tau_h} \int_{\mathbb{R}^n} D_\mu H(\mu_r)(z) \cdot D_\mu \varphi(r, \mu_r)(z) d\mu_r(z) dK_r \right] \geq \alpha \mathbb{E}[(\tau_h - t)].$$

As  $\mathbb{E}[(\tau_h - t)/h] \rightarrow 1$ , we find a contradiction.  $\square$

In order to establish a kind of reverse statement, we now prove a Feynman-Kac formula.

**Proposition 21.** *Assume that  $u$  is a classical solution to (99). Then  $u$  is given by (97).*

*Proof.* Let us set  $\mu_s := [X_s^{t_0, X_0} | W]$ . Then, by Itô's formula (16):

$$\begin{aligned} u(T, \mu_T) &= u(t_0, \mu_{t_0}) + \int_{t_0}^T (\partial_t + \mathcal{A})u(s, \mu_s) + \int_{t_0}^T \int_{\mathbb{R}^n} D_\mu u(s, \mu_s)(y) \cdot D_\mu H(\mu_s)(y) \mu_s(dy) dK_s \\ &\quad + \int_{t_0}^T \int_{\mathbb{R}^n} \sigma_1^T(s, x) D_\mu H(\mu_s)(x) \mu_s(dx) \cdot dW_s, \end{aligned}$$

where, by (99)-(ii),

$$\begin{aligned} &\int_{t_0}^T \int_{\mathbb{R}^n} D_\mu u(s, \mu_s)(y) \cdot D_\mu H(\mu_s)(y) \mu_s(dy) dK_s \\ &= \int_{t_0}^T \mathbf{1}_{\{\mu_s \in \partial \mathcal{O}\}} \left( \int_{\mathbb{R}^n} D_\mu u(s, \mu_s)(y) \cdot D_\mu H(\mu_s)(y) \mu_s(dy) \right) dK_s = 0. \end{aligned}$$

From the equation satisfied by  $u$ , we obtain

$$u(t_0, \mu_0) = u(t_0, \mu_{t_0}) = \mathbb{E}[u(T, \mu_T)] = \mathbb{E}[G(\mu_T)].$$

$\square$

#### 4. Backward SDEs with normal reflexion in law.

In this part, we are interested in Backward SDE constrained in law with normal reflexion. Namely, on  $[0, T]$ ,  $T > 0$  we consider the following problem:

$$\begin{cases} Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dB_s + \int_t^T D_\mu H([Y_s])(Y_s) dK_s, & 0 \leq t \leq T, \\ H([Y_t]) \geq 0, & 0 \leq t \leq T, \quad \int_0^T H([Y_s]) dK_s = 0, \end{cases} \quad (104)$$

where  $B$  is a Brownian motion on a given probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  endowed with the natural filtration  $\mathcal{F}^B$ , the processes  $Y, Z$  are of dimension  $n$  and  $n \times d$  respectively,  $f : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \rightarrow \mathbb{R}^n$  and  $H : \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$ .

We call solution of (104) a triple of progressively measurable processes  $(Y, Z, K)$  taking values in  $\mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathbb{R}$  such that  $K$  is deterministic, continuous and nondecreasing with  $K_0 = 0$ . We study the system under the following assumptions, referred as assumptions **(A)** in the following:

**(A1)** The functions  $f : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \rightarrow \mathbb{R}^n$  is Lipschitz w.r.t.  $(y, z)$  uniformly in time and  $\omega$  and adapted to the filtration  $\mathcal{F}^B$  for fixed  $(y, z)$  and:

$$\mathbb{E} \left[ \int_0^T |f(s, 0, 0)|^2 ds \right] < +\infty.$$

**(A2)** The function  $H : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}$  is fully  $\mathcal{C}^2$  (see the introduction for the definition) and

- there exist  $0 < \beta \leq M$  and  $\eta > 0$  such that

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n), \quad \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx) \leq M^2, \quad (105)$$

and

$$\forall \mu \in \mathcal{P}^2(\mathbb{R}^n) \text{ with } -\eta \leq H(\mu) \leq 0, \quad \beta^2 \leq \int_{\mathbb{R}^n} |D_\mu H(\mu)|^2(x) \mu(dx), \quad (106)$$

- there exists  $C \geq 0$  such that

$$\forall X \in L^2, \quad \mathbb{E} [|\mathrm{D}_\mu H([X])(X) - \mathrm{D}_\mu H([Y])(Y)|^2] \leq C \mathbb{E} [|X - Y|^2]. \quad (107)$$

**(A3)** The function  $H$  is concave: for all  $\mu, \nu$  in  $\mathcal{P}_2(\mathbb{R}^n)$

$$\forall X \sim \mu, Y \sim \nu, \quad H(\nu) - H(\mu) - \mathbb{E} [\mathrm{D}_\mu H([X])(X) \cdot (Y - X)] \leq 0. \quad (108)$$

**(A4)** The terminal value  $\xi$  is  $\mathcal{F}_T$ -measurable and

$$H([\xi]) \geq 0, \quad \mathcal{M}_2([\xi]) < +\infty.$$

*Remark 22.* Let us note that:

- (i) It follows from (106) and **(A4)** that there exists  $\tilde{\lambda}$  in  $\mathcal{P}^2(\mathbb{R}^n)$  such that  $H(\tilde{\lambda}) > 0$ ;
- (ii) ‘‘Conversely’’, if  $\{H \geq -\eta\}$  is bounded and if there exists  $\tilde{\lambda} \in \mathcal{P}^2(\mathbb{R}^n)$  such that  $H(\tilde{\lambda}) > 0$  then (106) holds.

*Remark 23.* In the following, we often use a r.v.  $\tilde{\Lambda}$  with law  $\tilde{\lambda}$  as a pivotal point to obtain certain estimates. This leads us to use several times Itô’s formula on quantities like  $|Y_t - \tilde{\Lambda}|^2$ . To do so, we work as follows: we pick a r.v.  $\tilde{\Lambda} \sim \tilde{\lambda}$  independent of  $\mathcal{F}^B$  on an atomless probability space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}})$ . Then, we apply Itô’s formula on the space  $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \geq 0}, \tilde{\mathbb{P}}) = (\Omega \times \tilde{\Omega}, \mathcal{F} \otimes \tilde{\mathcal{F}}, (\mathcal{F}_t \otimes \tilde{\mathcal{F}}_t)_{t \geq 0}, \mathbb{P} \otimes \tilde{\mathbb{P}})$  and denote by  $\tilde{\mathbb{E}}$  and  $\tilde{\mathbb{E}}$  the associated expectations.

In the following, this framework will be referred as *we work on the enlarged filtered probability space of Remark 23*.

*Proof of Remark 22.* The first assertion is obvious pushing forward the measure  $[\xi]$  along  $\mathrm{D}_\mu H$ . Concerning the second assertion, let  $\lambda$  be such that  $-\eta \leq H(\lambda) \leq 0$ . Let  $X, X_0 \in L^2$  be such that  $[X] = \lambda$ ,  $[X_0] = \tilde{\lambda}$ . Then, by the concavity of  $H$ ,

$$\begin{aligned} H(\tilde{\lambda}) &\leq H(\lambda) + \mathbb{E} [\partial_\mu H(\lambda)(X) \cdot (X_0 - X)] \\ &\leq \mathbb{E}^{1/2} [|\partial_\mu H(\lambda)(X)|^2] \mathbb{E}^{1/2} [|X_0 - X|^2]. \end{aligned}$$

As  $X$  is bounded in  $L^2$  by some constant  $C$  and  $H(\tilde{\lambda}) > 0$ , this proves (106) with  $\beta = H(\tilde{\lambda})/C > 0$ .

Next, let us come back to the concavity assumption **(A3)**. Let now  $\mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{R}^n)^N$  and  $H^N$  be the finite dimensional projection of  $H$ :

$$H^N(\mathbf{x}) := H\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i}\right).$$

We have that for all  $\mathbf{x}, \mathbf{y}$  in  $(\mathbb{R}^n)^N$ :

$$H^N(\mathbf{y}) - H^N(\mathbf{x}) - \frac{1}{N} \sum_{i=1}^N \mathrm{D}_\mu H(\mu_{\mathbf{x}}^N)(x_i)(x_i - y_i) \leq 0, \quad (109)$$

since

$$\partial_{x_i} H^N(\mathbf{x}) = \frac{1}{N} \mathrm{D}_\mu H\left(\frac{1}{N} \sum_{\ell=1}^N \delta_{x_\ell}\right)(x_i), \quad (110)$$

this means that the mapping  $\mathbf{x} \mapsto H^N(\mathbf{x})$  is concave in the classical sense. Hence,  $D_{\mathbf{x}}^2 H^N(\mathbf{x})$  is non-positive, where

$$\partial_{x_i x_j}^2 H^N(\mathbf{x}) = \frac{1}{N} \partial_z \mathrm{D}_\mu H\left(\frac{1}{N} \sum_{\ell=1}^N \delta_{x_\ell}\right)(x_i) \delta_{i,j} + \frac{1}{N^2} \mathrm{D}_\mu^2 H\left(\frac{1}{N} \sum_{\ell=1}^N \delta_{x_\ell}\right)(x_i, x_j). \quad (111)$$

□

#### 4.1. Existence and uniqueness of the solution.

**Theorem 24.** *Under (A) the BSDE with normal reflexion in law (104) has a unique square integrable solution such that  $K$  is deterministic.*

**Proposition 25.** *Let  $p \geq 2$ . Assume that assumptions (A) hold true and assume in addition that  $\mathcal{M}_p([\xi]) + \mathcal{H}_p$  is finite. Suppose further that  $\mathbb{E} \left[ \int_0^T |f(s, 0, 0)|^p ds \right] < +\infty$ . Then, there exists a constant  $C_{p,T} := C_{p,T} \left( \mathcal{M}_p([\xi]), \mathcal{H}_p, \mathbb{E} \left[ \int_0^T |f(s, 0, 0)|^p ds \right] \right) > 0$  such that*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t|^p \right] \leq C_{p,T}. \quad (112)$$

*Proof.* Let us start with the uniqueness result. Set  $(Y, Z, K)$  and  $(Y', Z', K')$  two solutions of (104) and let us denote  $\Delta Y_t = Y_t - Y'_t$ ,  $\Delta Z_t = Z_t - Z'_t$  and  $\Delta K_t = K_t - K'_t$ . Let  $\alpha \in \mathbb{R}$ , applying Itô's formula on  $e^{\alpha t} |\Delta Y_t|^2$  one obtains

$$\begin{aligned} e^{\alpha t} |\Delta Y_t|^2 &= \int_t^T (-\alpha e^{\alpha s} |\Delta Y_s|^2 + 2e^{\alpha s} (\Delta Y_s) \cdot \Delta f(s, Y_s, Z_s)) ds - 2 \int_t^T e^{\alpha s} (\Delta Y_s) \cdot (\Delta Z_s dB_s) \\ &\quad - \int_t^T e^{\alpha s} |\Delta Z_s|^2 ds + 2 \int_t^T e^{\alpha s} (\Delta Y_s) \cdot (\Delta [D_\mu H([Y_s])(Y_s) dK_s]), \end{aligned}$$

where  $\Delta [D_\mu H([Y_s])(Y_s) dK_s] = D_\mu H([Y_s])(Y_s) dK_s - D_\mu H([Y'_s])(Y'_s) dK'_s$ . By using classical arguments and assuming that  $\alpha \geq 2\|f\|_{\text{Lip}} + 2\|f\|_{\text{Lip}}^2$ , we derive:

$$\begin{aligned} e^{\alpha t} |\Delta Y_t|^2 + \frac{1}{2} \int_t^T e^{\alpha s} |\Delta Z_s|^2 ds &\leq -2 \int_t^T e^{\alpha s} (\Delta Y_s) \cdot (\Delta Z_s dB_s) \\ &\quad + 2 \int_t^T e^{\alpha s} (\Delta Y_s) \cdot (\Delta [D_\mu H([Y_s])(Y_s) dK_s]). \end{aligned}$$

From the concavity on  $H$ , thanks to the Skorokhod condition and since  $\forall s \in [0, T]$ ,  $H([Y'_s]) \geq 0$  we have

$$\int_t^T e^{\alpha s} \mathbb{E}[(\Delta Y_s) \cdot D_\mu H([Y_s])(Y_s) dK_s] \leq \int_t^T e^{\alpha s} (H([Y_s]) - H([Y'_s])) dK_s \leq 0,$$

and the same arguments lead to

$$- \int_t^T e^{\alpha s} \mathbb{E}[(\Delta Y_s) \cdot D_\mu H([Y'_s])(Y'_s) dK'_s] \leq 0.$$

Hence,

$$\mathbb{E} \left[ e^{\alpha t} |\Delta Y_t|^2 + \frac{1}{2} \int_t^T e^{\alpha s} |\Delta Z_s|^2 ds \right] \leq 0,$$

and uniqueness of  $\{Y, Z\}$  follows.

Let us now deal with the uniqueness of the processes  $K, K'$ . We aim at reproducing the approach implemented in the proof of Theorem 8. However, we cannot use the chain rule on the Wasserstein space because of the lack of needed integrability of the processes  $Z$  and  $Z'$ . To overcome this problem, we are lead to apply classical Itô's formula on i.i.d. copies of  $(Y, Z)$ . Define  $\{(\bar{Y})^i, (\bar{Z})^i\}_{1 \leq i \leq N}$  as  $N$  copies of  $(Y, Z)$ . Writing  $\bar{\mathbf{Y}} = (\bar{Y}^1, \dots, \bar{Y}^N)^*$ ,  $\bar{\mathbf{Z}} = (\bar{Z}^1, \dots, \bar{Z}^N)^*$  and  $\bar{\mu}_t^N = N^{-1} \sum_{i=1}^N \delta_{(\bar{Y}_t)^i}$  we have, from

classical Itô's formula

$$\begin{aligned}
\frac{1}{N} \sum_{i=1}^N \int_s^t |D_\mu H(\bar{\mu}_u^N)(\bar{Y}_u^i)|^2 dK_u &= H(\bar{\mu}_s^N) - H(\bar{\mu}_t^N) - \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^N)(\bar{Y}_u^i) \cdot f(u, Y_u, Z_u) du \\
&\quad + \frac{1}{2} \int_s^t \text{Tr} [D_x^2 H^N(\bar{Y}_u) \bar{Z}_u (\bar{Z}_u)^*] du \\
&\quad + \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^N)(\bar{Y}_u^i) \cdot (\bar{Z}_u)^i dB_u^i \\
&= \frac{1}{N} \sum_{i=1}^N \int_s^t |D_\mu H(\bar{\mu}_u^N)(\bar{Y}_u^i)|^2 dK'_u.
\end{aligned}$$

We have, passing to the limit  $N \rightarrow +\infty$ , that

$$\int_s^t \mathbb{E} \left[ |D_\mu H(\mu_u)(Y_u)|^2 \right] dK_u = \int_s^t \left[ |D_\mu H(\mu_u)(Y_u)|^2 \right] dK'_u.$$

We can thus repeat the end of the proof of uniqueness in Theorem 8 to deduce that  $K = K'$ . The result follows.

Let us now handle the existence part. The proof is divided into three parts: we first assume that the driver  $f$  is space independent and bounded and construct a solution thanks to a penalization approximation; we then extend the result thanks to a truncation argument when it is in  $\mathbb{L}^2(\Omega, \mathbb{L}^2([0, T], \mathbb{R}^n))$ ; we finally show, thanks to a Picard iteration, that the result holds true under **(A1)**.

*Step 1 : Existence for bounded and space independent generator.* In what follows we start by assuming that

$$\text{(T1), } \forall (\omega, s, y, z) \in \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}, f(\omega, s, y, z) = f(\omega, s), \quad \exists \kappa > 0, |f(\omega, s)| \leq \kappa, \mathbb{P} - a.s..$$

In this case, we construct a solution through a penalization approach. For  $k \geq 1$ , let  $\psi_k : \mathbb{R} \rightarrow \mathbb{R}_+$  be the function defined by

$$\psi_k(x) = r \text{ if } x \leq -1/k, \quad \psi_k(x) = -krx, \text{ if } -1/k \leq x \leq 0, \quad \psi_k(x) = 0, \text{ if } x \geq 0.$$

Note that the function  $\psi_k$  depends on the constant  $r > 0$  which will be chosen later. Let  $(Y^k, Z^k)$  be the solution to the following BSDE:

$$Y_t^k = \xi + \int_t^T f(s) ds - \int_t^T Z_s^k dB_s + \int_t^T D_\mu H([Y_s^k])(Y_s^k) \psi_k(H([Y_s^k])) ds, \quad 0 \leq t \leq T. \quad (113)$$

We have the following Proposition whose proof will be given at the end of this Step.

**Proposition 26.** *Under (A), for any  $k \geq 1$  there exists a unique solution of (113) satisfying*

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t^k|^2 + \int_0^T |Z_s^k|^2 ds \right] \leq C_r.$$

Moreover, under **(T1)**, there exists  $r > 0$  such that for all  $k \geq 1$  and  $0 \leq t \leq T$  we have

$$H([Y_t^k]) \geq -1/k. \quad (114)$$

Introducing

$$K_t^k = \int_0^t \psi_k(H([Y_s^k])) ds,$$

we rewrite the previous BSDE as

$$Y_t^k = \xi + \int_t^T f(s) ds - \int_t^T Z_s^k dB_s + \int_t^T D_\mu H([Y_s^k])(Y_s^k) dK_s^k, \quad 0 \leq t \leq T.$$

Let  $k, \ell$  in  $\mathbb{N}^*$  be fixed and set  $\Delta Y_t := Y_t^k - Y_t^\ell$  and  $\Delta Z_t = Z_t^k - Z_t^\ell$ . Applying Itô's formula to  $|\Delta Y_t|^2$ , and using Young's inequality, we obtain:

$$\begin{aligned}
 |\Delta Y_t|^2 + \int_t^T |\Delta Z_s|^2 ds &\leq 2 \int_t^T \Delta Y_s \cdot D_\mu H([Y_s^k])(Y_s^k) dK_s^k \\
 &\quad - 2 \int_t^T \Delta Y_s \cdot D_\mu H([Y_s^\ell])(Y_s^\ell) dK_s^\ell - 2 \int_t^T \Delta Y_s \cdot \Delta Z_s dB_s.
 \end{aligned} \tag{115}$$

From the  $L$ -concavity of  $H$ , Proposition 26 and the definition of  $K^k$  we have

$$\begin{aligned}
 \mathbb{E} \left[ \int_t^T \Delta Y_s \cdot D_\mu H([Y_s^k])(Y_s^k) dK_s^k \right] &\leq \mathbb{E} \left[ \int_t^T (H([Y_s^k]) - H([Y_s^\ell])) \psi_k(H([Y_s^k])) ds \right] \\
 &\leq 0 + \frac{r}{\ell}(T-t),
 \end{aligned}$$

since  $x\psi_k(x) \leq 0$  and  $\psi_k$  is bounded by  $r$ . Arguing similarly we have

$$-\mathbb{E} \left[ \int_t^T \Delta Y_s \cdot D_\mu H([Y_s^\ell])(Y_s^\ell) dK_s^\ell \right] \leq \frac{r}{k}(T-t).$$

Hence we deduce that for some  $C_{T,r} > 0$ :

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ |\Delta Y_t|^2 + \int_0^T |\Delta Z_s|^2 ds \right] \leq C_{T,r} \left( \frac{1}{k} + \frac{1}{\ell} \right),$$

and coming back to (115), we get from BDG's inequality together with the fact that  $\sup_k \psi_k \leq r$  that there exists  $C'_{T,r} > 0$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\Delta Y_t|^2 + \int_0^T |\Delta Z_s|^2 ds \right] \leq C'_{T,r} \left( \frac{1}{\sqrt{k}} + \frac{1}{\sqrt{\ell}} \right).$$

Thus,  $\{(Y^k, Z^k)\}_{k \geq 1}$  is a Cauchy sequence in  $\mathcal{S}^2 \times \mathcal{M}^2$ . Let us denote by  $(Y, Z)$  its limit. Since  $\psi_k$  is bounded by  $r$  for all  $k$ ,  $K^k$  is Lipschitz with  $|K^k|_{\text{Lip}} \leq r$ . Hence, by Ascoli-Arzelà theorem, up to a subsequence,  $(K^k)$  converges towards a non decreasing, Lipschitz continuous function  $K$  in  $\mathcal{C}([0, T], \mathbb{R})$ .

It is straightforward to check that  $(Y, Z, K)$  solves (104). Indeed,

- $H([Y_t]) = \lim_{k \rightarrow \infty} H([Y_t^k]) \geq 0$  by Proposition 26.
- the Skorokhod condition is also satisfied : since  $x\psi_k(x) \leq 0$ ,

$$0 \leq \int_0^T H([Y_s]) dK_s = \lim_{k \rightarrow \infty} \int_0^T H([Y_s^k]) dK_s^k = \lim_{k \rightarrow \infty} \int_0^T H([Y_s^k]) \psi_k(H([Y_s^k])) ds \leq 0.$$

*Proof of Proposition 26.* First note that existence and uniqueness of a solution follows from [8]. We work on the enlarged filtered probability space of Remark 23. Arguing as in the proof of uniqueness, for  $\alpha$  large enough, we have, for  $0 \leq t \leq T$ ,

$$\begin{aligned}
 &e^{\alpha t} |Y_t^k - \tilde{\Lambda}|^2 + \frac{1}{2} \int_t^T e^{\alpha s} |Z_s^k|^2 ds \\
 &\leq e^{\alpha T} |\xi - \tilde{\Lambda}|^2 + \int_t^T e^{\alpha s} |f(s, \tilde{\Lambda}, 0)|^2 ds + 2 \int_t^T e^{\alpha s} (Y_s^k - \tilde{\Lambda}) \cdot D_\mu H([Y_s^k])(Y_s^k) dK_s^k - 2 \int_t^T e^{\alpha s} (Y_s^k - \tilde{\Lambda}) \cdot Z_s^k dB_s.
 \end{aligned} \tag{116}$$

Using the  $L$ -concavity of  $H$  together with the fact that  $H(\tilde{\lambda}) \geq 0$  we have,

$$\mathbb{E} \left[ \int_t^T e^{\alpha s} (Y_s^k - \tilde{\Lambda}) \cdot D_\mu H([Y_s^k])(Y_s^k) dK_s^k \right] \leq \int_t^T e^{\alpha s} (H([Y_s^k]) - H(\tilde{\lambda})) \psi_k(H([Y_s^k])) ds \leq 0,$$

since  $x\psi_k(x) \leq 0$ . It follows that there exists  $C_\lambda := C(\mathcal{M}_2(\tilde{\lambda}), T, (\mathbf{A})) > 0$  independent of  $k$  and  $r$  such that

$$\sup_{0 \leq t \leq T} \mathbb{E} \left[ |Y_t^k|^2 + \int_0^T |Z_s^k|^2 ds \right] \leq C_\lambda. \tag{117}$$



Coming back to the estimate (116) and using BDG's inequality, we deduce that, for some constant  $C_{a,r} := C(a, T, r, (\mathbf{A})) > 0$  independent of  $k$ ,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t^k|^2 + \int_0^T |Z_s^k|^2 ds \right] \leq C_{a,r},$$

in other words, the sequence  $\{(Y^k, Z^k)\}_{k \geq 1}$  is bounded in  $\mathcal{S}^2 \times H^2$ .

Let us now prove (114). Our approach here is in the same spirit as the proof of estimate (48) in paragraph 2.2. Again, due to the lack of needed integrability of the process  $Z^k$ , we apply classical Itô's formula and then use the concavity property of the function  $H$ , see the computations below.

Assume that there exists  $0 < t_0 < T$  such that  $H([Y_{t_0}^k]) < -1/k$ . Define  $t = \inf\{u \geq t_0 : H([Y_u^k]) \geq -1/k\}$  and  $s = \sup\{u \leq t : H([Y_u^k]) \leq H([Y_{t_0}^k]) \vee -\eta\}$ . Since  $H([\xi]) \geq 0$ , and  $-\eta < -1/k$  we have that  $0 < s < t < T$  and  $-\eta \leq H([Y_u^k]) \leq -1/k$  on  $[s, t]$ .

Let  $\{(\bar{Y}^k)^i, (\bar{Z}^k)^i, \bar{f}^i\}_{1 \leq i \leq N}$  be  $N$  copies of  $(Y^k, Z^k, f)$ . Writing  $\bar{\mathbf{Y}}^k = ((\bar{Y}^k)^1, \dots, (\bar{Y}^k)^N)^*$  and  $\bar{\mathbf{Z}}^k = ((\bar{Z}^k)^1, \dots, (\bar{Z}^k)^N)^*$  and  $\bar{\mu}_t^{N,k} = N^{-1} \sum_{i=1}^N \delta_{(\bar{Y}_t^k)^i}$  we have, from classical Itô's formula

$$\begin{aligned} H(\bar{\mu}_s^{N,k}) &= H(\bar{\mu}_t^{N,k}) + \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i \cdot f^i(u) du \\ &\quad - \frac{1}{2} \int_s^t \text{Tr} [D_{\mathbf{x}}^2 H^N(\bar{\mathbf{Y}}_u^k) \bar{\mathbf{Z}}_u^k (\bar{\mathbf{Z}}_u^k)^*] du \\ &\quad - \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i \cdot (\bar{Z}_u^k)^i dB_u^i \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_s^t |D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i|^2 dK_u. \end{aligned}$$

Hence we have from (111):

$$\begin{aligned} H(\bar{\mu}_s^{N,k}) &\geq H(\bar{\mu}_t^{N,k}) + \frac{1}{2} \int_s^t \text{Tr} [D_{\mathbf{x}}^2 H^{N,k}(\bar{\mathbf{Y}}_u^k) \bar{\mathbf{Z}}_u^k (\bar{\mathbf{Z}}_u^k)^*] du \\ &= H(\bar{\mu}_t^{N,k}) + \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i \cdot f^i(u) du \\ &\quad - \frac{1}{N} \sum_{i=1}^N \int_s^t D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i \cdot (\bar{Z}_u^k)^i dB_u^i \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_s^t |D_\mu H(\bar{\mu}_u^{N,k})(\bar{Y}_u^k)^i|^2 dK_u^k. \end{aligned} \tag{118}$$

Taking the expectation and then the limit over  $N$  on both sides leads, see the proof of Theorem 5.98 of [12], to

$$\begin{aligned} H([Y_s^k]) &\geq H([Y_t^k]) + \int_s^t \mathbb{E} [D_\mu H([Y_u^k])(Y_u^k) \cdot f(u)] du \\ &\quad + \int_s^t \mathbb{E} [|D_\mu H([Y_u^k])(Y_u^k)|^2] \psi_k(H([Y_u^k])) du. \end{aligned}$$

From **(A)** we get that there exists a constant  $C := C(a, T, (\mathbf{A})) > 0$  such that

$$H([Y_s^k]) \geq H([Y_t^k]) - C(t-s) + \int_s^t \mathbb{E} [|D_\mu H([Y_u^k])(Y_u^k)|^2] \psi_k(H([Y_u^k])) du.$$

Using now **(T1)** and the definition of  $\psi_k$  we get

$$H([Y_s^k]) \geq H([Y_t^k]) - C(t-s) + \beta^2(t-s)r.$$

Since  $H([Y_s^k]) \leq H([Y_{t_0}^k]) \vee -\eta < -1/k \leq H([Y_t^k])$  this implies  $-C + \beta^2 r < 0$  which cannot hold for all  $r$  large enough and then leads to a contradiction.  $\square$

*Remark 27.* Note that from the proof of the above Proposition and Step1 remain true for a generator  $f$  satisfying  $\sup_{t \leq T} \mathbb{E}[|f(t)|^2] < +\infty$ .

*Step 2 : from space independent generator in  $\mathbb{L}^\infty$  to  $H^2$ .* We now prove, using a truncation argument, that the BSDE (104) admits a solution when

$$\mathbf{(T2)}, \quad \forall (\omega, s, y, z) \in \Omega \times [0, T] \times \mathbb{R}^n \times \mathbb{R}^{n \times d}, \quad f(\omega, s, y, z) = f(\omega, s), \quad \mathbb{E} \left[ \int_0^T |f|^2(s) ds \right] < +\infty.$$

Let  $Y_t^m$  be given by

$$Y_t^m = \xi + \int_t^T f(s) \mathbf{1}_{|f(s)| \leq m} ds - \int_t^T Z_s^m dB_s + \int_t^T D_\mu H([Y_s^m])(Y_s^m) dK_s^m.$$

For each  $m$  in  $\mathbb{N}$ , this BSDE admits a unique solution from *step 1*. Moreover we have for any  $m, \ell \geq 0$ , by applying Itô's formula on  $e^{\alpha t} |Y_t^\ell - Y_t^m|^2$ ,  $\alpha \geq 1$ , that:

$$\begin{aligned} & e^{\alpha t} |Y_t^\ell - Y_t^m|^2 + \int_t^T e^{\alpha s} |Z_s^\ell - Z_s^m|^2 ds \\ &= 2 \int_t^T e^{\alpha s} (Y_s^\ell - Y_s^m) \cdot (f(s) \mathbf{1}_{|f(s)| \leq \ell} - f(s) \mathbf{1}_{|f(s)| \leq m}) ds - 2 \int_t^T e^{\alpha s} (Y_s^\ell - Y_s^m) \cdot (Z_s^\ell - Z_s^m) dB_s \\ & \quad + 2 \int_t^T e^{\alpha s} (Y_s^\ell - Y_s^m) \cdot D_\mu H([Y_s^\ell])(Y_s^\ell) dK_s^\ell - 2 \int_t^T e^{\alpha s} (Y_s^\ell - Y_s^m) \cdot D_\mu H([Y_s^m])(Y_s^m) dK_s^m \\ & \quad - \alpha \int_t^T e^{\alpha s} |Y_s^\ell - Y_s^m|^2 ds. \end{aligned} \tag{119}$$

Taking now the expectation, using  $L$ -concavity, Skorokhod property and the fact that for each  $k$  the marginals  $\{[Y_t^k]\}_{0 \leq t \leq T}$  satisfy the constraint, we deduce that:

$$\begin{aligned} & \int_t^T \mathbb{E} [e^{\alpha s} (Y_s^\ell - Y_s^m) \cdot D_\mu H([Y_s^\ell])(Y_s^\ell)] dK_s^\ell - \int_t^T e^{\alpha s} \mathbb{E} [(Y_t^\ell - Y_t^m) \cdot D_\mu H([Y_s^m])(Y_s^m)] dK_s^m \\ & \leq 2 \mathbb{E} \left[ \int_t^T e^{\alpha s} (H([Y_s^\ell]) - H([Y_s^m])) (dK_s^\ell - dK_s^m) \right] \leq 0, \end{aligned} \tag{120}$$

so that,

$$\begin{aligned} & \sup_{t \leq T} \mathbb{E} \left[ e^{\alpha t} |Y_t^\ell - Y_t^m|^2 + \int_t^T e^{\alpha s} |Z_s^\ell - Z_s^m|^2 ds \right] \\ & \leq C \int_0^T \mathbb{E} [ |f(s) \mathbf{1}_{|f(s)| \leq \ell} - f(s) \mathbf{1}_{|f(s)| \leq m}|^2 ] ds. \end{aligned}$$

Let us now show that  $K_t^m$  is bounded, uniformly in  $m$ . To do so, we work on the enlarged filtered probability space of Remark 23. One has,

$$\begin{aligned} e^{\alpha t} |Y_t^m - \tilde{\Lambda}|^2 + \int_t^T e^{\alpha s} |Z_s^m|^2 ds &= e^{\alpha T} |\xi - \tilde{\Lambda}|^2 + 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot f(s) \mathbf{1}_{|f(s)| \leq m} ds \\ & \quad - 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot Z_s^m dB_s + 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot D_\mu H([Y_s^m])(Y_s^m) dK_s^m \\ & \quad - \alpha \int_t^T e^{\alpha s} |Y_s^m - \tilde{\Lambda}|^2 ds. \end{aligned} \tag{121}$$

Taking the expectation on both side, using Young's inequality and  $L$ -concavity together with the Skorokhod condition, we obtain for  $\alpha$  large enough that

$$\sup_{t \leq T} \bar{\mathbb{E}} [e^{\alpha t} |Y_t^m - \tilde{\Lambda}|^2] + \mathbb{E} \left[ \int_0^T e^{\alpha s} |Z_s^m|^2 ds \right] + e^{\alpha T} H(\tilde{\lambda})(K_T^m - K_t^m) \leq C \left( \bar{\mathbb{E}} [|\xi - \tilde{\Lambda}|^2] + \mathbb{E} \left[ \int_0^T |f(s)|^2 ds \right] \right).$$

Hence, coming back to (119), taking the supremum in time, the expectation and choosing  $\alpha$  large enough lead to

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} e^{\alpha t} |Y_t^\ell - Y_t^m|^2 + \int_0^T e^{\alpha s} |Z_s^\ell - Z_s^m|^2 ds \right] \\ & \leq C \left( \mathbb{E} \left[ \int_0^T |f(s) \mathbf{1}_{|f(s)| \leq \ell} - f(s) \mathbf{1}_{|f(s)| \leq m}|^2 ds \right] + \mathbb{E} \left[ \sup_{t \leq T} \left| \int_t^T e^{\alpha s} (Y_t^\ell - Y_t^m) \cdot (Z_s^\ell - Z_s^m) dB_s \right| \right] \right. \\ & \quad + \int_0^T \mathbb{E}^{1/2} [ |Y_s^\ell - Y_s^m|^2 ] \mathbb{E}^{1/2} [ |D_\mu H([Y_s^\ell])(Y_s^\ell)|^2 ] dK_s^\ell \\ & \quad \left. + \int_t^T \mathbb{E}^{1/2} [ |Y_t^\ell - Y_t^m|^2 ] \mathbb{E}^{1/2} [ |D_\mu H([Y_s^m])(Y_s^m)|^2 ] dK_s^m \right), \end{aligned} \quad (122)$$

thanks to Cauchy-Schwarz's inequality and since  $K^m, K^\ell$  are deterministic. Since from BDG's inequality and Young's inequality one has

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} \left| \int_t^T e^{\alpha s} (Y_t^\ell - Y_t^m) \cdot (Z_s^\ell - Z_s^m) dB_s \right| \right] \\ & \leq \epsilon \mathbb{E} \left[ \sup_{t \leq T} e^{\alpha t} |Y_t^\ell - Y_t^m|^2 \right] + \frac{C}{\epsilon} \mathbb{E} \left[ \int_0^T e^{\alpha s} |Z_s^\ell - Z_s^m|^2 ds \right], \end{aligned} \quad (123)$$

we deduce that  $\{Y^m, Z^m\}$  is a Cauchy sequence in  $\mathcal{S}^2 \times \mathbb{H}^2$  that converges to some limit  $\{Y, Z\}$ .

Let us now deal with the convergence of the process  $K^m$ . By the convergence of  $\{Y_t^m, Z_t^m\}$  it is clear that the sequence of processes  $(L^m)_{m \geq 0}$  defined by

$$\forall t \geq 0, \quad L_t^m := \int_0^t D_\mu H(\mu_s^m)(Y_s^m) dK_s^m,$$

converges in  $\mathcal{S}^2$  to some process  $L$ . Set  $h(\mu) := (\int |D_\mu H(\mu)(\cdot)|^2 d\mu \vee \beta^2)^{-1}$ , for all  $t \geq 0$ ,  $\varphi_t := [D_\mu H(\mu_t)(Y_t)]^*$ . Let us denote by  $(\varphi_t^M, h^M(\mu_t))_{t \geq 0}$  the discretized version of  $(\varphi_t, h(\mu_t))_{t \geq 0}$  along a subdivision  $(t_k)_{0 \leq k \leq M}$  of  $[0, T]$  of stepsize  $1/M$ . Let  $M > 0$ , for  $l, m \geq 0$  we have

$$\begin{aligned} |K_t^m - K_t^l| &= \left| \mathbb{E} \left[ \int_0^t h(\mu_s^m) \varphi_s^m dL_s^m \right] - \mathbb{E} \left[ \int_0^t h(\mu_s^l) \varphi_s^l dL_s^l \right] \right| \\ &= \left| \mathbb{E} \left[ \int_0^t \{h(\mu_s^m) \varphi_s^m - h(\mu_s^l) \varphi_s^l\} dL_s^m \right] + \mathbb{E} \left[ \int_0^t \{h(\mu_s^l) \varphi_s^l - h(\mu_s) \varphi_s\} dL_s^m \right] \right. \\ & \quad + \mathbb{E} \left[ \int_0^t \{h(\mu_s) \varphi_s - h^M(\mu_s) \varphi_s^M\} dL_s^m \right] \\ & \quad \left. + \mathbb{E} \left[ \int_0^t \{h^M(\mu_s) \varphi_s^M\} (dL_s^m - dL_s^l) \right] + \mathbb{E} \left[ \int_0^t \{h^M(\mu_s) \varphi_s^M - h(\mu_s^l) \varphi_s^l\} dL_s^l \right] \right|. \end{aligned}$$

Note now that for all  $k \geq 0$ , we have  $dL_s^k = D_\mu H(\mu_s^k)(Y_s^k) dK_s^k$  with  $K^k$  deterministic. Writing  $L^m$  and  $L^l$  as this, one can inverse the integration and expectation operators in the first, second, third and fifth terms of the above r.h.s. We can thus deduce, taking first the limit superior in  $m, l$  and then the limit superior over  $M$ , that the process  $K^m$  converges to some deterministic continuous process  $K$ .

It is not hard to see that this limit  $(Y, Z, K)$  is a solution of (104) under the standing assumption on the generator  $f$  assumed in this part.

*Step 3: from space independent generator to space dependent generator.* We now only assume that the generator  $f$  satisfies our assumptions **(A)**. We are now going to construct a solution to BSDE (104) by using a Picard iteration. Set  $(Y^0, Z^0) = (0, 0)$  and define recursively  $(Y^m, Z^m)$  as the unique solution of

$$\begin{aligned} Y_t^m &= \xi + \int_t^T f(s, Y_s^{m-1}, Z_s^{m-1}) ds - \int_t^T Z_s^m dB_s + \int_t^T D_\mu H([Y_s^m])(Y_s^m) dK_s^m, \\ \forall t \geq 0, \quad H([Y_t^m]) &\geq 0, \quad \int_t^T H([Y_s^m]) dK_s^m = 0, \quad m \geq 1. \end{aligned}$$

Let us denote for any  $m \geq 0$ :  $\Delta Y^{m+1} = Y^{m+1} - Y^m$  and  $\Delta Z^{m+1} = Z^{m+1} - Z^m$ .

On the one hand, applying Itô's formula on  $e^{\alpha t} |\Delta Y_t^{m+1}|^2$ ,  $\alpha \geq 1$ , we have

$$\begin{aligned} & e^{\alpha t} |\Delta Y_t^{m+1}|^2 + \int_t^T e^{\alpha s} |\Delta Z_s^{m+1}|^2 ds \\ &= -\alpha \int_t^T e^{\alpha s} |\Delta Y_s^{m+1}|^2 ds + 2 \int_t^T e^{\alpha s} (\Delta Y_s^{m+1}) \cdot (f(s, Y_s^m, Z_s^m) - f(s, Y_s^{m-1}, Z_s^{m-1})) ds \\ & \quad - 2 \int_t^T e^{\alpha s} (\Delta Y_s^{m+1}) \cdot (\Delta Z_s^{m+1}) dB_s + \int_t^T e^{\alpha s} (\Delta Y_s^{m+1}) D_\mu H([Y_s^{m+1}]) (Y_s^{m+1}) dK_s^{m+1} \\ & \quad - \int_t^T e^{\alpha s} (\Delta Y_s^{m+1}) D_\mu H([Y_s^m]) (Y_s^m) dK_s^m. \end{aligned} \tag{124}$$

Then, using Young's inequality and arguing as in (120) it can be deduced that for a suitable choice of  $\alpha$  that

$$\mathbb{E} \left[ e^{\alpha t} |\Delta Y_t^{m+1}|^2 + \int_t^T e^{\alpha s} |\Delta Z_s^{m+1}|^2 ds \right] \leq \frac{1}{4} \mathbb{E} \left[ \int_t^T e^{\alpha s} (|\Delta Y_t^m|^2 + |\Delta Z_s^m|^2) ds \right]. \tag{125}$$

Let us now *work on the enlarged filtered probability space of Remark 23*. We get that

$$\begin{aligned} e^{\alpha t} |Y_t^m - \tilde{\Lambda}|^2 + \int_t^T e^{\alpha s} |Z_s^m|^2 ds &= e^{\alpha T} |\xi - \tilde{\Lambda}|^2 + 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot f(s, Y_s^{m-1}, Z_s^{m-1}) ds \\ & \quad - \alpha \int_t^T e^{\alpha s} |Y_s^m - \tilde{\Lambda}|^2 ds - 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot Z_s^m dB_s \\ & \quad + 2 \int_t^T e^{\alpha s} (Y_s^m - \tilde{\Lambda}) \cdot D_\mu H([Y_s^m]) (Y_s^m) dK_s^m. \end{aligned}$$

By using the Lipschitz property of  $f$  together with Young's inequality and the Skorokhod condition, we derive, for  $\alpha$  large enough:

$$\begin{aligned} & \sup_{t \leq T} \bar{\mathbb{E}} [e^{\alpha t} |Y_t^m - \tilde{\Lambda}|^2] + \mathbb{E} \left[ \int_0^T e^{\alpha s} |Z_s^m|^2 ds \right] + e^{\alpha T} H(\tilde{\Lambda})(K_T^m - K_t^m) \\ & \leq C \left( \bar{\mathbb{E}} [|\xi - \tilde{\Lambda}|^2] + \bar{\mathbb{E}} \left[ \int_0^T |f(s, \tilde{\Lambda}, 0)|^2 ds \right] \right), \end{aligned} \tag{126}$$

so that  $K_T^m$  is uniformly bounded.

Finally, coming back to (124), we can argue as in (120) (122) to deduce from (125) and (126) that  $\{Y_t^m, Z_t^m\}$  is a converging sequence in  $\mathcal{S}^2 \times \mathcal{H}^2$  that converges to some limit  $\{Y, Z\}$ . We obtain that  $K^m$  converges to some process  $K$  by using the same scheme as we did at the end of *step 2*. Again, it follows from standard computations that the limit  $\{Y, Z, K\}$  is a solution (104).

□

*Proof of Proposition 25.* Applying Itô's formula to  $e^{\alpha t}|Y_t|^p$  for some  $p \geq 2$  we have

$$\begin{aligned} & e^{\alpha t}|Y_t|^p + \frac{p(p-1)}{2} \int_t^T e^{\alpha s}|Y_s|^{p-2}|Z_s|^2 ds \\ = & |\xi|^p + p \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot f(s, Y_s, Z_s) ds - p \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot Z_s dB_s \\ & + p \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot D_\mu H([Y_s])(Y_s) dK_s - \alpha \int_t^T e^{\alpha s}|Y_s|^p ds. \end{aligned}$$

On the one hand, Lipschitz continuity assumption on  $f$  together with Young's inequality yields

$$\begin{aligned} & \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot f(s, Y_s, Z_s) ds \\ \leq & C_p \left\{ \int_t^T |f(s, 0, 0)|^p ds + \int_t^T \left(1 + \frac{1}{\epsilon}\right) e^{\alpha s}|Y_s|^p + \epsilon e^{\alpha s}|Y_s|^{p-2}|Z_s|^2 ds \right\}, \end{aligned}$$

for any  $\epsilon > 0$ . On the other hand, still from Young's inequality we have

$$\sup_{t \leq T} \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot D_\mu H([Y_s])(Y_s) dK_s \leq \int_0^T e^{\alpha s} \frac{p}{p-1} |Y_s|^p + \frac{1}{p} |D_\mu H([Y_s])(Y_s)|^p dK_s.$$

Choosing  $\epsilon$  small enough and then  $\alpha$  large enough gives

$$\begin{aligned} & e^{\alpha t}|Y_t|^p + \frac{p(p-1)}{4} \int_t^T e^{\alpha s}|Y_s|^{p-2}|Z_s|^2 ds \\ \leq & |\xi|^p + C_p \left\{ \int_t^T e^{\alpha s}|Y_s|^p + |D_\mu H([Y_s])(Y_s)|^p dK_s + \int_t^T |f(s, 0, 0)|^p ds \right\} \\ & - p \int_t^T e^{\alpha s}|Y_s|^{p-2}(Y_s) \cdot Z_s dB_s. \end{aligned} \tag{127}$$

Taking the expectation in the above inequality and using Gronwall's lemma applied to the continuous maps  $s \mapsto K_s$  (see Lemma 4 in [21] or Theorem 17.1 in [2]) eventually lead to

$$\begin{aligned} & \sup_{t \leq T} \mathbb{E} \left[ e^{\alpha t}|Y_t|^p + \frac{p(p-1)}{4} \int_t^T e^{\alpha s}|Y_s|^{p-2}|Z_s|^2 ds \right] \\ \leq & C_{p,T} \left\{ \mathbb{E}[|\xi|^p] + \mathcal{H}_p K_T + \mathbb{E} \left[ \int_0^T |f(s, 0, 0)|^p ds \right] \right\}. \end{aligned}$$

Coming back to (127), taking the supremum in time, the expectation and using then BDG' inequality and Young's inequality yield, together with the above estimate, to the result.  $\square$

**4.2. Interacting particle system constrained in mean field.** Let us consider the following Skorohod problem in mean field:

$$\begin{cases} Y_t^i = \tilde{\xi}^i + \int_t^T f(s, Y_s^i, Z_s^{i,i}) ds - \int_t^T \sum_{j=1}^N Z_s^{i,j} dB_s^j + \int_t^T D_\mu H(\mu_s^N)(Y_s^i) dK_s^N, \\ \forall t \in [0, T]: \quad \mu_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}, \quad H(\mu_t^N) \geq 0, \quad \int_0^T H(\mu_s^N) dK_s^N = 0, \quad 1 \leq i \leq N, \end{cases} \tag{128}$$

where for each  $i, j$ ,  $Z_s^{i,j}$  is a  $n \times d$  matrix,  $\{B^i\}_{1 \leq i \leq N}$  are  $N$  independent  $d$ -dimensional Brownian motions and  $K^N$  is a continuous non decreasing process. The terminal conditions  $\{\tilde{\xi}^i\}_{1 \leq i \leq N}$  are  $\mathcal{F}_T^i$ -measurable ( $\{\mathcal{F}_t^i\}$  being the augmented natural filtration of  $B^i$ ) independent r.v. having second order moment and satisfying  $H\left(N^{-1} \sum_{i=1}^N \delta_{\tilde{\xi}^i}\right) \geq 0$ .

*Remark 28.* Making the analogy with classical mean field approximation of McKean-Vlasov processes, a natural choice for the family of terminal conditions  $\{\tilde{\xi}^i\}$  is the family  $\{\xi^i\}$  of copies of the r.v.  $\xi$  in (104). Nevertheless, the condition  $H([\xi]) \geq 0$  is not sufficient to ensure that (128) is indeed a solution of the Skorokhod problem in mean field since it does not imply  $H\left(N^{-1} \sum_{i=1}^N \delta_{\xi^i}\right) \geq 0$ . To overcome the problem, we may work only on the set  $\Omega_N = \left\{H\left(N^{-1} \sum_{i=1}^N \delta_{\xi^i}\right) \geq -\eta_N\right\}$  which becomes, for a suitable choice of  $\eta_N$ , of full measure when  $N \rightarrow +\infty$ . This hence leads to obtain an asymptotic solution to the Skorokhod problem in mean field, but not for any  $N$  (see the approach in paragraph 3.2). In this part, we decide to tackle the problem from a different point of view: we show how to construct, from the family of  $\{\xi^i\}$ , a family  $\{\tilde{\xi}^i\}$  satisfying the Skorokhod condition and whose empirical measure tends to  $[\xi]$ , provided some additional integrability conditions on the data.

By “some additional integrability conditions”, we mean that we assume fourth-order moment on the data  $\xi$  and  $D_\mu H$ . In comparison with the assumptions needed in the forward case (see paragraph 3.2), such conditions may appear to be the price to pay to solve the Skorokhod problem for each  $N$ . It seems to us that this is not the case, but something more related to the backward setting. For instance, still in comparison with the result obtained in the forward case, the set  $\Omega_N$  is not independent of the  $B^i$  and it is *a priori* quite intricate to obtain higher order moment on the process  $K^N$  without any additional assumptions on  $D_\mu H$  (recalling the lack of integrability of  $Z$ ).

**Initialization, well posedness of the particle system constrained in mean field and estimates.**

We first state the following Lemma which shows how to construct the  $\{\tilde{\xi}^i\}$  from the  $\{\xi^i\}$ .

**Lemma 29.** *Suppose that assumptions (A) hold. Given  $N$  copies  $\{\xi^i\}$  of  $\xi$ , there exists a family of random variable  $\{\tilde{\xi}^i\}$  satisfying*

$$H\left(\frac{1}{N} \sum_{i=1}^N \delta_{\tilde{\xi}^i}\right) \geq 0,$$

and if we assume in addition that  $\mathcal{M}_4([\xi]) + \mathcal{H}_4$  is finite then there exists  $C := C((\mathbf{A}), \mathcal{M}_4([\xi]), \mathcal{H}_4) > 0$  such that

$$\mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2\right] \leq C \mathbb{E}^{1/2} [W_2^2(\bar{\mu}_T^N, [\xi])], \quad \bar{\mu}_T^N := \frac{1}{N} \sum_{i=1}^N \delta_{\xi^i}. \quad (129)$$

*Proof.* To guarantee that the constraint is satisfied, the main idea consists in transporting the initial condition along the normal vector up to the set of constraint. Since we only assumed that the normal vector is non zero around the zero of  $H$  we have to proceed in two steps : consider first the case where the empirical measure is not “too far” from the constraint set and when it is too far away. Let  $\Omega_N = \{H(\bar{\mu}_0^N) \geq \eta_N\}$  where  $\eta_N \rightarrow 0$  and  $\eta_N^{-2} \mathbb{E} [W_2^2(\bar{\mu}_0^N, [\xi])] \rightarrow 0$  and let us push forward the empirical measure  $\bar{\mu}_T^N$  along the flow of  $dl_t^x = D_\mu H(\mu_{t^x}^N)(l_t^x) dt$ ,  $l_0^x = x$ ,  $x \in \mathbb{R}^{nN}$ , where  $\mu_{t^x}^N = N^{-1} \sum_{i=1}^N \delta_{l_t^{i,x}}$  and  $l_t^{i,x}$  denotes the  $i^{\text{th}}$   $n$ -dimensional component of  $l_t^x$ . For  $N$  large enough we have  $-\eta_N \geq -\eta$  so that on  $\Omega_N$  (106) is satisfied and it is therefore possible to find a positive  $\tilde{t} \leq \eta_N/\beta^2$  such that  $H(\mu_{\tilde{t}}^N) \geq 0$  where  $\mu_{\tilde{t}}^N := (\mathbb{R}^{nN} \ni x \mapsto l_{\tilde{t}}^x \in \mathbb{R}^{nN}) \# \bar{\mu}_T^N$ . We write  $\{\zeta^i\}$  the family of associated r.v.

Let us now handle the case when we are in  $\Omega_N^c$ . From Remark 22 there exists  $\tilde{\lambda}$  such that  $H(\tilde{\lambda}) \geq 0$ . We then set  $\tilde{\xi}^i = \zeta^i \mathbf{1}_{\Omega_N} + \tilde{\lambda} \mathbf{1}_{\Omega_N^c}$ . We have

$$\mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2\right] \leq M \frac{\eta_N^2}{\beta^4} + \mathbb{E}^{1/2} [\mathbf{1}_{\Omega_N^c}] \frac{1}{N} \sum_{i=1}^N \mathbb{E}^{1/2} [|\tilde{\lambda} - \xi^i|^4],$$

thanks to Cauchy-Schwarz inequality. Since  $H([\xi]) \geq 0$ ,

$$\mathbb{P}(\Omega_N^c) \leq \mathbb{P}(|H(\bar{\mu}_T^N) - H([\xi])| \geq \eta_N) \leq \frac{\mathbb{E} [W_2^2(\bar{\mu}_T^N, [\xi])]}{\eta_N^2},$$

we have, choosing  $\eta_N^2 = \mathbb{E}^{1/2} [W_2^2(\bar{\mu}_T^N, [\xi])]$ , that there exists a  $C := C((\mathbf{A}), \mathcal{M}_4([\xi]), \mathcal{H}_4) > 0$  such that

$$\mathbb{E}\left[\frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2\right] \leq C \mathbb{E}^{1/2} [W_2^2(\bar{\mu}_0^N, [\xi])].$$

□

Next, we have the following result.

**Proposition 30.** *Suppose that assumptions **(A)** hold and that the initial data  $\{\tilde{\xi}^i\}$  in (128) are given by Lemma 29. Then, the system (128) is well posed.*

*Proof.* The proof of this result is not new and relies on classical results. Let us indeed remark that (128) reads as a classical reflected BSDE in  $\mathbb{R}^{nN}$ , with normal reflexion in the constraint

$$\mathcal{K} = \left\{ \mathbf{x} = (x_1, \dots, x_N) \in (\mathbb{R}^n)^N, H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) \geq 0 \right\}.$$

Indeed, let us recall that, if

$$H^N(\mathbf{x}) = H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right),$$

then (see [11])

$$\partial_{x_i} H^N(\mathbf{x}) = \frac{1}{N} D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_i).$$

Therefore the vector  $-\left(D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_1), \dots, D_\mu H \left( \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right) (x_N)\right)$  is proportional to the outward normal to the set  $\mathcal{K}$  at the point  $(x_1, \dots, x_N) \in \partial\mathcal{K}$ . Existence and uniqueness of the solution are therefore immediate under our standing assumptions (cf. [18]).  $\square$

**Lemma 31** (Moment estimates). *Under assumption **(A)**, for any  $T > 0$ , there exists  $C := C((\mathbf{A}), T) > 0$  such that*

$$\mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |Y_t^i|^2 + \frac{1}{N} \sum_{i=1}^N \int_0^T \sum_{j=1}^N |Z_s^{i,j}|^2 ds \right] + \mathbb{E}[K_T^N] \leq C,$$

and

$$\sup_{N \geq 1} \mathbb{E}[(K_T^N)^2] \leq C.$$

*Proof.* Let  $p \geq 2$ , and let us work on the enlarged filtered probability space of Remark 23 (recalling that here, the filtration has also been enlarged to take into account the  $N$  independent Brownian motions). Applying Itô's formula on  $e^{\alpha t} |Y_t^i - \tilde{\Lambda}|^2$  we obtain, after summing over all the particles, that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N e^{\alpha t} |Y_t^i - \tilde{\Lambda}|^2 + \frac{1}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} \sum_{j=1}^N |Z_s^{i,j}|^2 ds \\ = & \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \tilde{\Lambda}|^2 + \frac{2}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot f(s, Y_s^i, Z_s^{i,i}) ds \\ & - \frac{2}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot \sum_{j=1}^N Z_s^{i,j} dB_s^j + \frac{2}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot D_\mu H(\mu_s^N)(Y_s^i) dK_s^N \\ & - \frac{\alpha}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} |Y_s^i - \tilde{\Lambda}|^2 ds. \end{aligned}$$

From  $L$ -concavity and Skorokhod condition we have,

$$\tilde{\mathbb{E}} \left[ \frac{1}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot D_\mu H(\mu_s^N)(Y_s^i) dK_s^N \right] \leq -\frac{1}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} H(\tilde{\lambda}) dK_s^N, \quad (130)$$

while Lipschitz continuity assumption on  $f$  together with Young's inequality yield

$$\begin{aligned} & 2 \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot f(s, Y_s^i, Z_s^{i,i}) ds \\ & \leq C \left\{ \int_t^T e^{\alpha s} |f(s, \tilde{\Lambda}, 0)|^2 ds + \int_t^T \left( 1 + \frac{1}{\epsilon} \right) e^{\alpha s} |Y_s^i - \tilde{\Lambda}|^2 + \epsilon e^{\alpha s} |Z_s^{i,i}|^2 ds \right\}, \end{aligned} \quad (131)$$

for any  $\epsilon > 0$ . Also, thanks to the construction in Lemma 29 we have

$$\frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \tilde{\Lambda}|^2 \leq |\xi|^2 + |\tilde{\Lambda}|^2 + \bar{t}M^2. \quad (132)$$

Choosing  $\epsilon$  small enough and  $\alpha$  large enough hence gives

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N e^{\alpha t} \mathbb{E}[|Y_t^i - \tilde{\Lambda}|^2] + \frac{1}{2N} \sum_{i=1}^N \int_t^T e^{\alpha s} \sum_{j=1}^N |Z_s^{i,j}|^2 ds + 2H(\tilde{\lambda})(K_T^N - K_t^N) \\ & \leq \frac{e^{\alpha T}}{N} \sum_{i=1}^N \mathbb{E}[|\tilde{\xi}^i - \tilde{\Lambda}|^2] + C \int_t^T e^{\alpha s} \mathbb{E}[|f(s, \tilde{\Lambda}, 0)|^2] ds - \frac{2}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} (Y_s^i - \tilde{\Lambda}) \cdot \sum_{j=1}^N Z_s^{i,j} dB_s^j. \end{aligned} \quad (133)$$

Taking expectation on both side of the above estimate we can deduce the first estimate. Let us now prove the second assertion. We come back to (118) and obtain

$$\begin{aligned} & H(\mu_t^N) - H(\mu_0^N) - \frac{1}{N} \sum_{i=1}^N \int_0^t D_\mu H(\mu_u^N)(Y_u^i) \cdot f(u, Y_u^i, Z_u^{i,i}) du \\ & + \frac{1}{N} \sum_{i=1}^N \int_0^t D_\mu H(\mu_u^N)(Y_u^i) \cdot \sum_{j=1}^N Z_u^{i,j} dB_u^j \geq \frac{1}{N} \sum_{i=1}^N \int_0^t |D_\mu H(\mu_u^N)(Y_u^i)|^2 dK_u^N. \end{aligned} \quad (134)$$

Note that  $dK^N(\{s, \text{ s.t. } H(\mu_s^N) > 0\}) = 0$  so that by (106) we obtain

$$\begin{aligned} & \beta^2 K_t^N \leq H(\mu_t^N) - H(\mu_0^N) - \frac{1}{N} \sum_{i=1}^N \int_0^t D_\mu H(\mu_u^N)(Y_u^i) \cdot f(u, Y_u^i, Z_u^{i,i}) du \\ & + \frac{1}{N} \sum_{i=1}^N \int_0^t D_\mu H(\mu_u^N)(Y_u^i) \cdot \sum_{j=1}^N Z_u^{i,j} dB_u^j. \end{aligned} \quad (135)$$

On the one hand

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} \left( \frac{1}{N} \sum_{i=1}^N \int_0^t D_\mu H(\mu_u^N)(Y_u^i) \cdot \sum_{j=1}^N Z_u^{i,j} dB_u^j \right)^2 \right] \\ & \leq \mathbb{E} \left[ \frac{4}{N^2} \int_0^T \sum_{j=1}^N \left| \sum_{i=1}^N |D_\mu H(\mu_u^N)(Y_u^i)| |Z_u^{i,j}| \right|^2 du \right] \\ & \leq \mathbb{E} \left[ \frac{4}{N^2} \int_0^T \sum_{j=1}^N \sum_{i=1}^N \{ |D_\mu H(\mu_u^N)(Y_u^i)|^2 \} \sum_{i=1}^N \{ |Z_u^{i,j}|^2 \} du \right] \\ & \leq 4\mathcal{H}_2 \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \int_0^T \sum_{j=1}^N |Z_u^{i,j}|^2 du \right]. \end{aligned}$$



On the other hand, using the Lipschitz continuity of  $f$  and Cauchy-Schwarz's inequality,

$$\begin{aligned} & \left( \int_0^T \frac{1}{N} \sum_{i=1}^N D_\mu H(\mu_u^N)(Y_u^i) \cdot f(s, Y_s^i, Z_s^{i,i}) du \right)^2 \\ & \leq C(\mathcal{H}_2) \left\{ \int_0^T |f(s, 0, 0)|^2 ds + \int_0^T \frac{1}{N} \sum_{i=1}^N |Y_s^i|^2 + \frac{1}{N} \sum_{i=1}^N |Z_s^{i,i}|^2 ds \right\}. \end{aligned}$$

Hence, choosing  $t = T$ , then squaring the expression and eventually taking the expectation lead, thanks to the two above estimates as well as the first assertion of this Lemma (using also (107)), to the result.  $\square$

**The mean field limit** Let  $\{B^i\}_{1 \leq i \leq N}$  be  $N$  independent  $d$ -dimensional Brownian motions,  $\{\xi^i\}_{1 \leq i \leq N}$  be  $N$  independent copies of  $\xi$  and set  $\{(\bar{Y}^i, \bar{Z}^i, K)\}_{1 \leq i \leq N}$  the solution of (104) with data  $\{(B^i, \xi^i)\}_{1 \leq i \leq N}$  (note that  $K$  does not depend on  $i$  since it only depends on the law of the data, which are identically distributed thanks to Remark 12). Let  $\{(Y^i, Z^i, K^N)\}_{1 \leq i \leq N}$  be the solution of (128) where the terminal data  $\{\tilde{\xi}^i\}$  therein are built from the  $\{\xi^i\}_{1 \leq i \leq N}$  thanks to Lemma 29. We aim at showing that the solution  $\{(Y^i, Z^i, K^N)\}$  is close to  $\{(\bar{Y}^i, \bar{Z}^i, K)\}$ .

Let us denote by  $\Delta Y^i := Y^i - \bar{Y}^i$  and  $\Delta Z^{i,j} := Z^{i,j} - \bar{Z}^{i,j}$  where  $\bar{Z}^{i,j} = 0_{\mathbb{R}^{n \times d}}$  when  $i \neq j$ . We have:

**Theorem 32.** *Suppose that assumptions **(A)** hold. Then, there exist  $C_T := C(T, \mathbf{(A)}) > 0$  such that*

$$\begin{aligned} & \mathbb{E} \left[ \sup_{t \leq T} \frac{1}{N} \sum_{i=1}^N \left\{ |\Delta Y_t^i|^2 + \int_t^T \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \right] \\ & \leq C_T \left\{ \mathbb{E}^{1/2} \left[ \sup_{0 \leq s \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] + \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] \right\}, \\ & \mathbb{E} \left[ \sup_{s \leq t} \left\{ |\Delta Y_s^i|^2 + \int_s^T \sum_{j=1}^N |\Delta Z_u^{i,j}|^2 du \right\} \right] \\ & \leq C_T \left\{ \mathbb{E}^{1/4} \left[ \sup_{t \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] + \mathbb{E}^{1/2} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] \right\}. \end{aligned}$$

**Lemma 33.** *Suppose that assumptions **(A)** hold. Assume further that there exists  $q > 4$  such that  $\mathcal{M}_q([\xi])$  is finite as well as  $\mathcal{H}_q$  and  $\int_0^T \mathbb{E}[|f(s, 0, 0)|^q] ds$ . Assume moreover that the data of the system are such that there exists  $p \geq 4$  and  $C_{p,Z} > 0$  such that  $\sup_{t \leq T} \mathbb{E}[|Z_t|^p] \leq C_{p,Z}$ . Then, there exists  $C_T := C(T, \mathbf{(A)}, \mathcal{M}_q([\xi]), \mathcal{H}_q, C_{p,Z}) > 0$  such that*

$$\left( \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] \right)^2 + \mathbb{E} \left[ \sup_{0 \leq s \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] \leq C_T \epsilon_N,$$

where  $\epsilon_N$  is given in Lemma 1 (see Remark 2).

*Remark 34.* The assumption on the integrability of  $Z$  could seem strange at first sight, especially because this process is part of the solution and not an input of the problem. This assumption relies on the control done for the discretization of the process  $Y$  to handle the convergence of the supremum in time of the Wasserstein distance. For instance, such property is satisfied when the terminal condition  $\xi$  and the driver  $f$  are Malliavin differentiable (see also Lemma 4.4 in [6] for further details).

*Proof of Lemma 33.* Let us first emphasize that the result is straightforward for the first term in the l.h.s. of the above estimate: it follows from Lemma 29 and Theorem 2 in [17] (see also Theorem 5.8 and Remark 5.9 in [12]).

Let us now deal with the second term in the above l.h.s. To handle this part, we have to check that assumptions in Lemma 1 are satisfied. Note first that, under our considered assumption, one can apply Proposition 25. It thus remains to check that (19) hold after taking into account Remark 2. These are

straightforward consequences of our standing assumptions provided the process  $K$  is Lipschitz continuous. Let us now explain why this is latter fact is true.

Let  $(Y, Z)$  be a given (part of) the unique solution of (104). It then holds that  $\sup_{t \leq T} \mathbb{E}[|f(t, Y_t, Z_t)|^2] < +\infty$  so that one can come back to *Step 1* of the proof of Theorem 24 and build the process  $K$  in such a way it is Lipschitz. By uniqueness, this proves that  $K$  is Lipschitz. The result then follows.  $\square$

*Proof of Theorem 32.* For  $\alpha$  large enough we have

$$\begin{aligned}
 & e^{\alpha t} |\Delta Y_t^i|^2 + \frac{1}{2} \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \\
 \leq & e^{\alpha T} |\tilde{\xi}^i - \xi^i|^2 + 2 \int_t^T e^{\alpha s} \Delta Y_s^i \cdot D_\mu H(\mu_s^N)(Y_s^i) dK_s^N \\
 & - 2 \int_t^T e^{\alpha s} \Delta Y_s^i \cdot D_\mu H(\mu_s)(\bar{Y}_s^i) dK_s - 2 \int_t^T e^{\alpha s} \Delta Y_s^i \cdot \sum_{j=1}^N \Delta Z_s^{i,j} dB_s^j.
 \end{aligned} \tag{136}$$

So that, summing over  $i$  leads to

$$\begin{aligned}
 & \frac{1}{N} \sum_{i=1}^N \left\{ e^{\alpha t} |\Delta Y_t^i|^2 + \frac{1}{2} \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \\
 \leq & \frac{e^{\alpha T}}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 + 2 \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \{ \Delta Y_s^i \cdot D_\mu H(\mu_s^N)(Y_s^i) \} dK_s^N \\
 & - 2 \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \{ \Delta Y_s^i \cdot D_\mu H(\mu_s)(\bar{Y}_s^i) \} dK_s - 2 \frac{1}{N} \sum_{i=1}^N \int_t^T e^{\alpha s} \Delta Y_s^i \cdot \sum_{j=1}^N \Delta Z_s^{i,j} dB_s^j \\
 =: & J_1^N + J_2^N(t, T) + J_3^N(t, T) + M_N(t, T).
 \end{aligned} \tag{137}$$

We first deal with the second term in the r.h.s. of (137). We have

$$\begin{aligned}
 J_2^N(t, T) &= \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \Delta Y_s^i \cdot D_\mu H(\mu_s^N)(Y_s^i) dK_s^N \leq \int_t^T e^{\alpha s} (H(\mu_s^N) - H(\bar{\mu}_s^N)) dK_s^N \\
 &\leq \int_t^T e^{\alpha s} (H(\mu_s) - H(\bar{\mu}_s^N)) dK_s^N
 \end{aligned} \tag{138}$$

from Skorokhod condition and since  $H(\mu_s) \geq 0$  for all  $0 \leq s \leq T$ . Thus,

$$J_2^N(t, T) \leq C e^{\alpha T} \sup_{t \leq s \leq T} W_2(\mu_s, \bar{\mu}_s^N) (K_T^N - K_t^N).$$

Next we handle the third term in the right hand side of (137) and split it into two parts

$$\begin{aligned}
 J_3^N(t, T) &= -2 \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \Delta Y_s^i \cdot D_\mu H(\mu_s)(\bar{Y}_s^i) dK_s \\
 &= -2 \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \Delta Y_s^i \cdot (D_\mu H(\mu_s)(\bar{Y}_s^i) - (D_\mu H(\bar{\mu}_s^N)(\bar{Y}_s^i))) dK_s \\
 &\quad - 2 \int_t^T e^{\alpha s} \frac{1}{N} \sum_{i=1}^N \Delta Y_s^i \cdot D_\mu H(\bar{\mu}_s^N)(\bar{Y}_s^i) dK_s =: I_1^N(t, T) + I_2^N(t, T).
 \end{aligned}$$

On the one hand we have from Cauchy-Schwarz's inequality, (84) and Young's inequality

$$\begin{aligned}
I_1^N(t, T) &\leq C \int_t^T e^{\alpha s} \left( \frac{1}{N} \sum_{i=1}^N |\Delta Y_s^i|^2 \right)^{1/2} W_2(\bar{\mu}_s^N, \mu_s) dK_s \\
&\leq C \int_t^T \left( \frac{1}{N} \sum_{i=1}^N e^{\alpha s} |\Delta Y_s^i|^2 \right)^{1/2} W_2(\bar{\mu}_s^N, \mu_s) dK_s \\
&\leq C \int_t^T \left\{ \frac{1}{N} \sum_{i=1}^N e^{\alpha s} |\Delta Y_s^i|^2 + W_2^2(\bar{\mu}_s^N, \mu_s) \right\} dK_s \\
&\leq C \int_t^T \frac{1}{N} \sum_{i=1}^N e^{\alpha s} |\Delta Y_s^i|^2 dK_s + C \sup_{t \leq s \leq T} W_2^2(\bar{\mu}_s^N, \mu_s) K_T.
\end{aligned}$$

On the other hand we have from  $L$ -concavity and since for all  $0 \leq s \leq T$ ,  $H(\mu_s^N) \geq 0$  and  $H(\mu_s) = 0$   $dK_s$ -a.e.

$$\begin{aligned}
I_2^N(t, T) &\leq 2 \int_t^T e^{\alpha s} (H(\bar{\mu}_s^N) - H(\mu_s^N)) dK_s \\
&\leq 2 \int_t^T e^{\alpha s} (H(\bar{\mu}_s^N) - H(\mu_s)) dK_s \\
&\leq 2e^{\alpha T} \sup_{t \leq s \leq T} W_2(\bar{\mu}_s^N, \mu_s) K_T.
\end{aligned}$$

Hence,

$$J_3^N(t, T) \leq C \int_t^T \frac{1}{N} \sum_{i=1}^N e^{\alpha s} |\Delta Y_s^i|^2 dK_s + C \sup_{t \leq s \leq T} \{W_2^2(\bar{\mu}_s^N, \mu_s) + W_2(\bar{\mu}_s^N, \mu_s)\}.$$

Bringing together the estimates on  $J_i^N$ ,  $1 \leq i \leq 3$  in (137), we obtain

$$\begin{aligned}
&\frac{1}{N} \sum_{i=1}^N \left\{ e^{\alpha t} |\Delta Y_t^i|^2 + \frac{1}{2} \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \\
&\leq \frac{e^{\alpha T}}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 + C e^{\alpha T} \sup_{t \leq s \leq T} W_2(\mu_s, \bar{\mu}_s^N) (K_T^N - K_t^N) \\
&\quad + C \int_t^T \frac{1}{N} \sum_{i=1}^N e^{\alpha s} |\Delta Y_s^i|^2 dK_s + C \sup_{t \leq s \leq T} \{W_2^2(\bar{\mu}_s^N, \mu_s) + W_2(\bar{\mu}_s^N, \mu_s)\} + M_N(t, T).
\end{aligned} \tag{139}$$

Hence, taking the expectation in (139), using Cauchy-Schwarz's inequality, Gronwall's lemma and then using Lemmas 29 and 31 lead to

$$\begin{aligned}
&\mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N \left\{ e^{\alpha t} |\Delta Y_t^i|^2 + \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \right] \\
&\leq C_T \left\{ \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] + \mathbb{E}^{1/2} \left[ \sup_{t \leq s \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] + \mathbb{E} \left[ \sup_{t \leq s \leq T} \{W_2^2(\bar{\mu}_s^N, \mu_s) + W_2(\bar{\mu}_s^N, \mu_s)\} \right] \right\}.
\end{aligned} \tag{140}$$

Now we have from BDG and Young inequality

$$\begin{aligned}
 \mathbb{E}[\sup_{t \leq T} M_N(t)] &\leq C\mathbb{E} \left[ \left( \frac{1}{N^2} \int_t^T \sum_{j=1}^N e^{2\alpha s} \left( \sum_{i=1}^N |\Delta Y_s^i| |\Delta Z_s^{i,j}| \right)^2 ds \right)^{1/2} \right] \\
 &\leq C\mathbb{E} \left[ \left( \sup_{t \leq T} \left\{ \frac{1}{N} \sum_{i=1}^N \{e^{\alpha t} |\Delta Y_s^i|^2\} \right\} \frac{1}{N} \sum_{i=1}^N \left\{ \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \right)^{1/2} \right] \\
 &\leq C\varepsilon \mathbb{E} \left[ \sup_{t \leq T} \frac{1}{N} \sum_{i=1}^N \{e^{\alpha t} |\Delta Y_s^i|^2\} \right] + \frac{C^2}{4\varepsilon} \frac{1}{N} \sum_{i=1}^N \left\{ \mathbb{E} \left[ \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right] \right\} \quad (141)
 \end{aligned}$$

for any  $\varepsilon > 0$ . Hence, coming back to (139) and taking first the supremum in time and then expectation we deduce from the above estimate, Lemma 31 and Gronwall's lemma that

$$\begin{aligned}
 &\mathbb{E} \left[ \sup_{t \leq T} \frac{1}{N} \sum_{i=1}^N \left\{ e^{\alpha t} |\Delta Y_t^i|^2 + \frac{1}{2} \int_0^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right\} \right] \quad (142) \\
 &\leq C_T \left\{ \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] + \mathbb{E}^{1/2} \left[ \sup_{t \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] + \mathbb{E} \left[ \sup_{t \leq s \leq T} \{W_2^2(\bar{\mu}_s^N, \mu_s) + W_2(\bar{\mu}_s^N, \mu_s)\} \right] \right\}.
 \end{aligned}$$

Note that we also have from above results, thanks to the exchangeability of the  $(Y^i, Z^i)$ , that

$$\mathbb{E} [e^{\alpha t} |\Delta Y_t^i|^2] + \frac{1}{2} \mathbb{E} \left[ \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 ds \right] \leq C_T \left\{ \mathbb{E} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] + \mathbb{E}^{1/2} \left[ \sup_{0 \leq t \leq T} W_2^2(\mu_t, \bar{\mu}_t^N) \right] \right\}. \quad (143)$$

Coming back to (136), taking first the supremum then the expectation we get, thanks to BDG's inequality and (141), that

$$\begin{aligned}
 &\mathbb{E} \left[ \sup_{t \leq T} \left\{ e^{\alpha t} |\Delta Y_t^i|^2 + \frac{1}{2} \int_t^T e^{\alpha s} \sum_{j=1}^N |\Delta Z_s^{i,j}|^2 du \right\} \right] \quad (144) \\
 &\leq \mathbb{E} [|\tilde{\xi}^i - \xi^i|^2] + 2\mathbb{E} \left[ \int_0^T e^{\alpha s} |\Delta Y_s^i| |D_\mu H(\mu_s^N)(Y_s^i)| (dK_s^N + dK_s) \right] + \mathbb{E}^{1/2} \left[ \sup_{t \leq T} W_2^2(\mu_t, \bar{\mu}_t^N) \right].
 \end{aligned}$$

Since by exchangeability of the  $(Y^i, Z^i)$ , Cauchy-Schwarz's inequality and (105)

$$\begin{aligned}
 &\mathbb{E} \left[ \int_t^T e^{\alpha s} |\Delta Y_s^i| |D_\mu H(\mu_s^N)(Y_s^i)| (dK_s^N + dK_s) \right] \quad (145) \\
 &= \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \int_t^T e^{\alpha s} |\Delta Y_s^i| |D_\mu H(\mu_s^N)(Y_s^i)| (dK_s^N + dK_s) \right] \\
 &\leq M^2 \mathbb{E} \left[ \sup_{s \leq T} e^{\alpha s} \left( \frac{1}{N} \sum_{i=1}^N |\Delta Y_s^i|^2 \right)^{1/2} ((K_T^N - K_t) + (K_T - K_t)) \right] \\
 &\leq 2e^{\alpha T} M^2 \mathbb{E}^{1/2} \left[ \sup_{s \leq T} \frac{1}{N} \sum_{i=1}^N |\Delta Y_s^i|^2 \right] \mathbb{E}^{1/2} [(K_T^N - K_t)^2 + (K_T - K_t)^2],
 \end{aligned}$$

we obtain, thanks to estimate in Proposition 25 and (142) that

$$\mathbb{E} \left[ \sup_{s \leq t} \left\{ e^{\alpha s} |\Delta Y_s^i|^2 + \frac{1}{2} \int_s^T e^{\alpha u} \sum_{j=1}^N |\Delta Z_u^{i,j}|^2 du \right\} \right] \leq C \left\{ \mathbb{E}^{1/4} \left[ \sup_{t \leq T} W_2^2(\mu_s, \bar{\mu}_s^N) \right] + \mathbb{E}^{1/2} \left[ \frac{1}{N} \sum_{i=1}^N |\tilde{\xi}^i - \xi^i|^2 \right] \right\}.$$

□

**4.3. Associated Obstacle problem on the Wasserstein space.** As we did for the (forward) normally constrained in law SDE, we aim at connecting our backward system with a PDE. To do so, we consider the Markovian setup, where the final condition (104) is of the form  $\phi(X_T, [X_T])$ , for some  $\phi : \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  and where  $X$  is the solution of an SDE driven by Lipschitz coefficients. Namely we consider for  $T > 0$ , for any r.v.  $X_0$  having second order moment, independent of the Brownian motion and any  $t$  in  $[0, T]$ , the Mean-Reflected Forward-Backward SDE (MR-FBSDE):

$$\left\{ \begin{array}{l} X_s^{t, X_0} = X_0 + \int_t^s b(r, X_r^{t, X_0}) dr + \int_t^s \sigma(r, X_r^{t, X_0}) dB_r, \quad s \in [t, T], \quad X_s^{t, X_0} = X_0, \quad s \in [0, t] \\ Y_s^{t, X_0} = \phi(X_T^{t, X_0}, [X_T^{t, X_0}]) + \int_s^T f(r, X_r^{t, X_0}, Y_r^{t, X_0}) dr - \int_s^T Z_r^{t, X_0} dB_r \\ \quad + \int_s^T D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, X_0}) dK_s^{t, X_0}, \quad s \in [t, T], \\ \forall s \in [0, T], \quad H([Y_s^{t, X_0}]) \geq 0, \quad \int_0^T H([Y_s^{t, X_0}]) dK_s^{t, [X_0]} = 0, \quad s \in [0, T], \end{array} \right. \quad (146)$$

where the superscript  $(t, X_0)$  stands for the initial condition of the SDE associated with  $X$  whose coefficients are supposed to be continuous in time and Lipschitz continuous in space (uniformly in time), where the coefficients of the backward component satisfy **(A)** (with  $d = 1$ ) and where  $\phi : \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  is a continuous and bounded function. Also, in this section we further assume that the following assumptions hold:

**(A'1)** The functions  $f : \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  is continuous and there exists a positive  $C_f$  such that for all  $(s, x, y)$  in  $[0, T] \times \mathbb{R}^n \times \mathbb{R} : |f(s, x, y)| \leq C_f(1 + |x| + |y|)$ ,

**(A'2)** There exists  $\beta > 0$  such that for all  $(x, \mu) \in \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n)$  we have  $\beta \leq D_\mu H(\mu)(x)$ .

In the following we say that assumptions **(A')** are in force if assumptions **(A'1)**, **(A'2)** and **(A)** hold.

We start by recalling that, according to Remark 12, uniqueness in law holds for (146). In this case, we aim at proving that there exists a decoupling field  $u : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \ni (t, x, \mu) \mapsto u(t, x, \mu) \in \mathbb{R}$  such that for any  $t$  in  $[0, T]$ , any r.v.  $X_0$  having second order moment and independent of the Brownian motion, we have for all  $s$  in  $[t, T]$   $u(s, X_s^{t, X_0}, [X_s^{t, X_0}]) = Y_s^{t, X_0}$  a.s..

Especially we are going to prove that  $u$  solves, in the viscosity sense, the following obstacle problem on the Wasserstein space:

$$\left\{ \begin{array}{l} \min \left\{ \{(\partial_t + \mathcal{L})u(t, x, \mu) + f(t, x, u(t, x, \mu))\}; \right. \\ \left. H(u(t, \cdot, \mu) \# \mu) \right\} = 0, \quad \text{on } [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n), \\ u(T, \cdot, \cdot) = \phi, \end{array} \right. \quad (147)$$

where  $\mathcal{L}$  is given by, for all smooth  $\varphi : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$

$$\begin{aligned} \mathcal{L}\varphi(t, x, \mu) &= \frac{1}{2} \int_{\mathbb{R}^n} \text{Tr} [((\sigma\sigma^*)(t, y)) \partial_y D_\mu \varphi(\mu)(t, x, \mu)(y)] \mu(dy) + \int_{\mathbb{R}^n} D_\mu \varphi(\mu)(t, x, \mu)(y) \cdot b(t, y) \mu(dy) \\ &\quad + \frac{1}{2} \text{Tr} [((\sigma\sigma^*)(t, x)) D_x^2 \varphi(t, x, \mu)] + D_x \varphi(t, x, \mu) \cdot b(t, x). \end{aligned}$$

Inspired from [13], we define a viscosity solution of (147) as follows.

**Definition 35.** We say that a continuous function  $u : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  is a viscosity solution of (147) if

- (i) The function  $u$  is jointly continuous and bounded;
- (ii) for any  $(t, x, \mu)$  in  $[0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n)$ , for any test functions (see definition 11.18 of [13] for the class of test functions)  $\varphi : (0, T) \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \rightarrow \mathbb{R}$  such that  $u - \varphi$  have a global minimum (resp. max) in  $(t, x, \mu)$  we have

$$\left\{ \begin{array}{l} \min \left\{ \{(\partial_t + \mathcal{L})\varphi(t, x, \mu) + f(t, x, \varphi(t, x, \mu))\}; \right. \\ \left. H(u(t, \cdot, \mu) \# \mu) \right\} \leq 0, \quad (\text{resp. } \geq 0), \quad [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n); \end{array} \right.$$

(iii)  $u(T, x, \mu) = \phi(x, \mu)$  on  $\mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n)$ .

The result is the following :

**Theorem 36.** *Suppose that assumption (A) holds.*

- (1) *Then there exists a decoupling field  $u : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \ni (t, x, \mu) \mapsto u(t, x, \mu) \in \mathbb{R}$  such that for any  $t$  in  $[0, T]$ , any r.v.  $X_0$  having second order moment and independent of the Brownian motion, we have for all  $s$  in  $[t, T]$   $u(s, X_s^{t, X_0}, [X_s^{t, X_0}]) = Y_s^{t, X_0}$  a.s..*
- (2) *Suppose in addition that assumption (A') holds. Then, the mapping  $u : [0, T] \times \mathbb{R}^n \times \mathcal{P}^2(\mathbb{R}^n) \ni (t, x, \mu) \mapsto u(t, x, \mu) \in \mathbb{R}$  is a continuous function and solves, in the sense of Definition 35, the obstacle problem on Wasserstein space (147).*

*Remark 37.* In (146), the driver  $f$  does not depend on  $Z$ . This assumption again relies to the lack of comparison principle for FBSDE (146) (see [8]) which is required to tackle driver depending on the  $Z$  argument when investigating the existence of a viscosity solution connected with FBSDEs.

*Proof.* Let us first emphasize that strong uniqueness holds for (146) under our standing assumptions. Considering the equation with the reflection process  $K^{t, X_0}$  as an entry, one can deduce from classical FBSDEs results that for any given initial  $t$  in  $[0, T]$  there exists a measurable function  $v_{t, X_0}(t, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}$ , called a decoupling field, such that for any  $s$  in  $[t, T]$ ,  $v_{t, X_0}(s, X_s^{t, X_0}) = Y_s^{t, X_0}$  a.s. (see also chapter 4 of [12]). To make the connection between the classical decoupling field  $v$  with a decoupling field  $u : [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ , let us first introduce the following decoupled flow of (146): for any  $x \in \mathbb{R}^n$ , we set on  $[0, T]$ :

$$\left\{ \begin{array}{l} X_s^{t, x, X_0} = x + \int_t^s b(r, X_r^{t, x, X_0}) dr + \int_t^s \sigma(r, X_r^{t, x, X_0}) dB_r, \quad s \geq t, \text{ and } X_s^{t, x, X_0} = x, \quad s < t, \\ Y_s^{t, x, X_0} = \phi(X_T^{t, x, X_0}, [X_T^{t, X_0}]) + \int_s^T f(r, X_r^{t, x, X_0}, Y_r^{t, x, X_0}) dr - \int_s^T Z_r^{t, x, X_0} dB_r \\ \quad + \int_s^T D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) dK_s^{t, [X_0]}. \end{array} \right. \quad (148)$$

Note that this equation is not of a MR-FBSDE and not of McKean-Vlasov type as well. Indeed the coefficients do not depend on the law of the solution of the above system but on the law of the solution of (146). Under our current assumptions, since the processes  $(X^{X_0}, Y^{X_0}, K^{X_0})$  are given, it is not hard to see that this system is well posed. Moreover, from weak uniqueness, the solution  $(X^{t, x, X_0}, Y^{t, x, X_0}, Z^{t, x, X_0})$  only depends on  $X_0$  through its law. Denoting this law by  $\mu$ , it is hence possible to write the solution  $(X^{t, x, \mu}, Y^{t, x, \mu}, Z^{t, x, \mu})$  without specifying the choice of the lifted random variable  $X_0$  with law  $\mu$ . Defining the mapping  $u(t, x, \mu) := Y_t^{t, x, \mu} : \mathbb{R}^+ \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$  we can show, arguing as in [14] (see proof of Proposition 2.2), that for any  $[X_0] = \mu$ , for all  $s$  in  $[t, T]$  :  $v_{t, \mu}(s, x) = u(s, x, [X_s^{t, X_0}])$ ,  $u(s, X_s^{t, x, [X_0]}, [X_s^{t, X_0}]) = Y_s^{t, x, [X_0]}$  and  $u(s, X_s^{t, X_0}, [X_s^{t, X_0}]) = Y_s^{t, X_0}$  a.s..

We now have the following lemma whose proof is postponed at the end of the current section.

**Lemma 38.** *The decoupling field  $u$  defined above is a continuous function from  $[0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ .*

It thus remains to check that this function solves, in the viscosity sense, (147). Let  $\varphi$  be a test function and  $(t, x, \mu) \in [0, T] \times \mathbb{R}^n \times \mathcal{P}_2(\mathbb{R}^n)$  be such that  $u - \varphi$  has a global minimum at  $(t, x, \mu)$ . For any  $s > t$  we have from the flow property that

$$\mathbb{E}u(s, X_s^{t, x, \mu}, [X_s^{t, \mu}]) = u(t, x, \mu) - \mathbb{E} \int_t^s f(r, X_r^{t, x, \mu}, Y_r^{t, x, \mu}) dr - \mathbb{E} \int_t^s D_\mu H([Y_r^{t, \mu}]) (Y_r^{t, x, \mu}) dK_r^{t, \mu}, \quad (149)$$

and from Itô's formula (12) that:

$$\mathbb{E}\varphi(s, X_s^{t, x, \mu}, [X_s^{t, \mu}]) = \varphi(t, x, \mu) + \mathbb{E} \int_t^s (\partial_t + \mathcal{L})\varphi(r, X_r^{t, x, \mu}, [X_r^{t, \mu}]) dr. \quad (150)$$

Up to a translation of  $\varphi$  we can assume that  $u(t, x, \mu) = \varphi(t, x, \mu)$ . Since  $(t, x, \mu)$  is a global minimum we have

$$\mathbb{E}u(s, X_s^{t, x, \mu}, [X_s^{t, \mu}]) - \mathbb{E}\varphi(s, X_s^{t, x, \mu}, [X_s^{t, \mu}]) \geq 0, \quad s \in [t, T].$$

Hence we have:

$$\begin{aligned} \forall s \in [t, \tilde{s}], \quad & \mathbb{E} \int_t^s (\partial_t + \mathcal{L})\varphi(r, X_r^{t,x,\mu}, [X_r^{t,\mu}]) dr \\ & + \mathbb{E} \int_t^s f(r, X_r^{t,x,\mu}, Y_r^{t,x,\mu}) dr + \mathbb{E} \int_t^s D_\mu H([Y_r^{t,\mu}]) (Y_r^{t,x,\mu}) dK_r^{t,\mu} \leq 0. \end{aligned}$$

Assume now that  $\mu$  is such that  $H(u(t, \cdot, \mu)\sharp\mu) > 0$  hence, by continuity, there exists  $\tilde{s} > t$  such for  $s$  in  $[t, \tilde{s}]$  we have  $H([u(s, X_s^{t,\mu}, [X_s^{t,\mu}])]) = H([Y_s^{t,X_0}]) > 0$ ,  $s \in [t, \tilde{s}]$  so that  $dK^{t,\mu}([t, s]) = 0$ ,  $s \in [t, \tilde{s}]$ . Thus

$$\mathbb{E} \int_t^s (\partial_t + \mathcal{L})\varphi(r, X_r^{t,x,\mu}, [X_r^{t,\mu}]) dr + \mathbb{E} \int_t^s f(r, X_r^{t,x,\mu}, Y_r^{t,x,\mu}) dr \leq 0.$$

Dividing by  $(s - t)$  and then letting  $s \rightarrow t$  we deduce

$$(\partial_t + \mathcal{L})\varphi(t, x, \mu) + f(t, x, \varphi(t, x, \mu)) \leq 0, \quad (151)$$

and

$$\min \left\{ (\partial_t + \mathcal{L})\varphi(t, x, \mu) + f(t, x, \varphi(t, x, \mu)) ; H(u(t, \cdot, \mu)\sharp\mu) \right\} \leq 0.$$

If  $\mu$  is now such that  $H(u(t, \cdot, \mu)\sharp\mu) = 0$  then

$$\min \left\{ (\partial_t + \mathcal{L})\varphi(t, x, \mu) + f(t, x, \varphi(t, x, \mu)) ; H(u(t, \cdot, \mu)\sharp\mu) \right\} \leq 0.$$

□

*Proof of Lemma 38.* Let  $(t, x, \mu) := (t^0, x^0, \mu^0)$  and  $(t^m, x^m, \mu^m)_m$  tends to  $(t, x, \mu)$  as  $m \rightarrow 0$ . Let  $X_0^m$  and  $X_0$  be of law  $\mu^m$  and  $\mu$  respectively. Recall that

$$\begin{cases} X_s^{t^m, x^m, X_0^m} = x^m + \int_t^s b(r, X_r^{t^m, x^m, X_0^m}) dr + \int_t^s \sigma(r, X_r^{t^m, x^m, X_0^m}) dB_r, & s \geq t \text{ and} \\ X_s^{t^m, x^m, X_0^m} = x^m, & s < t, \\ Y_s^{t^m, x^m, X_0^m} = \phi(X_T^{t^m, x^m}, [X_T^{t^m, X_0^m}]) + \int_s^T f(r, X_r^{t^m, x^m}, Y_r^{t^m, x^m, X_0^m}) dr - \int_s^T Z_r^{t^m, x^m, X_0^m} dB_r \\ \quad + \int_s^T D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) dK_s^{t^m, [X_0^m]}, & s \in [0, T]. \end{cases} \quad (152)$$

It is straightforward to see that  $X^{t^m, x^m, X_0^m} \rightarrow X^{t, x, X_0}$  in  $\mathcal{S}^2$ . Denoting by  $(Y^{t^m, X_0^m}, Z^{t^m, X_0^m})$  the solution of (146) with initial condition  $(t^m, X_0^m)$ , it is also easily to see (see proof of the uniqueness part of Theorem 24) that  $Y^{t^m, X_0^m} \rightarrow Y^{t, X_0}$  in  $\mathcal{S}^2$  and that  $Z^{t^m, X_0^m} \rightarrow Z^{t, X_0}$  in  $M^2$ . The tricky part here consists in proving that the solution  $Y_t^{t^m, x^m, X_0^m}$  of (152) tends to the solution  $Y_t^{t, x, X_0}$  of (148). Recall indeed that when considering the decoupled flow (148) of (146) we are not considering solutions of MR-FBSDE since the process  $K$  is only an input in these equations so that the Skorokhod condition is not satisfied anymore.

To prove that  $Y_{t^m}^{t^m, x^m, X_0^m} \rightarrow Y_t^{t, x, X_0}$  we first notice that these quantities are deterministic so

$$\begin{aligned} Y_{t^m}^{t^m, x^m, X_0^m} - Y_t^{t, x, X_0} &= \mathbb{E} \left[ Y_{t^m}^{t^m, x^m, X_0^m} - Y_t^{t, x, X_0} \right] \\ &\leq \left| \mathbb{E} \left[ Y_{t^m}^{t^m, x^m, X_0^m} - Y_t^{t^m, x^m, X_0^m} \right] \right| + \left| \mathbb{E} \left[ Y_t^{t^m, x^m, X_0^m} - Y_t^{t, x, X_0} \right] \right|. \end{aligned} \quad (153)$$

We have

$$\begin{aligned} \left| \mathbb{E} \left[ Y_{t^m}^{t^m, x^m, X_0^m} - Y_t^{t^m, x^m, X_0^m} \right] \right| &\leq \left| \mathbb{E} \int_{t^m}^t \left| f(r, X_r^{t^m, x^m, X_0^m}, Y_r^{t^m, x^m, X_0^m}) \right| dr \right| \\ &\quad + \left| \mathbb{E} \int_{t^m}^t |D_\mu H([Y_r^{t^m, X_0^m}]) (Y_r^{t^m, x^m, X_0^m})| dK_r^m \right|. \end{aligned}$$

Note now that since  $f$  satisfies the linear growth assumption in **(A')** (so that  $\sup_{t \leq T} \mathbb{E}[|f(t, X_t, Y_t)|^2] < +\infty$ ) we can deduce from the proof of *Step 1* and Proposition 26 (see Remark 27) that the map  $[0, T] \ni r \mapsto K_r^m$  is Lipschitz continuous (uniformly in  $m$ ). Hence, there exists  $k^m$  satisfying  $\sup_{r \in [0, T]} k_r^m \leq c$  (with  $c := c((A')) > 0$  independent of  $m$ ) such that for all  $r$  in  $[0, T]$  we have  $dK_r^m = k_r^m dr$ . Using this property, together with assumptions in **(A)** and **(A')** it clear that the above contribution also tends to 0.

It then remains to show that the second term in the right hand side of (153) also tends to 0. This last fact is a little bit more involved. We start by setting for all  $s$  in  $[0, T]$ ,

$$L_s^m = \int_0^s D_\mu H([Y_r^{t^m}, X_0^m]) (Y_r^{t^m, x^m, X_0^m}) dK_r^{t^m, X_0^m},$$

and

$$L_s = \int_0^s D_\mu H([Y_r^t, X_0]) (Y_r^{t, x, X_0}) dK_r^{t, X_0}.$$

From the FBSDE (146), we have, for all  $s$  in  $[0, T]$ ,

$$L_s^m = -Y_s^{t^m, X_0^m} + Y_{t^m}^{t^m, X_0^m} + \int_{t^m}^s f(r, X_r^{t^m, X_0^m}, Y_r^{t^m, X_0^m}) dr - \int_{t^m}^s Z_r^{t^m, x^m, X_0^m} dB_r,$$

and

$$L_s = -Y_s^{t, X_0} + Y_t^{t, X_0} + \int_t^s f(r, X_r^t, X_0, Y_r^t, X_0) dr - \int_t^s Z_r^{t, X_0} dB_r,$$

from which we deduce that for every  $s$  in  $[0, T]$ :

$$\lim_{m \rightarrow +\infty} L_s^m = L_s. \quad (154)$$

Let us now rewrite the BSDE (152) in the following form:

$$\begin{aligned} Y_s^{t^m, x^m, X_0^m} &= \phi(X_T^{t^m, x^m}, [X_T^{t^m, X_0^m}]) + \int_s^T f(r, X_r^{t^m, x^m, X_0^m}, Y_r^{t^m, x^m, X_0^m}) dr - \int_s^T Z_r^{t^m, x^m, X_0^m} dB_r \\ &\quad + \int_s^T D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) \left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) \right)^{-1} dL_s^m. \end{aligned} \quad (155)$$

Setting

$$\begin{aligned} A_s^m &= \int_0^s D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t, x, X_0}) \left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) \right)^{-1} dL_s^m \\ &\quad - \int_0^s D_\mu H([Y_s^t, X_0]) (Y_s^{t, x, X_0}) \left( D_\mu H([Y_s^t, X_0]) (Y_s^{t, x, X_0}) \right)^{-1} dL_s, \end{aligned}$$

and

$$U^m = Y^{t^m, x^m, X_0^m} - Y^{t, x, X_0} + A^m, \quad V^m = Z^{t^m, x^m, X_0^m} - Z^{t, x, X_0},$$

we deduce that  $(U^m, V^m)$  satisfy the following BSDE,

$$U_s^m = \eta^m + \int_s^T (\alpha_r^m U_r^m dr + \gamma_r^m U_r^m dL_r^m) - \int_s^T V_r^m dB_r + \int_s^T dF_r^m, \quad s \in [0, T],$$

where:

$$F_r^m = \int_0^r ((-\alpha_v^m A_v^m + \beta_v^m) dv - \gamma_v^m A_v^m dL_v^m),$$



with

$$\begin{aligned}
\eta^m &= \phi(X_T^{t^m, x^m, X_0^m}, [X_T^{t^m, X_0^m}]) - \phi(X_T^{t, x, X_0}, [X_T^{t, X_0}]) \\
&\quad + \int_0^T D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t, x, X_0}) \left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) \right)^{-1} dL_s^m \\
&\quad - \int_0^T D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) \left( D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) \right)^{-1} dL_s, \\
\alpha^m &= \frac{f(r, X_r^{t^m, x^m, X_0^m}, Y_r^{t^m, x^m, X_0^m}) - f(r, X_r^{t, x, X_0}, Y_r^{t, x, X_0})}{Y_s^{t^m, x^m, X_0^m} - Y_s^{t, x, X_0}}, \\
\beta^m &= f(r, X_r^{t^m, x^m, X_0^m}, Y_r^{t, x, X_0}) - f(r, X_r^{t, x, X_0}, Y_r^{t, x, X_0}), \\
\gamma^m &= \frac{\left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) - D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) \right)}{Y_s^{t^m, x^m, X_0^m} - Y_s^{t, x, X_0}} \\
&\quad \times \left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) \right)^{-1},
\end{aligned}$$

using the convention  $0/0 = 0$ . The definition of  $L^m$  and the Lipschitz property of  $K^m$  yield

$$\gamma_r^m U_r^m dL_r^m = \frac{\left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) - D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) \right)}{Y_s^{t^m, x^m, X_0^m} - Y_s^{t, x, X_0}} U_r^m k_r^m dr, \quad (156)$$

with  $k_r^m \leq c$  where  $c > 0$  does not depend on  $m$ . Note also that  $\alpha^m$ ,  $\beta^m$  and  $\gamma^m$  are bounded uniformly in  $m$ . We then obtain for every  $1 < p < 2$ :

$$\mathbb{E}|U_s^m|^p = \mathbb{E}[|\mathbb{E}[U_s^m | \mathcal{F}_s]|^p] \leq C \left\{ \mathbb{E}|\eta^m|^p + \int_s^T \mathbb{E}|U_r^m|^p dr + \mathbb{E} \left[ \left( \int_0^T d|F^m|_r \right)^p \right] \right\}.$$

Hence, applying Gronwall's inequality, we obtain that there exists  $c := c(T) > 0$  such that

$$\mathbb{E}|U_s^m|^p \leq c (\mathbb{E}|\eta^m|^p) + \mathbb{E} \left[ \left( \int_0^T d|F^m|_r \right)^p \right].$$

Let us now prove that the right hand side of the above equation tends to 0.

Note that from (154) we obtain that  $\eta^m \rightarrow 0$  and from Lipschitz property of  $r \mapsto K_r^m$  we know that there exists  $C := C((\mathbf{A}'), T) > 0$  such that  $\mathbb{E} \left[ \int_0^T \left| D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t, x, X_0}) dK_r^m \right|^2 \right] \leq C$  so that  $\mathbb{E} [|\eta^m|^p] \rightarrow 0$ .

We consider now the term  $\mathbb{E}[|F^m|_T^p]$ . First note that from the Lipschitz property of  $r \mapsto K_r^m$  and the definition of  $\gamma^m$  and  $L^m$ :

$$\gamma_r^m A_r^m dL_r^m = \frac{\left( D_\mu H([Y_s^{t^m, X_0^m}]) (Y_s^{t^m, x^m, X_0^m}) - D_\mu H([Y_s^{t, X_0}]) (Y_s^{t, x, X_0}) \right)}{Y_s^{t^m, x^m, X_0^m} - Y_s^{t, x, X_0}} A_r^m k_r^m dr, \quad (157)$$

so that there exists  $C := C((\mathbf{A}), T) > 0$  (that may change from line to line) such that

$$\begin{aligned}
\mathbb{E}[|F^m|_T^p] &\leq C \left\{ \mathbb{E} \left[ \int_0^T (|A_r^m|^p + |\beta_r^m|^p) dr \right] + \mathbb{E} \left[ \int_0^T A_r^m dK_r^m \right]^p \right\} \\
&\leq C \mathbb{E} \left[ \int_0^T (|A_r^m|^p + |\beta_r^m|^p) dr \right].
\end{aligned}$$

It is clear that  $\mathbb{E}[|\beta_r^m|^p] \rightarrow 0$ . Let us now investigate the convergence of  $\mathbb{E}[|A_r^m|^p]$ . We have from the definition of  $A_s^m$  and  $L_s^m$  that

$$\begin{aligned} A_s^m &= \int_0^s D_\mu H([Y_s^{t^m}, X_0^m])(Y_s^{t,x}, X_0) dK_r^m - \int_0^s D_\mu H([Y_s^{t,x}, X_0])(Y_s^{t,x}, X_0) dK_r \\ &= \int_0^s \left( D_\mu H([Y_s^{t^m}, X_0^m])(Y_s^{t,x}, X_0) - D_\mu H([Y_s^{t,x}, X_0])(Y_s^{t,x}, X_0) \right) k_r^m dr \\ &\quad + \int_0^s D_\mu H([Y_s^{t,x}, X_0])(Y_s^{t,x}, X_0) (dK_r^m - dK_r). \end{aligned} \quad (158)$$

The first term in the right hand side tends to 0 by continuity and uniform boundedness of  $k_r^m$ . To prove the convergence to 0 of the second term in the right hand side it is sufficient to prove, thanks to an approximation argument, that this holds when the integrand is any step function. The problem hence reduces to prove the convergence of  $K_s^m$  to  $K_s$  for every  $s$  in  $[0, T]$ . Recall

$$K_s^m = \int_0^s [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} dL_s^m, \quad K_s = \int_0^s [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} dL_s.$$

We then have

$$\begin{aligned} K_s^m - K_s &= \int_0^s [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} - [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} dL_r^m \\ &\quad + \int_0^s [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} d(L_s^m - L_s). \end{aligned}$$

On the one hand we have  $r \mapsto [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1}$  is continuous and we also have that for every step function  $\varphi$ ,  $\int_0^s \varphi(s) d(L_s^m - L_s) \rightarrow 0$  so that

$$\int_0^s [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} d(L_s^m - L_s) \rightarrow 0.$$

On the other hand,

$$\begin{aligned} &\mathbb{E} \sup_{s \in [0, T]} \left| \int_0^s [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} - [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} dL_r^m \right| \\ &\leq \mathbb{E} \left[ \int_0^T \left| [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} - [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} \right| d|L_r^m| \right] \\ &\leq C_T \int_0^T \mathbb{E}^{1/2} \left[ \left| [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} - [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} \right|^2 \right] \\ &\quad \times \mathbb{E}^{1/2} \left[ \left| [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)] \right|^2 \right] dr \\ &\leq C_T \int_0^T \mathbb{E}^{1/2} \left[ \left| [D_\mu H([Y_r^{t^m}, X_0^m])(Y_r^{t^m, x^m}, X_0^m)]^{-1} - [D_\mu H([Y_r^{t,x}, X_0])(Y_r^{t,x}, X_0)]^{-1} \right|^2 \right] M dr, \end{aligned}$$

and it is clear that the last term in the above right hand side also tends to 0.

For any subsequence  $(m')$ , there exists  $(m'') \subset (m')$  such that for every  $s$  in  $[0, T]$ ,  $K_s^{m''} \rightarrow K_s$ , and then  $A_s^{m''} \rightarrow 0$ . Finally, we deduce from our assumptions and the Lipschitz regularity of  $K$  that the family  $(A_s^{m''})_{m'' \geq 0}$  is uniformly bounded in  $L^2([0, T] \times \Omega)$  so that it converges to 0 in  $L^p([0, T] \times \Omega)$  for any  $1 < p < 2$ . This means that  $(A_s^m)_{m \geq 0}$  converges to 0 in  $L^p([0, T] \times \Omega)$  for any  $1 < p < 2$ .

We obtain that for all  $s$  in  $[0, T]$ ,  $1 < p < 2$ :  $\mathbb{E}[|U_s^m|^p] \rightarrow 0$ , which means that in  $L^p$ , for every  $s$  in  $[0, T]$ ,

$$Y_s^{t^m, x^m, X_0^m} \rightarrow Y_s^{t,x, X_0}.$$

Using this in (153), this concludes the proof.  $\square$

## References.

- [1] L. Ambrosio, N. Gigli and G. Savaré *Gradient flows in metric spaces and in the space of probability measures*. Second edition. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.
- [2] D. D., Bainov and P. S., Simeonov, *Integral inequalities and applications*, **57**. Springer Science & Business Media, 2013.
- [3] Ph. Briand, P.-É. Chaudru de Raynal, A. Guillin and C. Labart, *Particles systems and numerical schemes for mean reflected stochastic differential equations*. arXiv:1612.06886, 2016.
- [4] Ph. Briand, R. Elie and Y. Hu, *BSDEs with mean reflexion*. Ann. Appl. Probab. **28** (2018), no. 1, 482–510.
- [5] Ph. Briand, A. Ghannoum and C. Labart, *Mean Reflected Stochastic Differential Equations with Jumps*. arXiv:1803.10165, 2018.
- [6] Ph. Briand, H. Hibon, *Particles Systems for Mean Reflected BSDEs*. arXiv:1712.07385, 2017.
- [7] B. Bouchard, B. Djehiche and I. Kharroubi, *Quenched mass transport of particles towards a target*. arXiv:1707.07869, 2017.
- [8] R. Buckdahn, J. Li and S. Peng, *Mean-field backward stochastic differential equations and related partial differential equations*. Stochastic Process. Appl. **119** (2009), no. 10, 3133–3154.
- [9] R. Buckdahn, J. Li, S. Peng and C. Rainer, *Mean-field stochastic differential equations and associated PDEs*. Ann. Probab. **45** (2017), no. 2, 824–878.
- [10] D. L. Burkholder, É. Pardoux and A. Sznitman, *École d'Été de Probabilités de Saint-Flour XIX—1989, 1–66*, Lecture Notes in Math. **1464**, Springer, Berlin, 1991.
- [11] P. Cardaliaguet, F. Delarue, J.-M. Lasry and P.-L. Lions, *The master equation and the convergence problem in mean field games*. Annals of Mathematics Studies, 2019.
- [12] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications. I. Mean-field FBSDEs, control, and games*. Probability Theory and Stochastic Modelling, **83**. Springer, Cham, 2018.
- [13] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications. II. Mean-field games with common noise and master equations*. Probability Theory and Stochastic Modelling, **83**. Springer, Cham, 2018.
- [14] J.-F. Chassagneux, D. Crisan and F. Delarue, *A Probabilistic approach to classical solutions of the master equation for large population equilibria*. arXiv:1411.3009, 2014.
- [15] M. G. Crandall, H. Ishii and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*. Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1.
- [16] I. Fabbri, F. Gozzi and A. Swiech, *Stochastic optimal control in infinite dimension. Dynamic programming and HJB equations*. Probability Theory and Stochastic Modelling, **82**. Springer, Cham, 2017.
- [17] N. Fournier and A. Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*. Probab. Theory Related Fields **162** (2015), no. 3-4, 707–738.
- [18] A. Gégout-Petit and E. Pardoux, *Équations différentielles stochastiques rétrogrades réfléchies dans un convexe*. Stochastics Stochastics Rep. **57**, no. 1-2, 111–128.
- [19] H. Hibon, Y. Hu, Y. Lin, P. Luo, F. Wang *Quadratic BSDEs with mean reflection*. Mathematical Control & Related Fields **8**, no. 3-4, 721–738, 2018.
- [20] J.-F. Jabir (2017). *Diffusion processes with weak constraint through penalization approximation*. arXiv preprint arXiv:1704.01505.
- [21] G. Jones, *Fundamental Inequalities for Discrete and Discontinuous Functional Equations*. J. Soc. Indust. Appl. Math **12** (1964), no. 1, 43–57.
- [22] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, M. C. Quenez, *Reflected solutions of backward SDE's, and related obstacle problems for PDE's*. Ann. Probab. **25** (1997), no. 2, 702–737.
- [23] V. N. Kolokoltsov (2010). *Nonlinear Markov processes and kinetic equations* (Vol. 182). Cambridge University Press.
- [24] P.-L. Lions, In Cours au Collège de France. www.college-de-france.fr (2008).
- [25] P.-L. Lions and A.-S. Sznitman, *Stochastic differential equations with reflecting boundary conditions*. Comm. Pure Appl. Math. **37** (1984), no. 4, 511–537.
- [26] P. L. Lions, J. L. Menaldi and A. S. Sznitman, *Construction de processus de diffusion réfléchis par pénalisation du domaine*. CR Acad. Sci. Paris Sér. I Math, 292(11) (1981), 559–562.
- [27] S. Mischler and C. Mouhot (2013). *Kac's program in kinetic theory*. Inventiones mathematicae, 193(1), 1-147.
- [28] H. Pham and X. Wei, *Bellman equation and viscosity solutions for mean-field stochastic control problem*. arXiv:1512.07866, 2015.
- [29] S. T. Rachev and L. Rüschendorf, *Mass transportation problems. Vol. I. Theory*. Probability and its Applications (New York). Springer-Verlag, New York, 1998.
- [30] C. Wu, J. Zhang, *Viscosity solutions to parabolic master equations and McKean-Vlasov SDEs with closed-loop controls*. arXiv:1805.02639, 2018.

UNIV. GRENOBLE ALPES, UNIV. SAVOIE MONT BLANC, CNRS, LAMA, 73000 CHAMBÉRY, FRANCE  
*E-mail address:* [philippe.briand@univ-smb.fr](mailto:philippe.briand@univ-smb.fr)

UNIVERSITÉ PARIS-DAUPHINE, PSL RESEARCH UNIVERSITY, CNRS, CEREMADE, 75016 PARIS, FRANCE.  
*E-mail address:* [cardaliagnet@ceremade.dauphine.fr](mailto:cardaliagnet@ceremade.dauphine.fr)

UNIV. GRENOBLE ALPES, UNIV. SAVOIE MONT BLANC, CNRS, LAMA, 73000 CHAMBÉRY, FRANCE  
*E-mail address:* [pe.deraynal@univ-smb.fr](mailto:pe.deraynal@univ-smb.fr)

UNIV RENNES, CNRS, IRMAR-UMR6625, F-35000, RENNES, FRANCE.  
*E-mail address:* [ying.hu@univ-rennes1.fr](mailto:ying.hu@univ-rennes1.fr)