Approximation schemes for McKean-Vlasov and Boltzmann type equations (error analysis in total variation distance)

Yifeng Qin

Approximation schemes for McKean-Vlasov and Boltzmann type equations (error analysis in total variation distance)

Yifeng Qin*

*Université Gustave Eiffel, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. Email address: gzevonqin@gmail.com

Jan. 2023

Abstract We deal with Mckean-Vlasov and Boltzmann type jump equations. This means that the coefficients of the stochastic equation depend on the law of the solution, and the equation is driven by a Poisson point measure with intensity measure which depends on the law of the solution as well. In [3], Alfonsi and Bally have proved that under some suitable conditions, the solution X_t of such equation exists and is unique. One also proves that X_t is the probabilistic interpretation of an analytical weak equation. Moreover, the Euler scheme $X_t^{\mathcal{P}}$ of this equation converges to X_t in Wasserstein distance. In this paper, under more restricted assumptions, we show that the Euler scheme $X_t^{\mathcal{P}}$ converges to X_t in total variation distance and X_t has a smooth density (which is a function solution of the analytical weak equation). On the other hand, in view of simulation, we use a truncated Euler scheme $X_t^{\mathcal{P},M}$ which has a finite numbers of jumps in any compact interval. We prove that $X_t^{\mathcal{P},M}$ also converges to X_t in total variation distance. Finally, we give an algorithm based on a particle system associated to $X_t^{\mathcal{P},M}$ in order to approximate the density of the law of X_t. Complete estimates of the error are obtained.

Key words: Mckean-Vlasov equation, Boltzmann equation, Malliavin calculus, Total variation distance, Wasserstein distance, Particle system

Contents

1 Introduction

In this paper, we consider a $d-$dimensional Mckean-Vlasov and Boltzmann type jump equation as follows.

$$X_t = X_0 + \int_0^t b(r, X_r, \rho_r)dr + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(r, v, z, X_{r-}, \rho_{r-}) N_{\rho_{r-}}(dv, dz, dr), \quad (1)$$

where $\rho_t(dv) = \mathbb{P}(X_t \in dv)$ is the law of X_t, $t \in [0, T]$, N_{ρ_t} is a Poisson point measure on the state space $\mathbb{R}^d \times \mathbb{R}^d$ with intensity measure $\rho_t(dv)\mu(dz)dr$, μ is a positive σ-finite measure on $\mathbb{R}^d$, X_0 is the initial random variable independent of the Poisson point measure N_{ρ_t}, and b, c are functions which verify some regularity and ellipticity conditions (see **Hypotheses 2.1∼2.4** in Section 2.2 for precise statements). In particular, we assume that for every multi-indices β_1, β_2, there exists a non-negative function $\bar{c} : \mathbb{R}^d \to \mathbb{R}_+$ such that

$$|c(r, v, z, x, \rho)| + |\partial_z^{\beta_2} \partial_x^{\beta_1} c(r, v, z, x, \rho)| \leq \bar{c}(z),$$

with $\int_{\mathbb{R}^d} |\bar{c}(z)|^p \mu(dz) < \infty, \ \forall p \geq 1$. We also assume that there exists a non-negative function $\underline{c} : \mathbb{R}^d \to \mathbb{R}_+$ such that for every $\zeta \in \mathbb{R}^d$,

$$\sum_{j=1}^d \langle \partial_{z_j} c(r, v, z, x, \rho), \zeta \rangle^2 \geq \underline{c}(z)|\zeta|^2.$$

We remark that we use the notations from [25] and we refer to [6], [16], [25], [34], [35], [43] and [44] for the basic theory of the classical jump equations. We stress that our equation is a more general kind of jump equation (than the classical one) in the following sense. The coefficients b and c depend on the law of the solution, so our equation is of Mckean-Vlasov type. One can see for example [23] for a mathematical approach to this kind of equation and see [5], [12], [22], [24], [39], [40] and [47] for the approximation schemes of a Mckean-Vlasov equation. Moreover, the intensity of the Poisson point measure N_{ρ_t} depends on the law of the solution as well, so our equation is also of Boltzmann type. The probabilistic approach to the Boltzmann equation is initiated by Tanaka in [48], [49], and followed by many others in [9], [17], [18], [19], [36], [38] and [46] for example. One can also see [2] and [50] for the analytical Boltzmann equation and [15] for the physical background. Recently, there is also some work on inhomogeneous Boltzmann equations (see for instance [1], [20] and [21]). We have to mention however that our equation (1) does not cover the general physical Boltzmann equation for the following reason. In that equation, the intensity of the jumps $\mu(dz)$ is replaced by $\gamma(r, v, z, x, \rho)\mu(dz)$ which depends on the position $x = X_{r-}$ of the solution of the equation. At least at this time, we are not able to include this case in our study. The simplified model that we treat in our paper corresponds to Maxwell molecules (see [18] for example).

Now we construct the Euler scheme. For any partition $\mathcal{P} = \{0 = r_0 < r_1 < \cdots < r_{n-1} < r_n = T\}$ of the time interval $[0, T]$, we define $\tau(r) = r_k$ when $r \in [r_k, r_{k+1})$, and we consider the equation

$$\begin{aligned} X_t^{\mathcal{P}} &= X_0 + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P}}, \rho_{\tau(r)}^{\mathcal{P}})dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P}}, \rho_{\tau(r)-}^{\mathcal{P}}) N_{\rho_{\tau(r)-}^{\mathcal{P}}}(dv, dz, dr), \end{aligned} \quad (2)$$

where $\rho_t^{\mathcal{P}}$ is the law of $X_t^{\mathcal{P}}$, and $N_{\rho_t^{\mathcal{P}}}(dv, dz, dr)$ is a Poisson point measure with intensity $\rho_t^{\mathcal{P}}(dv)\mu(dz)dr$, independent of X_0. We remark that for the classical jump equations (the coefficients and the Poisson point measures do not depend on the law of the solution), there is a huge amount of work on the convergence of the Euler scheme. One can see for example [4], [27], [28], [29], [30], [31], [42] and the references therein. For the equation (1), [3] has proved recently that under some regularity conditions on the coefficients b and c, the solution of the equation (1) exists and is unique, and X_t is the probabilistic interpretation of

the following analytical weak equation.

$$\forall \phi \in C_b^1(\mathbb{R}^d)(\textit{the space of differentiable and bounded functions with bounded derivatives}),$$
$$\int_{\mathbb{R}^d} \phi(x)\rho_t(dx) = \int_{\mathbb{R}^d} \phi(x)\rho_0(dx) + \int_0^t \int_{\mathbb{R}^d} \langle b(r,x,\rho_r), \nabla\phi(x)\rangle \rho_r(dx)dr$$
$$+ \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} \rho_r(dx)\rho_r(dv) \int_{\mathbb{R}^d} (\phi(x + c(r,v,z,x,\rho_r)) - \phi(x))\mu(dz)dr. \tag{3}$$

Moreover, [3] has proved that the Euler scheme $X_t^{\mathcal{P}}$ converges to X_t in Wasserstein distance (of order 1) W_1. In our paper, under supplementary hypotheses, we prove a stronger result. We prove (see **Theorem 2.1**) that the Euler scheme $X_t^{\mathcal{P}}$ converges to X_t in total variation distance: for any $\varepsilon > 0$, there exists a constant C such that

$$d_{TV}(X_t^{\mathcal{P}}, X_t) \leq C|\mathcal{P}|^{1-\varepsilon} \to 0, \tag{4}$$

as $|\mathcal{P}| \to 0$, with $|\mathcal{P}| := \max_{k\in\{0,\cdots,n-1\}} (r_{k+1} - r_k)$. We also show that the law of X_t has a smooth density $p_t(x)$, which is a function solution of the analytical weak equation (3).

Since we have infinite numbers of jumps (due to **Hypothesis 2.4** in Section 2.2), we have $\mu(\mathbb{R}^d) = \infty$. In view of simulation, we need to work with a truncated Poisson point measure, which has a finite number of jumps in any compact time interval. For $M \in \mathbb{N}$, we denote $B_M = \{z \in \mathbb{R}^d : |z| \leq M\}$, $c_M(r,v,z,x,\rho) := c(r,v,z,x,\rho)\mathbb{1}_{B_M}(z)$ and $a_T^M := \sqrt{T\int_{\{|z|>M\}} \underline{c}(z)\mu(dz)}$. Now we cancel the jumps of size $|z| > M$ and we replace them by a Gaussian random variable.

$$\begin{aligned} X_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})dr \\ &+ \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) N_{\rho_{\tau(r)-}^{\mathcal{P},M}}(dv,dz,dr), \end{aligned} \tag{5}$$

where $\rho_t^{\mathcal{P},M}$ is the law of $X_t^{\mathcal{P},M}$, $N_{\rho_t^{\mathcal{P},M}}(dv,dz,dr)$ is a Poisson point measure independent of X_0 with intensity $\rho_t^{\mathcal{P},M}(dv)\mu(dz)dr$, Δ is a $d-$dimensional standard Gaussian random variable independent of X_0 and of $N_{\rho_t^{\mathcal{P},M}}$. We prove (see **Theorem 2.2**) that $X_t^{\mathcal{P},M}$ converges to X_t in total variation distance: for any $\varepsilon > 0$, there exists a constant C such that

$$d_{TV}(X_t^{\mathcal{P},M}, X_t) \leq C(\sqrt{\varepsilon_M} + |\mathcal{P}|)^{1-\varepsilon} \to 0, \tag{6}$$

as $|\mathcal{P}| \to 0$ and $M \to \infty$, with $\varepsilon_M := \int_{\{|z|>M\}} |\bar{c}(z)|^2\mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^2$. Moreover, the law of $X_t^{\mathcal{P},M}$ has a smooth density.

In order to construct an approximation scheme which is appropriate for simulation, we need to compute $\rho_t^{\mathcal{P},M}$ as well, so we use the following particle system. We take an initial vector $(X_0^1, \cdots, X_0^N)$ with components which are independent and identically distributed with common law ρ_0 (which is the law of X_0), and $(\Delta^1, \cdots, \Delta^N)$ which is a $N \times d-$dimensional standard Gaussian random variable independent of $(X_0^1, \cdots, X_0^N)$. Then we construct the particle system $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M} = (X_t^{\mathcal{P},M,1}, \cdots, X_t^{\mathcal{P},M,N})$:

$$\begin{aligned} X_t^{\mathcal{P},M,i} &= X_0^i + a_T^M \Delta^i + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M,i}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}_{\tau(r)}^{\mathcal{P},M}))dr \\ &+ \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M,i}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}_{\tau(r)-}^{\mathcal{P},M}))N^i_{\widehat{\rho}(\overrightarrow{\boldsymbol{X}}_{\tau(r)-}^{\mathcal{P},M})}(dv,dz,dr),\ i = 1,\cdots,N, \end{aligned} \tag{7}$$

where

$$\widehat{\rho}(\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M})(dv) = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{\mathcal{P},M,i}}(dv)$$

is the empirical measure of $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M}$ (with $\delta_x(dv)$ the Dirac measure), $N^i_{\widehat{\rho}(\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M})}(dv,dz,dr),\ i=1,\cdots,N$ are Poisson point measures that are independent each other conditionally to $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M}$ and independent of $(X_0^1,\cdots,X_0^N,\Delta^1,\cdots,\Delta^N)$ with intensity $\widehat{\rho}(\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M})(dv)\mu(dz)dr$. It is clear that $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M}$ may be simulated in an explicit way (see (31)).

We denote

$$V_N := \mathbb{1}_{d=1}N^{-\frac{1}{2}} + \mathbb{1}_{d=2}N^{-\frac{1}{2}}\log(1+N) + \mathbb{1}_{d\geq 3}N^{-\frac{1}{d}},$$

and we consider the following $d-$dimensional regularization kernels

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}}e^{-\frac{|x|^2}{2}}, \quad \varphi_\delta(x) = \frac{1}{\delta^d}\varphi(\frac{x}{\delta}), \quad 0<\delta\leq 1.$$

We have proved in **Theorem 2.1** that the law of X_t has a density function $p_t(x)$. Now we obtain in **Theorem 2.3** the following results concerning the approximation of the density $p_t(x)$. We take

$$\delta = (|\mathcal{P}|+\sqrt{\varepsilon_M})^{\frac{1}{d+3}}, \quad \text{and take } N \text{ such that} \quad V_N \leq |\mathcal{P}|+\sqrt{\varepsilon_M}.$$

Then we have

$$p_t(x) = \frac{1}{N}\sum_{i=1}^N \mathbb{E}\varphi_\delta(X_t^{\mathcal{P},M,i}-x) + \mathcal{O}((|\mathcal{P}|+\sqrt{\varepsilon_M})^{\frac{2}{d+3}}), \tag{8}$$

where $\mathcal{O}(\bullet)$ is the big $\mathcal{O}$ notation (i.e. for a strictly positive function g defined on $\mathbb{R}_+$, $\exists C>0$, s.t. $|\mathcal{O}(g(y))|\leq Cg(y)$). If we take

$$\delta = (|\mathcal{P}|+\sqrt{\varepsilon_M})^{\frac{1}{d+5}}, \quad \text{and take } N \text{ such that} \quad V_N \leq |\mathcal{P}|+\sqrt{\varepsilon_M},$$

then we get moreover by Romberg method that

$$p_t(x) = \frac{2}{N}\sum_{i=1}^N \mathbb{E}\varphi_{\delta/\sqrt{2}}(X_t^{\mathcal{P},M,i}-x) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}\varphi_\delta(X_t^{\mathcal{P},M,i}-x) + \mathcal{O}((|\mathcal{P}|+\sqrt{\varepsilon_M})^{\frac{4}{d+5}}). \tag{9}$$

It is clear that the approximation scheme based on Romberg method gives a better accuracy: we have the power $\frac{4}{d+5} > \frac{2}{d+3}$. So we are able to simulate the density function of X_t in an explicit way, with error $\mathcal{O}((|\mathcal{P}|+\sqrt{\varepsilon_M})^{\frac{4}{d+5}})$. We notice however, that the speed of convergence of the error depends on the dimension d, so it converges slowly when d is large. In **Theorem 2.4**, we prove an alternative approximation result. We give up the approximation of the density, and we focus on the approximation in total variation distance. We take supplementally $\widetilde{\Delta}$ a $d-$dimensional standard Gaussian random variable independent of $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M}$. For any $\varepsilon>0$, we take

$$\delta = (|\mathcal{P}|+\varepsilon_M)^{\frac{1}{2}(1-\varepsilon')} \quad \text{and take } N \text{ such that} \quad V_N \leq (|\mathcal{P}|+\varepsilon_M)^{\frac{d+3}{2}(1-\varepsilon'')},$$

with $\varepsilon' = \frac{\varepsilon}{2-\varepsilon}$ and $\varepsilon'' = \frac{(d+5)\varepsilon-2\varepsilon^2}{(d+3)(2-\varepsilon)}$. For every measurable and bounded function f, we prove that

$$\int_{\mathbb{R}^d} f(x)p_t(x)dx \quad = \quad \frac{1}{N}\sum_{i=1}^N \mathbb{E}f(X_t^{\mathcal{P},M,i}+\delta\widetilde{\Delta}) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}|+\sqrt{\varepsilon_M})^{1-\varepsilon}). \tag{10}$$

We notice that the speed of convergence in (10) no longer depends on the dimension d, so it still behaves well for large dimension. We also stress that the speed of convergence in (10) is the same as in (6) for the truncated Euler scheme. Moreover, for any $\varepsilon>0$, we take

$$\delta = (|\mathcal{P}|+\varepsilon_M)^{\frac{1}{4}(1-\varepsilon_\prime)} \quad \text{and take } N \text{ such that} \quad V_N \leq (|\mathcal{P}|+\varepsilon_M)^{\frac{d+5}{4}(1-\varepsilon_{\prime\prime})},$$

with $\varepsilon_{\prime} = \frac{\varepsilon^2}{2-\varepsilon}$ and $\varepsilon_{\prime\prime} = \frac{8\varepsilon+(d-3)\varepsilon^2}{(d+5)(2-\varepsilon)}$. Then for every measurable and bounded function f, we get by Romberg method that

$$\int_{\mathbb{R}^d} f(x)\rho_t(dx) \quad = \quad \frac{2}{N}\sum_{i=1}^{N}\mathbb{E}f(X_t^{\mathcal{P},M,i}+\frac{\delta}{\sqrt{2}}\widetilde{\Delta}) - \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}f(X_t^{\mathcal{P},M,i}++\delta\widetilde{\Delta}) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}|+\sqrt{\varepsilon_M})^{1-\varepsilon}). \quad (11)$$

We remark that (11) is even a better simulation scheme than (10) in the sense that the numbers of particles N is smaller than the one in (10) and δ is larger than the one in (10).

We give now a general view on the strategy used in the paper. Notice that the Poisson process which appears in the equation (1) has intensity $\mu(dz)$ which is an infinite measure. As we mentioned before, it is convenient, both from the point of view of Malliavin calculus and for simulation, to introduce an intermediary equation driven by a Poison point measure with intensity $\mathbb{1}_{\{|z|\leq M\}}\mu(dz)$ which is a finite measure. We denote by X_t^M the solution of this equation (which is a truncated version of (1), see (38) for precise expression). Since X_t^M depends only on a finite number of jumps in any compact time interval, this will be a "simple functional" in the Malliavin calculus with respect to the amplitudes of the jumps. We also replace the jumps larger then M (which have been canceled) by a Gaussian noise - this is necessary in order to obtain the non degeneracy for X_t^M. Moreover, in order to be able to establish integration by parts formulas, we assume (see **Hypothesis 2.4** b)) that the measure μ is absolutely continuous with respect to the Lebesgue measure: $\mu(dz) = h(z)dz$, where h is infinitely differentiable and $\ln h$ has bounded derivatives of any order. Using the convergence $X_t^M \to X_t$, we are able to prove that X_t is smooth in the sense of Malliavin calculus for jump processes. We use this calculus in order to prove that the law of X_t is absolutely continuous with respect to the Lebesgue measure, with smooth density $p_t(dx)$.

Moreover, we construct an explicit algorithm which allows us to use Monte Carlo simulation in order to approximate X_t and p_t. To do it, we consider the Euler scheme $X_t^{\mathcal{P}}$ and the truncated Euler scheme $X_t^{\mathcal{P},M}$ (see (2) and (5)). Now we focus on three equations with solutions X_t, X_t^M and $X_t^{\mathcal{P},M}$. There is a supplementary difficulty which appears here: the Poisson point measures which govern these equations have an intensity which depends on the law of the solution of each of these equations. It is convenient to use similar equations driven by the same Poisson point measure. This is obtained by a coupling procedure: we construct x_t, x_t^M and $x_t^{\mathcal{P},M}$ which have the same law as X_t, X_t^M and $X_t^{\mathcal{P},M}$ but are defined on the same probability space and verify equations driven by the same Poisson point measure (this is done in Section 2.7). This allows us to compare them by using an L^1 calculus. This is why all our computations will concern these last equations.

In [3], one obtains estimates of the Wasserstein distance between these processes. In order to estimate the total variation distance between them, we will use Malliavin integration by parts techniques (which are presented in Section 3) together with some results from [7] which allows us to pass from estimates in Wassestein distance to estimates in total variation distance. Consequently a large part of the technical effort in the paper will concern estimates of the Malliavin-Sobolev norms of x_t, x_t^M and $x_t^{\mathcal{P},M}$ as well as the proof of the non-degeneracy of these random variables (see Section 4).

Our paper is organized in the following way. In Section 2, we state our problems and give the hypotheses. We define the main equation X_t, the Euler scheme $X_t^{\mathcal{P}}$, the truncated Euler scheme $X_t^{\mathcal{P},M}$ and the particle system $X_t^{\mathcal{P},M,i}$, $i = 1, \cdots, N$. Then we state our main results: **Theorem 2.1, 2.2** (see (4) and (6)) and **Theorem 2.3, 2.4** (see (8), (9), (10) and (11)). We also give some typical examples to apply our main results. At the end of this section, we make a coupling argument to construct x_t^M, $x_t^{\mathcal{P},M}$ and x_t. In Section 3, we give an abstract integration by parts framework (of Malliavin type) and then apply these abstract results to the solutions of our equations. There are two types of results that we have to prove in order to make the integration by parts machinery works. First, we prove that the Malliavin-Sobolev norms of x_t^M, $x_t^{\mathcal{P},M}$ and x_t are bounded, uniformly with respect to $\mathcal{P}$ and M (see **Lemma 3.7**). Moreover we have to check the non-degeneracy condition for the Malliavin covariance matrix. This is done in **Lemma 3.8**. Both these two lemmas are rather technical so we leave the proofs for Section 4. Once these lemmas are proved, **Proposition 3.6.1** allows us to conclude that $X_t^{\mathcal{P},M} \to X_t$ in total variation distance. We also prove that the Euler scheme $X_t^{\mathcal{P}} \to X_t$ in total variation distance in a similar way. Furthermore, we obtain an algorithm based on the particle system $X_t^{\mathcal{P},M,i}$, $i = 1, \cdots, N$ in order to compute the density function $p_t(x)$ of the law of X_t, and we estimate the error.

2 Main results

2.1 Basic notations and the main equation

We give a time horizon $T > 0$ and let $0 < t \leq T$. To begin, we introduce some notations which will be used through our paper. For a multi-index β, we denote $|\beta|$ to be the length of β. We denote $C_b^l(\mathbb{R}^d)$ the space of $l-$times differential and bounded functions on $\mathbb{R}^d$ with bounded derivatives up to order l, and $\|f\|_{l,\infty} := \sum_{|\beta|\leq l} \left\|\partial^\beta f\right\|_\infty$ for a function $f \in C_b^l(\mathbb{R}^d)$. We also denote $\mathcal{P}_l(\mathbb{R}^d)$ the space of all probability measures on $\mathbb{R}^d$ with finite $l-$moment. For $\rho_1, \rho_2 \in \mathcal{P}_1(\mathbb{R}^d)$, we define the Wasserstein distance W_1 by

$$W_1(\rho_1, \rho_2) = \sup_{Lip(f)\leq 1} \big| \int_{\mathbb{R}^d} f(x)\rho_1(dx) - \int_{\mathbb{R}^d} f(x)\rho_2(dx)\big|, \tag{12}$$

with $Lip(f) := \sup_{x\neq y} \frac{|f(x)-f(y)|}{|x-y|}$ the Lipschitz constant of f, and we define the total variation distance d_{TV} by

$$d_{TV}(\rho_1, \rho_2) = \sup_{\|f\|_\infty \leq 1} \big| \int_{\mathbb{R}^d} f(x)\rho_1(dx) - \int_{\mathbb{R}^d} f(x)\rho_2(dx)\big|. \tag{13}$$

For $F, G \in L^1(\Omega)$, we also denote $W_1(F, G) = W_1(\mathcal{L}(F), \mathcal{L}(G))$ and $d_{TV}(F, G) = d_{TV}(\mathcal{L}(F), \mathcal{L}(G))$, with $\mathcal{L}(F)$(respectively $\mathcal{L}(G)$) the law of the random variable F(respectively G). In addition, along the paper, C will be a constant which may change from a line to another. It may depend on some parameters and sometimes the dependence is precised in the notation (ex. C_l is a constant depending on l).

In this paper, we consider the $d-$dimensional stochastic differential equation with jumps

$$X_t \quad = \quad X_0 + \int_0^t b(r, X_r, \rho_r)dr + \int_0^t \int_{\mathbb{R}^d\times\mathbb{R}^d} c(r, v, z, X_{r-}, \rho_{r-})N_{\rho_{r-}}(dv, dz, dr), \tag{14}$$

where $\rho_t(dv) = \mathbb{P}(X_t \in dv)$ is the law of X_t, N_{ρ_t} is a Poisson point measure on the state space $\mathbb{R}^d \times \mathbb{R}^d$ with intensity measure $\widehat{N}_{\rho_t}(dv, dz, dr) = \rho_t(dv)\mu(dz)dr$, X_0 is the initial random variable with law ρ_0 independent of the Poisson point measure N_{ρ_t}, μ is a positive σ-finite measure on $\mathbb{R}^d$, and $b : [0, T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$, $c : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$.

Remark. We remark that we will assume in the following that $\int_{\mathbb{R}^d} \sup_{r\in[0,T]} \sup_{v,x\in\mathbb{R}^d} \sup_{\rho\in\mathcal{P}_1(\mathbb{R}^d)} |c(r, v, z, x, \rho)|\mu(dz) < \infty$, so we are in the finite variation case. The integral with respect to the Poisson point measure is not compensated.

2.2 Hypotheses

Here we give our hypotheses.

Hypothesis 2.1 (Regularity) We assume that the function $x \mapsto b(r, x, \rho)$ is infinitely differentiable with bounded derivatives of any orders, and that $\rho_0 \in \bigcap_{p=1}^\infty \mathcal{P}_p(\mathbb{R}^d)$. We also assume that the function $(z, x) \mapsto c(r, v, z, x, \rho)$ is infinitely differentiable and for every multi-indices β_1, β_2, there exists a non-negative function $\bar{c} : \mathbb{R}^d \to \mathbb{R}_+$ depending on β_1, β_2 such that we have

$$\sup_{r\in[0,T]} \sup_{v,x\in\mathbb{R}^d} \sup_{\rho\in\mathcal{P}_1(\mathbb{R}^d)} (|c(r, v, z, x, \rho)| + |\partial_z^{\beta_2}\partial_x^{\beta_1} c(r, v, z, x, \rho)|) \leq \bar{c}(z), \quad \forall z \in \mathbb{R}^d,$$

$$with \quad \int_{\mathbb{R}^d} |\bar{c}(z)|^p \mu(dz) := \bar{c}_p < \infty, \quad \forall p \geq 1. \tag{15}$$

Moreover, there exists a constant $L_b > 0$ such that for any $r_1, r_2 \in [0, T], v_1, v_2, x \in \mathbb{R}^d, z \in \mathbb{R}^d, \rho_1, \rho_2 \in \mathcal{P}_1(\mathbb{R}^d)$,

$$|b(r_1, x, \rho_1) - b(r_2, x, \rho_2)| \leq L_b(|r_1 - r_2| + W_1(\rho_1, \rho_2)),$$

and $|c(r_1,v_1,z,x,\rho_1)-c(r_2,v_2,z,x,\rho_2)|$
$+|\nabla_z c(r_1,v_1,z,x,\rho_1)-\nabla_z c(r_2,v_2,z,x,\rho_2)|+|\nabla_x c(r_1,v_1,z,x,\rho_1)-\nabla_x c(r_2,v_2,z,x,\rho_2)|$
$\leq \bar{c}(z)(|r_1-r_2|+|v_1-v_2|+W_1(\rho_1,\rho_2))$.

Remark. We will use several times the following consequence of (15) and of Burkholder inequality (see for example the Theorem 2.11 in [34], see also in [35]): Let $\Phi(r,v,z,\omega,\rho):[0,T]\times\mathbb{R}^d\times\mathbb{R}^d\times\Omega\times\mathcal{P}_1(\mathbb{R}^d)\to\mathbb{R}_+$ and $\varphi(r,v,\omega,\rho):[0,T]\times\mathbb{R}^d\times\Omega\times\mathcal{P}_1(\mathbb{R}^d)\to\mathbb{R}_+$ be two functions such that

$$|\Phi(r,v,z,\omega,\rho)|\leq|\bar{c}(z)||\varphi(r,v,\omega,\rho)|.$$

Then for any $p\geq 2,\rho\in\mathcal{P}_1(\mathbb{R}^d)$,

$$\mathbb{E}\Big|\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\Phi(r,v,z,\omega,\rho)N_\rho(dv,dz,dr)\Big|^p\leq C\mathbb{E}\int_0^t\int_{\mathbb{R}^d}|\varphi(r,v,\omega,\rho)|^p\rho(dv)dr, \tag{16}$$

where C is a constant depending on p, $\bar{c}_1$, $\bar{c}_2$, $\bar{c}_p$ and T.

Proof. By compensating N_ρ and using Burkholder inequality and (15), we have

$$\begin{aligned}
&\mathbb{E}|\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}\Phi(r,v,z,\omega,\rho)N_\rho(dv,dz,dr)|^p\\
&\leq C[\mathbb{E}(\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}|\Phi(r,v,z,\omega,\rho)|^2\rho(dv)\mu(dz)dr)^{\frac{p}{2}}+\mathbb{E}\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}|\Phi(r,v,z,\omega,\rho)|^p\rho(dv)\mu(dz)dr\\
&+\mathbb{E}|\int_0^t\int_{\mathbb{R}^d\times\mathbb{R}^d}|\Phi(r,v,z,\omega,\rho)|\rho(dv)\mu(dz)dr|^p] \qquad (17)\\
&\leq C\mathbb{E}\int_0^t\int_{\mathbb{R}^d}|\varphi(r,v,\omega,\rho)|^p\rho(dv)dr.
\end{aligned}$$

□

For the sake of simplicity of notations, in the following, for a constant C, we do not precise the dependence on the regularity constants of the function b and c (such as $\|\nabla_x b\|_\infty$, L_b and $\bar{c}_p$).

Hypothesis 2.2 We assume that there exists a non-negative function $\breve{c}:\mathbb{R}^d\to\mathbb{R}_+$ such that $\int_{\mathbb{R}^d}|\breve{c}(z)|^p\mu(dz):=\breve{c}_p<\infty,\ \forall p\geq 1$, and

$$\big\|\nabla_x c(r,v,z,x,\rho)(I_d+\nabla_x c(r,v,z,x,\rho))^{-1}\big\|\leq\breve{c}(z),\quad\forall r\in[0,T],v,x\in\mathbb{R}^d,z\in\mathbb{R}^d,\rho\in\mathcal{P}_1(\mathbb{R}^d),$$

with I_d the $d-$dimensional identity matrix. To avoid overburdening notation, since both hypotheses 2.1 and 2.2 apply, we take $\breve{c}(z)=\bar{c}(z)$ and $\breve{c}_p=\bar{c}_p$.

Remark. We need this hypothesis to prove the regularity of the inverse tangent flow (see Section 4.2 (130)).

Hypothesis 2.3 (Ellipticity) There exists a non-negative function $\underline{c}:\mathbb{R}^d\to\mathbb{R}_+$ such that for every $r\in[0,T],v,x\in\mathbb{R}^d,z\in\mathbb{R}^d,\rho\in\mathcal{P}_1(\mathbb{R}^d),\zeta\in\mathbb{R}^d$, we have

$$\sum_{j=1}^d\langle\partial_{z_j}c(r,v,z,x,\rho),\zeta\rangle^2\geq\underline{c}(z)|\zeta|^2.$$

Remark. We notice that together with **Hypothesis 2.1**, we have $\underline{c}(z)\leq|\bar{c}(z)|^2,\ \forall z\in\mathbb{R}^d$.

Hypothesis 2.4
We give some supplementary hypotheses concerning the function $\underline{c}$ and the measure μ.
$a)$ We assume that there exists a $\theta>0$ such that

$$\underline{\lim}_{u\to+\infty}\frac{1}{\ln u}\nu\{\underline{c}\geq\frac{1}{u}\}:=\theta>0, \tag{18}$$

with

$$\nu(dz) = \sum_{k=1}^{\infty} \mathbb{1}_{[k-\frac{3}{4}, k-\frac{1}{4}]}(|z|)\mu(dz).$$

This means that $\underline{c}$ could not be too small so that we could have enough noises to deduce the non-degeneracy of the Malliavin covariance matrix (see Section 4.2 (137)).

Remark. If $\mu(\mathbb{R}^d) < \infty$, then $\theta = 0$. So (18) implies that $\mu(\mathbb{R}^d) = \infty$.

b) We assume that μ is absolutely continuous with respect to the Lebesgue measure: $\mu(dz) = h(z)dz$, where h is infinitely differentiable and $\ln h$ has bounded derivatives of any order.

Remark. We need this hypothesis to construct the integration by parts framework for the jump equations.

2.3 The Euler scheme

Now we construct the Euler scheme. For any partition $\mathcal{P} = \{0 = r_0 < r_1 < \cdots < r_{n-1} < r_n = T\}$ of the interval $[0, T]$, we define $\tau(r) = r_k$ when $r \in [r_k, r_{k+1})$, and we consider the equation

$$\begin{aligned} X_t^{\mathcal{P}} &= X_0 + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P}}, \rho_{\tau(r)}^{\mathcal{P}})dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P}}, \rho_{\tau(r)-}^{\mathcal{P}}) N_{\rho_{\tau(r)-}^{\mathcal{P}}}(dv, dz, dr), \end{aligned} \tag{19}$$

where $\rho_t^{\mathcal{P}}$ is the law of $X_t^{\mathcal{P}}$ and $N_{\rho_t^{\mathcal{P}}}(dv, dz, dr)$ is a Poisson point measure with intensity measure $\rho_t^{\mathcal{P}}(dv)\mu(dz)dr$, independent of X_0.

In [3](Theorem 3.5, 3.7, 3.8, Proposition 3.9), Alfonsi and Bally have proved that under some suitable regularity conditions on the coefficients b and c (which are some conditions weaker than the **Hypothesis 2.1** in this paper), the strong solution of the equation (14) exists and is unique, and the following statements are true.

a) There exists a constant C depending on T such that for every $0 < t \leq T$ and every partition $\mathcal{P}$ of $[0, T]$,

$$W_1(X_t^{\mathcal{P}}, X_t) \leq C|\mathcal{P}|, \quad \text{with} \quad |\mathcal{P}| := \max_{k \in \{0, \cdots, n-1\}} (r_{k+1} - r_k). \tag{20}$$

b) The solution of the following weak equation exists.

$$\begin{aligned} \forall \phi \in C_b^1(\mathbb{R}^d), \qquad & \int_{\mathbb{R}^d} \phi(x)\rho_t(dx) = \int_{\mathbb{R}^d} \phi(x)\rho_0(dx) + \int_0^t \int_{\mathbb{R}^d} \langle b(r, x, \rho_r), \nabla\phi(x)\rangle \rho_r(dx)dr \\ & + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} \rho_r(dx)\rho_r(dv) \int_{\mathbb{R}^d} (\phi(x + c(r, v, z, x, \rho_r)) - \phi(x))\mu(dz)dr. \end{aligned} \tag{21}$$

And the solution of the equation (14) is the probabilistic interpretation of (21) in the sense that $\rho_t = \mathcal{L}(X_t)$ (the law of X_t) solves (21).

We recall the notation θ in **Hypothesis 2.4**. One aim of this paper is to prove the following error estimate.

Theorem 2.1. *Under the **Hypothesis 2.1, Hypothesis 2.2, Hypothesis 2.3** and **Hypothesis 2.4**, we have*

a) For any $0 < t \leq T$, when $t > \frac{8d(l+d)}{\theta}$, the law of X_t has a $l-$times differentiable density p_t:

$$\mathbb{P}(X_t \in dx) = \rho_t(dx) = p_t(x)dx, \tag{22}$$

and the density p_t is a function solution of the equation (21).

b) For any $\varepsilon > 0$, there exists a constant C depending on ε, d and T such that for every partition $\mathcal{P}$ of $[0, T]$ with $|\mathcal{P}| \leq 1$, when $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$,

$$d_{TV}(X_t^{\mathcal{P}}, X_t) \leq C|\mathcal{P}|^{1-\varepsilon}. \tag{23}$$

Remark. In the case $\theta = \infty$, the results in **Theorem 2.1** hold for every $0 < t \leq T$.

The proof of this theorem will be given in Section 3.3. The main methods we will use in the proofs are the Malliavin calculus techniques introduced in [7]. We will discuss them in Section 3.

2.4 The truncated Euler scheme

Since we have $\mu(\mathbb{R}^d) = \infty$ (which is a consequence of (18)), we have infinitely many jumps. We use a truncation argument in order to have finite numbers of jumps and obtain a representation by means of a compound Poisson process. This is necessary in order to obtain a scheme which may be simulated. For $M \in \mathbb{N}$, we denote $B_M = \{z \in \mathbb{R}^d : |z| \leq M\}$, $c_M(r, v, z, x, \rho) := c(r, v, z, x, \rho)\mathbb{1}_{B_M}(z)$, and

$$a_T^M := \sqrt{T \int_{\{|z|>M\}} \underline{c}(z)\mu(dz)}. \tag{24}$$

This is a deterministic sequence such that $a_T^M \to 0$ as $M \to \infty$. We also denote $\Delta = (\Delta_1, \cdots, \Delta_d)$ to be a $d-$dimensional standard Gaussian random variable independent of X_0 and N_ρ. Now we cancel the "big jumps" (the jumps of size $|z| > M$) and replace them by a Gaussian random variable $a_t^M \Delta$.

$$\begin{aligned} X_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) N_{\rho_{\tau(r)-}^{\mathcal{P},M}}(dv, dz, dr), \end{aligned} \tag{25}$$

where $\rho_t^{\mathcal{P},M}$ is the law of $X_t^{\mathcal{P},M}$, and $N_{\rho_t^{\mathcal{P},M}}(dv, dz, dr)$ is a Poisson point measure with intensity measure $\rho_t^{\mathcal{P},M}(dv)\mu(dz)dr$, independent of X_0 and of Δ. We remark that Δ is necessary in order to obtain the non-degeneracy of the Malliavin covariance matrix, which will be discussed in detail in Section 4.2.

The advantage of considering $X_t^{\mathcal{P},M}$ is that we may represent it by means of compound Poisson processes. For $k \in \mathbb{N}$, we denote $I_1 = B_1$, $I_k = B_k \backslash B_{k-1}$ for $k \geq 2$ and take $(J_t^k)_{t\in[0,T]}$ a Poisson process of intensity $\mu(I_k)$. We denote by $(T_i^k)_{i\in\mathbb{N}}$ the jump times of $(J_t^k)_{t\in[0,T]}$ and we consider some sequences of independent random variables $Z_i^k \sim \mathbb{1}_{I_k}(z)\frac{\mu(dz)}{\mu(I_k)}$, and $V_{k,i}^{\mathcal{P},M} \sim \rho_{\tau(T_i^k)-}^{\mathcal{P},M}(dv)$, $k, i \in \mathbb{N}$. Moreover, $((J_t^k)_{t\in[0,T]}, (Z_i^k)_{k,i\in\mathbb{N}}, (V_{k,i}^{\mathcal{P},M})_{k,i\in\mathbb{N}}, X_0, \Delta)$ are taken to be independent. Then in order to do the simulation, we represent the jump's parts of the equation (25) by compound Poisson processes:

$$X_t^{\mathcal{P},M} = X_0 + a_T^M \Delta + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})dr + \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} c(\tau(T_i^k), V_{k,i}^{\mathcal{P},M}, Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M}). \tag{26}$$

Notice that the solution of the equation (26) may be constructed in an explicit way (except for $\rho_{\tau(r)}^{\mathcal{P},M}$ and $\rho_{\tau(T_i^k)-}^{\mathcal{P},M}$ which will be discussed in detail in Section 2.5).

We denote

$$\varepsilon_M := \int_{\{|z|>M\}} |\bar{c}(z)|^2 \mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^2, \tag{27}$$

and recall the notation θ in **Hypothesis 2.4**. We obtain the following error estimate for $X_t^{\mathcal{P},M}$.

Theorem 2.2. *Under the **Hypothesis 2.1, Hypothesis 2.2, Hypothesis 2.3** and **Hypothesis 2.4**, we have*

a) For any $0 < t \leq T$, the law of $X_t^{\mathcal{P},M}$ has a smooth density $p_t^{\mathcal{P},M}$.

b) For any $\varepsilon > 0$, there exists a constant C depending on ε, d and T such that for every partition $\mathcal{P}$ of $[0, T]$ with $|\mathcal{P}| \leq 1$, every $M \in \mathbb{N}$ with $\varepsilon_M \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$, when $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$,

$$d_{TV}(X_t^{\mathcal{P},M}, X_t) \leq C(\sqrt{\varepsilon_M} + |\mathcal{P}|)^{1-\varepsilon}. \tag{28}$$

Remark. In the case $\theta = \infty$, the results in **Theorem 2.2** hold for every $0 < t \leq T$.

The proof of this theorem will be given in Section 3.3 by using some Malliavin integration by parts techniques.

2.5 The particle system

We notice that we still cannot compute $\rho^{\mathcal{P},M}_{\tau(r)}$ and $\rho^{\mathcal{P},M}_{\tau(T^k_i)-}$ directly in (26), so we construct the particle system as follows in order to obtain an explicit scheme of simulation. For a random vector $X = (X^1, \cdots, X^N),\ X^i \in \mathbb{R}^d,\ i = 1, \cdots, N$ with a fixed dimension N, we associate the (random) empirical measure

$$\widehat{\rho}(X)(dv) = \frac{1}{N}\sum_{i=1}^{N} \delta_{X^i}(dv), \tag{29}$$

where $\delta_x(dv)$ is the Dirac measure. Now we consider an initial vector $(X^1_0, \cdots, X^N_0)$ with components which are independent and identically distributed with common law ρ_0 (we recall that ρ_0 is the law of X_0 in (14)), and we consider $(\Delta^1, \cdots, \Delta^N)$ which is a $N \times d-$dimensional standard Gaussian random variable independent of $(X^1_0, \cdots, X^N_0)$. Then we construct the particle system $\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_t = (X^{\mathcal{P},M,1}_t, \cdots, X^{\mathcal{P},M,N}_t)$:

$$\begin{aligned} X^{\mathcal{P},M,i}_t &= X^i_0 + a^M_T \Delta^i + \int_0^t b(\tau(r), X^{\mathcal{P},M,i}_{\tau(r)}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_{\tau(r)}))dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X^{\mathcal{P},M,i}_{\tau(r)-}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_{\tau(r)-}))N^i_{\widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_{\tau(r)-})}(dv, dz, dr),\ i = 1, \cdots, N, \end{aligned} \tag{30}$$

where $N^i_{\widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_t)}(dv, dz, dr),\ i = 1, \cdots, N$ are Poisson point measures that are independent each other conditionally to $\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_t$ and independent of $(X^1_0, \cdots, X^N_0, \Delta^1, \cdots, \Delta^N)$ with intensity $\widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_t)(dv)\mu(dz)dr$. We give now the representation of the above equation in terms of compound Poisson processes. This is necessary in order to obtain an explicit simulation algorithm. We recall that we denote $I_1 = B_1$, $I_k = B_k \backslash B_{k-1}$ for $k \geq 2$. Now for $i = 1, \cdots, N$, $k \in \mathbb{N}$, we take $(J^{k,i}_t)_{t\in[0,T]}$ a Poisson process of intensity $\mu(I_k)$. We denote by $(T^{k,i}_l)_{l\in\mathbb{N}}$ the jump times of $(J^{k,i}_t)_{t\in[0,T]}$ and we consider some sequences of independent random variables $Z^{k,i}_l \sim \mathbb{1}_{I_k}(z)\frac{\mu(dz)}{\mu(I_k)}$ and $U^{k,i}_l$ uniformly distributed on $\{1, \cdots, N\}$, for all $i = 1, \cdots, N$, $k, l \in \mathbb{N}$. Moreover, $((J^{k,i}_t)_{\substack{t\in[0,T] \\ i=1,\cdots,N,\ k\in\mathbb{N}}}, (Z^{k,i}_l)_{\substack{i=1,\cdots,N \\ k,l\in\mathbb{N}}}, (U^{k,i}_l)_{\substack{i=1,\cdots,N \\ k,l\in\mathbb{N}}}, (\Delta^1, \cdots, \Delta^N), (X^1_0, \cdots, X^N_0))$ are taken to be independent. Then we represent the jump's parts of the equation (30) by compound Poisson processes to give an explicit scheme of simulation.

$$\begin{aligned} X^{\mathcal{P},M,i}_t &= X^i_0 + a^M_T \Delta^i + \int_0^t b(\tau(r), X^{\mathcal{P},M,i}_{\tau(r)}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_{\tau(r)}))dr \\ &+ \sum_{k=1}^{M}\sum_{l=1}^{J^{k,i}_t} c(\tau(T^{k,i}_l), X^{\mathcal{P},M,U^{k,i}_l}_{\tau(T^{k,i}_l)-}, Z^{k,i}_l, X^{\mathcal{P},M,i}_{\tau(T^{k,i}_l)-}, \widehat{\rho}(\overrightarrow{\boldsymbol{X}}^{\mathcal{P},M}_{\tau(T^{k,i}_l)-})). \end{aligned} \tag{31}$$

So now the solution of the equation (31) is constructed in an explicit way.

We denote

$$V_N := \mathbb{1}_{d=1}N^{-\frac{1}{2}} + \mathbb{1}_{d=2}N^{-\frac{1}{2}}\log(1+N) + \mathbb{1}_{d\geq 3}N^{-\frac{1}{d}}, \tag{32}$$

and we consider the $d-$dimensional regularization kernels

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}}e^{-\frac{|x|^2}{2}}, \quad \varphi_\delta(x) = \frac{1}{\delta^d}\varphi(\frac{x}{\delta}), \quad 0 < \delta \leq 1. \tag{33}$$

We recall the notations ε_M in (27) and θ in **Hypothesis 2.4**. In **Theorem 2.1**, we proved that under appropriate hypotheses, $\mathcal{L}(X_t)(dx) = p_t(x)dx$. We give now some approximation results for $p_t(x)$.

Theorem 2.3. *Under the* ***Hypothesis 2.1, Hypothesis 2.2, Hypothesis 2.3*** *and* ***Hypothesis 2.4****, for every partition $\mathcal{P}$ of $[0,T]$ and every $M \in \mathbb{N}$ with $|\mathcal{P}| + \sqrt{\varepsilon_M} \leq 1$, we have the following:*

i) We take

$$\delta = (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{d+3}}, \quad \text{and take } N \text{ such that} \quad V_N \leq |\mathcal{P}| + \sqrt{\varepsilon_M}.$$

When $t > \frac{8d}{\theta}(2+d)$,

$$p_t(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\varphi_\delta(X_t^{\mathcal{P},M,i} - x) + \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{2}{d+3}}), \tag{34}$$

where $\mathcal{O}(\bullet)$ is the big $\mathcal{O}$ notation (i.e. for a strictly positive function g defined on $\mathbb{R}_+$, $\exists C > 0$, s.t. $|\mathcal{O}(g(y))| \leq Cg(y)$).

*ii) **(Romberg)** We take*

$$\delta = (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{d+5}}, \quad \text{and take } N \text{ such that} \quad V_N \leq |\mathcal{P}| + \sqrt{\varepsilon_M}.$$

When $t > \frac{8d}{\theta}(4+d)$,

$$p_t(x) = \frac{2}{N}\sum_{i=1}^{N} \mathbb{E}\varphi_{\delta/\sqrt{2}}(X_t^{\mathcal{P},M,i} - x) - \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\varphi_\delta(X_t^{\mathcal{P},M,i} - x) + \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{4}{d+5}}). \tag{35}$$

Theorem 2.4. *We suppose* ***Hypothesis 2.1, Hypothesis 2.2, Hypothesis 2.3*** *and* ***Hypothesis 2.4*** *hold true. We take supplementally $\widetilde{\Delta}$ to be a $d-$dimensional standard Gaussian random variable independent of $\overrightarrow{\boldsymbol{X}}_t^{\mathcal{P},M}$. Let $\mathcal{P}$ be a partition of $[0,T]$ with $|\mathcal{P}| \leq 1$, and let $M \in \mathbb{N}$ with $\varepsilon_M \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$. For any $\varepsilon > 0$, for every measurable and bounded function f, when $t > \frac{8d}{\theta}(\frac{16}{\varepsilon} + 1)$, we have the followings.*

i) We take

$$\delta = (|\mathcal{P}| + \varepsilon_M)^{\frac{1}{2}(1-\varepsilon')} \quad \text{and take } N \text{ such that} \quad V_N \leq (|\mathcal{P}| + \varepsilon_M)^{\frac{d+3}{2}(1-\varepsilon'')},$$

with $\varepsilon' = \frac{\varepsilon}{2-\varepsilon}$ and $\varepsilon'' = \frac{(d+5)\varepsilon - 2\varepsilon^2}{(d+3)(2-\varepsilon)}$. Then

$$\int_{\mathbb{R}^d} f(x)p_t(x)dx \quad = \quad \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}f(X_t^{\mathcal{P},M,i} + \delta\widetilde{\Delta}) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}). \tag{36}$$

*ii) **(Romberg)** We take*

$$\delta = (|\mathcal{P}| + \varepsilon_M)^{\frac{1}{4}(1-\varepsilon_\prime)} \quad \text{and take } N \text{ such that} \quad V_N \leq (|\mathcal{P}| + \varepsilon_M)^{\frac{d+5}{4}(1-\varepsilon_{\prime\prime})},$$

with $\varepsilon_\prime = \frac{\varepsilon^2}{2-\varepsilon}$ and $\varepsilon_{\prime\prime} = \frac{8\varepsilon + (d-3)\varepsilon^2}{(d+5)(2-\varepsilon)}$. Then

$$\int_{\mathbb{R}^d} f(x)\rho_t(dx) \quad = \quad \frac{2}{N}\sum_{i=1}^{N} \mathbb{E}f(X_t^{\mathcal{P},M,i} + \frac{\delta}{\sqrt{2}}\widetilde{\Delta}) - \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}f(X_t^{\mathcal{P},M,i} + +\delta\widetilde{\Delta}) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}). \tag{37}$$

Remark. In the case $\theta = \infty$, the results in **Theorem 2.3** and **Theorem 2.4** hold for every $0 < t \leq T$.

Remark. We remark that we have determined δ, N, and we obtain an explicit formula to simulate the density function $p_t(x)$, which is a function solution of the analytical equation (21). We also give the error of this simulation scheme explicitly. We notice that (35), the scheme based on Romberg method, gives a faster speed of convergence than (34): we have the power $\frac{4}{d+5} > \frac{2}{d+3}$.

Remark. We mention that we obtain the results of **Theorem 2.3** directly without using the previous estimates (**Theorem 2.2**), but the speed of convergence depends on the dimension d. So when d is large, the speed of convergence is very slow. However for **Theorem 2.4**, we need to use the previous estimates **Theorem 2.2** to obtain (36). The advantage of considering (36) is that the speed of convergence no longer depends on the dimension d. So we keep the speed of convergence even for large dimension. Finally, (37) is a better simulation scheme in the sense that the numbers of particles N we need is smaller than the one in (36) and δ is larger than the one in (36). We also stress that the speed of convergence in (36) and (37) is $(|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}$, the same as in **Theorem 2.2** (28) for the truncated Euler scheme.

The proof of this theorem will be given in Section 3.3 by using Malliavin integration by parts techniques.

2.6 Some examples

We give some typical examples to illustrate our main results.

Example 1 We take $h = 1$ so the measure μ is the Lebesgue measure. We consider two types of behaviour for c.

i) Exponential decay We assume that $|\bar{c}(z)|^2 = e^{-a_1|z|^p}$ and $\underline{c}(z) = e^{-a_2|z|^p}$ with some constants $0 < a_1 \leq a_2$, $p > 0$. We only check **Hypothesis 2.4** here. We have

$$\nu\{\underline{c} > \frac{1}{u}\} = \nu\{|z| < (\frac{\ln u}{a_2})^{\frac{1}{p}}\} \geq \frac{r_d}{2}(\frac{\ln(u-1)}{a_2})^{\frac{d}{p}},$$

with r_d the volume of the unit ball in $\mathbb{R}^d$, so that

$$\frac{1}{\ln u}\nu\{\underline{c} > \frac{1}{u}\} \geq \frac{r_d}{2(a_2)^{\frac{d}{p}}}\frac{(\ln(u-1))^{\frac{d}{p}}}{\ln u}.$$

We notice that $\theta = 0$ when $p > d$; $\theta = \infty$ when $0 < p < d$; and $\theta = \frac{r_d}{2a_2}$ when $p = d$. Therefore, when $p > d$, we can say nothing in **Theorem 2.1** and **Theorem 2.2**; when $0 < p < d$, all the results in **Theorem 2.1** and **Theorem 2.2** are true for every $0 < t \leq T$; and when $p = d$, (22) holds true for $t > \frac{8d(3l+2)a_2}{r_d}$, (23) and (28) hold true for $t > \frac{16da_2}{r_d}(\frac{2}{\varepsilon} + 1)$.

ii) Polynomial decay We assume that $|\bar{c}(z)|^2 = \frac{a_1}{1+|z|^p}$ and $\underline{c}(z) = \frac{a_2}{1+|z|^p}$ for some constants $0 < a_2 \leq a_1$ and $p > d$. Then

$$\nu\{\underline{c} > \frac{1}{u}\} = \nu\{|z| < (a_2u - 1)^{\frac{1}{p}}\} \geq \frac{r_d}{2}(a_2(u-1) - 1)^{\frac{d}{p}},$$

so that

$$\frac{1}{\ln u}\nu\{\underline{c} > \frac{1}{u}\} \geq \frac{r_d}{2}\frac{(a_2(u-1) - 1)^{\frac{d}{p}}}{\ln u}.$$

We notice that in this case, $\theta = \infty$. Thus, all the results in **Theorem 2.1** and **Theorem 2.2** holds for every $0 < t \leq T$.

Example 2 We consider the $(1-$dimensional) truncated $\alpha-$stable process: $X_t = X_0 + \int_0^t \sigma(X_{r-})dU_r$. Here $(U_t)_{t\geq 0}$ is a (pure jump) Lévy process with intensity measure

$$\widetilde{\mu}(dz) = \mathbb{1}_{\{|z|\leq 1\}}\frac{1}{|z|^{1+\alpha}}dz, \quad 0 \leq \alpha < 1.$$

We assume that $\sigma \in C_b^\infty(\mathbb{R})$, $0 < \underline{\sigma} \leq \sigma(x) \leq \bar{\sigma}$ and $-1 < \underline{a} \leq \sigma'(x) \leq \bar{\sigma}$, $\forall x \in \mathbb{R}$, for some universal constants $\bar{\sigma}, \underline{\sigma}, \underline{a}$, where σ' is the differential of σ in x. Then by a change of variable $z \mapsto \frac{1}{z}$, we come back to the setting of this paper with $c(r, v, z, x, \rho) = \sigma(x) \times \frac{1}{z}$ and $\mu(dz) = \mathbb{1}_{\{|z|\geq 1\}}\frac{1}{|z|^{1-\alpha}}dz$. In this case, $\underline{c}(z) = \underline{\sigma} \times \frac{1}{|z|^4}$, then

$$\frac{1}{\ln u}\nu\{\underline{c} > \frac{1}{u}\} \geq \frac{1}{\ln u}\int_1^{(\underline{\sigma}(u-1))^{\frac{1}{4}}} \frac{1}{|z|^{1-\alpha}}dz = \frac{(\underline{\sigma}(u-1))^{\frac{\alpha}{4}} - 1}{\alpha \ln u},$$

so that $\theta = \infty$. Thus, all the results in **Theorem 2.1** and **Theorem 2.2** hold for every $0 < t \leq T$.

2.7 Preliminaries: coupling

Before moving on to the next section, we make some preliminary computations here. For some technical reasons, besides the truncated Euler scheme (25), we also consider the truncation of the original equation (14) as follows (with a_T^M, Δ and c_M defined in Section 2.4).

$$\begin{aligned} X_t^M &= X_0 + a_T^M \Delta + \int_0^t b(r, X_r^M, \rho_r) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(r, v, z, X_{r-}^M, \rho_{r-}) N_{\rho_{r-}}(dv, dz, dr). \end{aligned} \tag{38}$$

We notice that we keep ρ_r (the law of X_r) instead of taking ρ_r^M (the law of X_r^M) to simplify the calculation below, so the equation (38) is just an intermediate equation (which is not used for simulation).

We notice that the jump's parts of $X_t^{\mathcal{P},M}$ and X_t^M solutions of (25), (38) are defined with respect to different Poisson point measures (on different probability spaces), so it is not possible to estimate the L^2 distance between them directly (we need to estimate the L^2 distance later in the proof of **Lemma 3.9**). To overcome this difficulty, we use similar equations driven by the same Poisson point measure. This is done by a coupling procedure. In this section, we make a coupling argument to construct $x_t, x_t^{\mathcal{P}}$ and $x_t^{\mathcal{P},M}$ which have the same law as $X_t, X_t^{\mathcal{P}}$ and $X_t^{\mathcal{P},M}$ but are defined on the same probability space and verify equations driven by the same Poisson point measure.

We remark that the basic distance which appears in our framework is W_1 (see (12)). However for technical reasons, we need to estimate the distance $W_{2+\varepsilon_*}$ (defined immediately below) for some small $\varepsilon_* > 0$. This is because we need L^2 estimate in **Lemma 3.9** and we have to use the Hölder inequality with conjugates $1 + \frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$. So now we take $\varepsilon_* > 0$ which is small enough. For $\rho_1, \rho_2 \in \mathcal{P}_{2+\varepsilon_*}(\mathbb{R}^d)$, we denote the Wasserstein distance of order $2 + \varepsilon_*$ by

$$W_{2+\varepsilon_*}(\rho_1, \rho_2) = \inf_{\pi \in \Pi(\rho_1, \rho_2)} \big\{ \big(\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^{2+\varepsilon_*} \pi(dx, dy) \big)^{\frac{1}{2+\varepsilon_*}} \big\},$$

where $\Pi(\rho_1, \rho_2)$ is the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ρ_1 and ρ_2. Some basic properties of $W_p, p \geq 1$ can be found in [37] and [51] for example, and we mention that $W_1(\rho_1, \rho_2) \leq W_{2+\varepsilon_*}(\rho_1, \rho_2)$.

Now we make the optimal coupling in $W_{2+\varepsilon_*}$ distance between $X_{\tau(t)-}^{\mathcal{P},M}$ and X_{t-}. We recall that $\rho_{\tau(t)-}^{\mathcal{P},M}$ is the law of $X_{\tau(t)-}^{\mathcal{P},M}$ and ρ_{t-} is the law of X_{t-}. For every partition $\mathcal{P}$, $M \in \mathbb{N}$ and time $0 < t \leq T$, one can easily check that $\rho_{\tau(t)-}^{\mathcal{P},M}$ and ρ_{t-} both belong to $\mathcal{P}_{2+\varepsilon_*}(\mathbb{R}^d)$. This is a consequence of **Hypothesis 2.1** and of (16) with

$$\Phi(r, v, z, \omega, \rho) = c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M})$$

and with

$$\Phi(r, v, z, \omega, \rho) = c(r, v, z, X_{r-}, \rho_{r-}).$$

Then we take $\Pi_t^{\mathcal{P},M}(dv_1, dv_2)$ to be the optimal $W_{2+\varepsilon_*}-$coupling of $\rho_{\tau(t)-}^{\mathcal{P},M}(dv_1)$ and $\rho_{t-}(dv_2)$. So we have

$$(W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t-}))^{2+\varepsilon_*} = \int_{\mathbb{R}^d \times \mathbb{R}^d} |v_1 - v_2|^{2+\varepsilon_*} \Pi_t^{\mathcal{P},M}(dv_1, dv_2).$$

We will need the representation of $\Pi_t^{\mathcal{P},M}(dv_1, dv_2)$ by means of the Lebesgue measure dw on $[0, 1]$. This will be done by using the following lemma.

Lemma 2.5. *There exists a measurable map $\Phi : [0, 1) \times \mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}^d$ such that for any $\rho \in \mathcal{P}_1(\mathbb{R}^d)$, any bounded and measurable function $\phi : \mathbb{R}^d \to \mathbb{R}$, we have*

$$\int_0^1 \phi(\Phi(w, \rho)) dw = \int_{\mathbb{R}^d} \phi(x) \rho(dx).$$

This result is stated in [14] and is useful when we estimate the L^p distance. We construct $(\eta_t^1(w), \eta_t^2(w))$ which represents $\Pi_t^{\mathcal{P},M}$ in the sense of **Lemma 2.5**, this means

$$\int_0^1 \phi(\eta_t^1(w), \eta_t^2(w))dw = \int_{\mathbb{R}^d\times\mathbb{R}^d} \phi(v_1, v_2)\Pi_t^{\mathcal{P},M}(dv_1, dv_2).$$

In particular, this gives for any measurable and bounded function $f : \mathbb{R}^d \to \mathbb{R}$,

$$\int_0^1 f(\eta_t^1(w))dw = \int_{\mathbb{R}^d} f(v_1)\rho_{\tau(t)-}^{\mathcal{P},M}(dv_1), \quad \int_0^1 f(\eta_t^2(w))dw = \int_{\mathbb{R}^d} f(v_2)\rho_{t-}(dv_2),$$
$$\int_0^1 |\eta_t^1(w) - \eta_t^2(w)|^{2+\varepsilon_*}dw = \int_{\mathbb{R}^d\times\mathbb{R}^d} |v_1 - v_2|^{2+\varepsilon_*}\Pi_t^{\mathcal{P},M}(dv_1, dv_2) = (W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t)-}))^{2+\varepsilon_*}. \quad (39)$$

Now we construct a Poisson point measure $\mathcal{N}(dw, dz, dr)$ on the state space $[0, 1] \times \mathbb{R}^d$ with intensity measure $dw\mu(dz)dr$. Then we consider the equations

$$x_t = X_0 + \int_0^t b(r, x_r, \rho_r)dr + \int_0^t \int_{[0,1]\times\mathbb{R}^d} c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-})\mathcal{N}(dw, dz, dr), \quad (40)$$

$$x_t^M = X_0 + a_T^M\Delta + \int_0^t b(r, x_r^M, \rho_r)dr + \int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(r, \eta_r^2(w), z, x_{r-}^M, \rho_{r-})\mathcal{N}(dw, dz, dr), \quad (41)$$

$$\begin{aligned} x_t^{\mathcal{P},M} &= X_0 + a_T^M\Delta + \int_0^t b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})dr \\ &+ \int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M})\mathcal{N}(dw, dz, dr). \end{aligned} \quad (42)$$

One can check by Itô formula that $x_t^{\mathcal{P},M}$ has the same law as $X_t^{\mathcal{P},M}$ (solution of (25)), x_t^M has the same law as X_t^M (solution of (38)) and x_t has the same law as X_t (solution of (19)). Then

$$\begin{aligned} (W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t-}))^{2+\varepsilon_*} = (W_{2+\varepsilon_*}(\mathcal{L}(X_{\tau(t)-}^{\mathcal{P},M}), \mathcal{L}(X_{t-})))^{2+\varepsilon_*} &= (W_{2+\varepsilon_*}(\mathcal{L}(x_{\tau(t)-}^{\mathcal{P},M}), \mathcal{L}(x_{t-})))^{2+\varepsilon_*} \\ &\leq \mathbb{E}|x_{\tau(t)-}^{\mathcal{P},M} - x_{t-}|^{2+\varepsilon_*}. \end{aligned} \quad (43)$$

Remark. We also have the following consequence of Burkholder inequality (as a variant of (16) and (17)): Let $\bar{\Phi}(r, w, z, \omega, \rho) : [0, T]\times[0, 1]\times\mathbb{R}^d\times\Omega\times\mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}_+$ and $\bar{\varphi}(r, w, \omega, \rho) : [0, T]\times[0, 1]\times\Omega\times\mathcal{P}_1(\mathbb{R}^d) \to \mathbb{R}_+$ be two non-negative functions.

a) Then for any $p \geq 2$,

$$\begin{aligned} &\mathbb{E}\Big|\int_0^t \int_{[0,1]\times\mathbb{R}^d} \bar{\Phi}(r, w, z, \omega, \rho)\mathcal{N}(dw, dz, dr)\Big|^p \\ &\leq C[\mathbb{E}\int_0^t \int_0^1 \int_{\mathbb{R}^d} |\bar{\Phi}(r, w, z, \omega, \rho)|^p\mu(dz)dwdr + \mathbb{E}\int_0^t \int_0^1 |\int_{\mathbb{R}^d} |\bar{\Phi}(r, w, z, \omega, \rho)|\mu(dz)|^p dwdr \\ &+\mathbb{E}\int_0^t \int_0^1 |\int_{\mathbb{R}^d} |\bar{\Phi}(r, w, z, \omega, \rho)|^2\mu(dz)|^{\frac{p}{2}} dwdr], \end{aligned} \quad (44)$$

where C is a constant depending on p, T.

b) If we have

$$|\bar{\Phi}(r, w, z, \omega, \rho)| \leq |\bar{c}(z)||\bar{\varphi}(r, w, \omega, \rho)|,$$

then for any $p \geq 2$,

$$\mathbb{E}\Big|\int_0^t \int_{[0,1]\times\mathbb{R}^d} \bar{\Phi}(r, w, z, \omega, \rho)\mathcal{N}(dw, dz, dr)\Big|^p \leq C\mathbb{E}\int_0^t \int_0^1 |\bar{\varphi}(r, w, \omega, \rho)|^p dwdr. \quad (45)$$

Then we obtain the following consequence. We recall by (27) that $\varepsilon_M = \int_{\{|z|>M\}} |\bar{c}(z)|^2\mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^2)$.

Lemma 2.6. *Assume that the **Hypothesis 2.1** holds true. Then there exists a constant C dependent on T and ε_*, for every M such that $\varepsilon_M \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$, we have*

$$i) \quad \mathbb{E}|x_t^M - x_t|^{2+\varepsilon_*} \leq C\varepsilon_M \to 0.$$

And for every partition $\mathcal{P}$ with $|\mathcal{P}| \leq 1$, we have

$$ii) \quad \mathbb{E}|x_t^{\mathcal{P},M} - x_t^M|^{2+\varepsilon_*} \leq C(|\mathcal{P}| + \varepsilon_M),$$

$$iii) \quad \mathbb{E}|x_t^{\mathcal{P},M} - x_t|^{2+\varepsilon_*} \leq C(|\mathcal{P}| + \varepsilon_M),$$

$$iv) \quad W_{2+\varepsilon_*}(x_t^{\mathcal{P},M}, x_t) \leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{1}{2+\varepsilon_*}}.$$

Proof. We only prove $i)$ and $iii)$, since $ii)$ is a direct consequence of $i)$ and $iii)$, and $iv)$ is an immediate consequence of $iii)$.

Proof of $i)$: We write $\mathbb{E}|x_t^M - x_t|^{2+\varepsilon_*} \leq E_0 + E_1 + E_2$, where $E_0 = |a_T^M|^{2+\varepsilon_*}\mathbb{E}|\Delta|^{2+\varepsilon_*} \leq C\varepsilon_M$, and

$$E_1 = \mathbb{E}|\int_0^t (b(r, x_r^M, \rho_r) - b(r, x_r, \rho_r))dr|^{2+\varepsilon_*},$$

$$E_2 = \mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} (c_M(r, \eta_r^2(w), z, x_{r-}^M, \rho_{r-}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}))\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*}$$

Firstly, by **Hypothesis 2.1**,

$$E_1 \leq C\int_0^t \mathbb{E}|x_r^M - x_r^M|^{2+\varepsilon_*}dr. \tag{46}$$

Then by **Hypothesis 2.1**, (44) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = |c_M(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-})|$$

and by (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = |c_M(r, \eta_r^2(w), z, x_{r-}^M, \rho_{r-}) - c_M(r, \eta_r^2(w), z, x_{r-}, \rho_{r-})|,$$

we have

$$\begin{aligned}
E_2 &\leq \mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} (c_M(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}))\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*} \\
&+ \mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} (c_M(r, \eta_r^2(w), z, x_{r-}^M, \rho_{r-}) - c_M(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}))\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*} \\
&\leq C[\int_{\{|z|>M\}} |\bar{c}(z)|^{2+\varepsilon_*}\mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^{2+\varepsilon_*} + |\int_{\{|z|>M\}} |\bar{c}(z)|^2\mu(dz)|^{\frac{2+\varepsilon_*}{2}} \\
&+ \int_0^t \mathbb{E}|x_r - x_r^M|^{2+\varepsilon_*}dr] \\
&\leq C[\varepsilon_M + \int_0^t \mathbb{E}|x_r - x_r^M|^{2+\varepsilon_*}dr].
\end{aligned} \tag{47}$$

Combining (46) and (47), we have

$$\mathbb{E}|x_t^M - x_t|^{2+\varepsilon_*} \leq C[\varepsilon_M + \int_0^t \mathbb{E}|x_r^M - x_r|^{2+\varepsilon_*}dr].$$

So we conclude by Gronwall lemma.

Proof of $iii)$ We write $\mathbb{E}|x_t^{\mathcal{P},M} - x_t|^{2+\varepsilon_*} \leq C[K_0 + K_1 + K_2]$, with $K_0 = |a_T^M|^{2+\varepsilon_*}\mathbb{E}|\Delta|^{2+\varepsilon_*} \leq C\varepsilon_M$, and

$$K_1 = \mathbb{E}|\int_0^t b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M}) - b(r, x_r, \rho_r)dr|^{2+\varepsilon_*},$$

$$K_2 = \mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-})\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*}.$$

Using **Hypothesis 2.1**,

$$\begin{aligned} K_1 &\leq C[|\mathcal{P}|^{2+\varepsilon_*} + \int_0^t \mathbb{E}|x_{\tau(r)}^{\mathcal{P},M} - x_r|^{2+\varepsilon_*}dr + \int_0^t (W_1(\rho_{\tau(r)}^{\mathcal{P},M}, \rho_r))^{2+\varepsilon_*}dr] \\ &\leq C[|\mathcal{P}|^{2+\varepsilon_*} + \int_0^t \mathbb{E}|x_{\tau(r)}^{\mathcal{P},M} - x_r|^{2+\varepsilon_*}dr]. \end{aligned} \tag{48}$$

By **Hypothesis 2.1**, (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = |c(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-})|$$

and by (44) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = |c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) - c(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M})|,$$

we have

$$\begin{aligned} K_2 &\leq C[\mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} (c(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) - c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}))\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*} \\ &+ \mathbb{E}|\int_0^t \int_{[0,1]\times\mathbb{R}^d} (c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) \\ &- c(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}))\mathcal{N}(dw, dz, dr)|^{2+\varepsilon_*}] \\ &\leq C[|\mathcal{P}|^{2+\varepsilon_*} + \int_0^t \int_0^1 |\eta_r^1(w) - \eta_r^2(w)|^{2+\varepsilon_*}dwdr + \int_0^t \mathbb{E}|x_{\tau(r)}^{\mathcal{P},M} - x_r|^{2+\varepsilon_*}dr + \int_0^t (W_1(\rho_{\tau(r)}^{\mathcal{P},M}, \rho_r))^{2+\varepsilon_*}dr \\ &+ \int_{\{|z|>M\}} |\bar{c}(z)|^{2+\varepsilon_*}\mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^{2+\varepsilon_*} + |\int_{\{|z|>M\}} |\bar{c}(z)|^2\mu(dz)|^{\frac{2+\varepsilon_*}{2}}] \\ &\leq C[|\mathcal{P}|^{2+\varepsilon_*} + \int_0^t \mathbb{E}|x_{\tau(r)}^{\mathcal{P},M} - x_r|^{2+\varepsilon_*}dr + \varepsilon_M], \end{aligned} \tag{49}$$

where the last inequality is obtained by (39), (43), and the fact that W_1 distance is upper bounded by $W_{2+\varepsilon_*}$ distance, and so upper bounded by the $L^{2+\varepsilon_*}$ distance.

We notice that by (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}),$$

we have

$$\mathbb{E}|x_{\tau(t)}^{\mathcal{P},M} - x_t^{\mathcal{P},M}|^{2+\varepsilon_*} \leq C|\mathcal{P}|. \tag{50}$$

Combining (48), (49) and (50),

$$\mathbb{E}|x_t^{\mathcal{P},M} - x_t|^{2+\varepsilon_*} \leq C[K_0 + K_1 + K_2] \leq C[|\mathcal{P}| + \int_0^t \mathbb{E}|x_r^{\mathcal{P},M} - x_r|^{2+\varepsilon_*}dr + \varepsilon_M].$$

So finally, we conclude by Gronwall lemma. □

We remark that we may represent the jump's parts of the equations (41) and (42) by means of compound Poisson processes. With all the random variables $((J_t^k)_{t\in[0,T]}, (Z_i^k)_{k,i\in\mathbb{N}}, X_0, \Delta)$ constructed in Section 2.4, we take moreover $(W_i^k)_{k,i\in\mathbb{N}}$ a sequence of independent random variables which are uniformly distributed on $[0,1]$ and independent of $((J_t^k)_{t\in[0,T]}, (Z_i^k)_{k,i\in\mathbb{N}}, X_0, \Delta)$. Then we have

$$x_t^M = X_0 + a_T^M \Delta + \int_0^t b(r, x_r^M, \rho_r)dr + \sum_{k=1}^{M}\sum_{i=1}^{J_t^k} c(T_i^k, \eta^2_{T_i^k}(W_i^k), Z_i^k, x^M_{T_i^k-}, \rho_{T_i^k-}), \tag{51}$$

$$x_t^{\mathcal{P},M} = X_0 + a_T^M \Delta + \int_0^t b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})dr + \sum_{k=1}^{M}\sum_{i=1}^{J_t^k} c(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-}). \tag{52}$$

We recall that the laws of x_t and X_t coincide, $x_t^{\mathcal{P},M}$ has the same law as $X_t^{\mathcal{P},M}$, and x_t^M has the same law as X_t^M. The advantage of considering x_t, x_t^M and $x_t^{\mathcal{P},M}$ is that the jump's parts of them are all defined with respect to the same Poisson point measure, which means that we are able to overcome the problems caused by the "Boltzmann term" (the Poisson point measure depends on the law of the solution). So in the following, instead of dealing with X_t, X_t^M and $X_t^{\mathcal{P},M}$ solutions of (14), (38) and (25), we deal with x_t, x_t^M and $x_t^{\mathcal{P},M}$ solutions of (40), (51) and (52).

3 Malliavin calculus

3.1 Abstract integration by parts framework

Here we recall the abstract integration by parts framework in [7].

We denote C_p^∞ to be the space of smooth functions which, together with all the derivatives, have polynomial growth. We also denote C_p^q to be the space of $q-$times differentiable functions which, together with all the derivatives, have polynomial growth.

We consider a probability space (Ω,$\mathcal{F}$,$\mathbb{P}$), and a subset $\mathcal{S} \subset \bigcap_{p=1}^{\infty} L^p(\Omega;\mathbb{R})$ such that for every $\phi \in C_p^\infty(\mathbb{R}^d)$ and every $F \in \mathcal{S}^d$, we have $\phi(F) \in \mathcal{S}$. A typical example of $\mathcal{S}$ is the space of simple functionals, as in the standard Malliavin calculus. Another example is the space of "Malliavin smooth functionals".

Given a separable Hilbert space $\mathcal{H}$, we assume that we have a derivative operator $D : \mathcal{S} \to \bigcap_{p=1}^{\infty} L^p(\Omega;\mathcal{H})$ which is a linear application which satisfies

$a)$

$$D_h F := \langle DF, h\rangle_{\mathcal{H}} \in \mathcal{S},\ for\ any\ h \in \mathcal{H}, \tag{53}$$

$b)$ $\underline{Chain\ Rule}$: For every $\phi \in C_p^1(\mathbb{R}^d)$ and $F = (F_1, \cdots, F_d) \in \mathcal{S}^d$, we have

$$D\phi(F) = \sum_{i=1}^{d} \partial_i \phi(F) DF_i, \tag{54}$$

Since $D_h F \in \mathcal{S}$, we may define by iteration the derivative operator of higher order $D^q : \mathcal{S} \to \bigcap_{p=1}^{\infty} L^p(\Omega;\mathcal{H}^{\otimes q})$ which verifies $\langle D^q F, \otimes_{i=1}^q h_i\rangle_{\mathcal{H}^{\otimes q}} = D_{h_q} D_{h_{q-1}} \cdots D_{h_1} F$. We also denote $D^q_{h_1,\cdots,h_q} F := \langle D^q F, \otimes_{i=1}^q h_i\rangle_{\mathcal{H}^{\otimes q}}$, for any $h_1, \cdots, h_q \in \mathcal{H}$. Then, $D^q_{h_1,\cdots,h_q} F = D_{h_q} D^{q-1}_{h_1,\cdots,h_{q-1}} F$ $(q \geq 2)$.

We notice that since $\mathcal{H}$ is separable, there exists a countable orthonormal base $(e_i)_{i\in\mathbb{N}}$. We denote

$$D_iF = D_{e_i}F = \langle DF, e_i\rangle_{\mathcal{H}}.$$

Then

$$DF = \sum_{i=1}^{\infty} D_iF \times e_i \quad \text{and} \quad D^qF = \sum_{i_1,\cdots,i_q} D_{i_1,\cdots,i_q}F \times \otimes_{j=1}^{q} e_j.$$

For $F = (F_1, \cdots, F_d) \in \mathcal{S}^d$, we associate the Malliavin covariance matrix

$$\sigma_F = (\sigma_F^{i,j})_{i,j=1,\cdots,d}, \quad \text{with} \quad \sigma_F^{i,j} = \langle DF_i, DF_j\rangle_{\mathcal{H}}. \tag{55}$$

And we denote

$$\Sigma_p(F) = \mathbb{E}(1/\det\sigma_F)^p. \tag{56}$$

We say that F is non-degenerated if $\Sigma_p(F) < \infty$, $\forall p \geq 1$.

We also assume that we have an Ornstein-Uhlenbeck (divergence) operator $L : \mathcal{S} \to \mathcal{S}$ which is a linear operator satisfying

$a)$ $\underline{Duality}$: For every $F, G \in \mathcal{S}$,

$$\mathbb{E}\langle DF, DG\rangle_{\mathcal{H}} = \mathbb{E}(FLG) = \mathbb{E}(GLF), \tag{57}$$

$b)$ $\underline{Chain\ Rule}$: For every $\phi \in C_p^2(\mathbb{R}^d)$ and $F = (F_1, \cdots, F_d) \in \mathcal{S}^d$, we have

$$L\phi(F) = \sum_{i=1}^{d} \partial_i\phi(F)LF_i - \sum_{i=1}^{d}\sum_{j=1}^{d} \partial_i\partial_j\phi(F)\langle DF_i, DF_j\rangle_{\mathcal{H}}.$$

As an immediate consequence of the duality formula, we know that $L : \mathcal{S} \subset L^2(\Omega) \to L^2(\Omega)$ is closable.

Definition 3.1. If $D^q : \mathcal{S} \subset L^2(\Omega) \to L^2(\Omega; \mathcal{H}^{\otimes q})$, $\forall q \geq 1$, are closable, then the triplet $(\mathcal{S}, D, L)$ is called an IbP (Integration by Parts) framework.

Now, we introduce the Sobolev norms. For any $l \geq 1$, $F \in \mathcal{S}$,

$$|F|_{1,l} = \sum_{q=1}^{l} |D^qF|_{\mathcal{H}^{\otimes q}}, \quad |F|_l = |F| + |F|_{1,l}, \tag{58}$$

We put $|F|_0 = |F|$, $|F|_l = 0$ for $l < 0$, and $|F|_{1,l} = 0$ for $l \leq 0$. For $F = (F_1, \cdots, F_d) \in \mathcal{S}^d$, we set

$$|F|_{1,l} = \sum_{i=1}^{d} |F_i|_{1,l}, \quad |F|_l = \sum_{i=1}^{d} |F_i|_l,$$

Moreover, we associate the following norms. For any $l \geq 0, p \geq 1$,

$$\begin{aligned} \|F\|_{l,p} &= (\mathbb{E}\,|F|_l^p)^{1/p}, \quad \|F\|_p = (\mathbb{E}\,|F|^p)^{1/p}, \\ \|F\|_{L,l,p} &= \|F\|_{l,p} + \|LF\|_{l-2,p}. \end{aligned} \tag{59}$$

With these notations, we have the following lemma from [8] (lemma 8 and lemma 10), which is a consequence of the chain rule.

Lemma 3.1. *Let $F \in \mathcal{S}^d$. For every $l \in \mathbb{N}$, if $\phi : \mathbb{R}^d \to \mathbb{R}$ is a $C^l(\mathbb{R}^d)$ function ($l-$times differentiable function), then there is a constant C_l dependent on l such that*

$$a) \quad |\phi(F)|_{1,l} \leq |\nabla\phi(F)||F|_{1,l} + C_l \sup_{2\leq|\beta|\leq l} |\partial^\beta\phi(F)||F|^l_{1,l-1}.$$

If $\phi \in C^{l+2}(\mathbb{R}^d)$, then

$$b) \quad |L\phi(F)|_l \leq |\nabla\phi(F)||LF|_l + C_l \sup_{2\leq|\beta|\leq l+2} |\partial^\beta\phi(F)|(1+|F|^{l+2}_{l+1})(1+|LF|_{l-1}).$$

For $l = 0$, we have

$$c) \quad |L\phi(F)| \leq |\nabla\phi(F)||LF| + \sup_{|\beta|=2} |\partial^\beta\phi(F)||F|^2_{1,1}.$$

We denote by $\mathcal{D}_{l,p}$ the closure of $\mathcal{S}$ with respect to the norm $\|\circ\|_{L,l,p}$:

$$\mathcal{D}_{l,p} = \overline{\mathcal{S}}^{\|\circ\|_{L,l,p}}, \tag{60}$$

and

$$\mathcal{D}_\infty = \bigcap_{l=1}^{\infty}\bigcap_{p=1}^{\infty} \mathcal{D}_{l,p}, \quad \mathcal{H}_l = \mathcal{D}_{l,2}. \tag{61}$$

For an IbP framework $(\mathcal{S}, D, L)$, we now extend the operators from $\mathcal{S}$ to $\mathcal{D}_\infty$. For $F \in \mathcal{D}_\infty$, $p \geq 2$, there exists a sequence $F_n \in \mathcal{S}$ such that $\|F - F_n\|_p \to 0$, $\|F_m - F_n\|_{q,p} \to 0$ and $\|LF_m - LF_n\|_{q-2,p} \to 0$. Since D^q and L are closable, we can define

$$D^qF = \lim_{n\to\infty} D^qF_n \quad in \quad L^p(\Omega;\mathcal{H}^{\otimes q}), \quad LF = \lim_{n\to\infty} LF_n \quad in \quad L^p(\Omega). \tag{62}$$

We still associate the same norms and covariance matrix introduced above for $F \in \mathcal{D}_\infty$.

Lemma 3.2. *The triplet $(\mathcal{D}_\infty, D, L)$ is an IbP framework.*

Proof. The proof is standard and we refer to the lemma 3.1 in [10] for details. □

The following lemma is useful in order to control the Sobolev norms and covariance matrices when passing to the limit.

Lemma 3.3. ***(A)*** *We fix $p \geq 2, l \geq 2$. Let $F \in L^1(\Omega)$ and let $F_n \in \mathcal{S}^d, n \in \mathbb{N}$ such that*

$$\begin{aligned} i) \quad \mathbb{E}\,|F_n - F| \quad &\to \quad 0, \\ ii) \quad \sup_n \|F_n\|_{L,l,p} \quad &\leq \quad K_{l,p} < \infty. \end{aligned}$$

Then for every $1 \leq \bar{p} < p$, we have $F \in \mathcal{D}^d_{l,\bar{p}}$ and $\|F\|_{L,l,\bar{p}} \leq K_{l,\bar{p}}$. Moreover, there exists a convex combination

$$G_n = \sum_{i=n}^{m_n} \gamma^n_i \times F_i \in \mathcal{S}^d,$$

with $\gamma^n_i \geq 0, i = n, ..., m_n$ and $\sum_{i=n}^{m_n} \gamma^n_i = 1$, such that

$$\|G_n - F\|_{L,l,2} \to 0.$$

(B) *For $F \in \mathcal{D}^d_\infty$, we denote*

$$\lambda(F) = \inf_{|\zeta|=1} \langle \sigma_F\zeta, \zeta\rangle$$

*the lowest eigenvalue of the covariance matrix σ_F. We consider some F and F_n which verify $i), ii)$ in **(A)**. We also suppose that*

$$iii) \quad (DF_n)_{n\in\mathbb{N}} \text{ is a Cauchy sequence in } L^2(\Omega;\mathcal{H}),$$

and for every $p \geq 1$,

$$iv) \quad \sup_n \mathbb{E}(\lambda(F_n))^{-p} \leq Q_p < \infty. \tag{63}$$

Then we have

$$\mathbb{E}(\lambda(F))^{-p} \leq Q_p < \infty, \quad \forall p \geq 1.$$

(C) *We suppose that we have $(F, \bar{F})$ and $(F_n, \bar{F}_n)$ which verify the hypotheses of **(A)**. If we also have*

$$v) \quad \sup_n \|DF_n - D\bar{F}_n\|_{L^2(\Omega;\mathcal{H})} \leq \bar{\varepsilon}, \tag{64}$$

then

$$\|DF - D\bar{F}\|_{L^2(\Omega;\mathcal{H})} \leq \bar{\varepsilon}.$$

Proof. **Proof of (A)** For the sake of the simplicity of notations, we only prove for the one dimensional case. The Hilbert space $\mathcal{H}_l = \mathcal{D}_{l,2}$ equipped with the scalar product

$$\begin{aligned} \langle U, V \rangle_{L,l,2} &:= \sum_{q=1}^{l} \mathbb{E}\langle D^q U, D^q V \rangle_{\mathcal{H}^{\otimes q}} + \mathbb{E}(UV) \\ &+ \sum_{q=1}^{l-2} \mathbb{E}\langle D^q LU, D^q LV \rangle_{\mathcal{H}^{\otimes q}} + \mathbb{E}(LU \times LV) \end{aligned}$$

is the space of the functionals which are $l-$times differentiable in L^2 sense. By $ii)$, for $p \geq 2$, $\|F_n\|_{L,l,2} \leq \|F_n\|_{L,l,p} \leq K_{l,p}$. Then, applying Banach Alaoglu theorem, there exists $G \in \mathcal{H}_l$ and a subsequence (we still denote it by n), such that $F_n \to G$ weakly in the Hilbert space $\mathcal{H}_l$. This means that for every $Q \in \mathcal{H}_l$, $\langle F_n, Q \rangle_{L,l,2} \to \langle G, Q \rangle_{L,l,2}$. Therefore, by Mazur theorem, we can construct some convex combination

$$G_n = \sum_{i=n}^{m_n} \gamma_i^n \times F_i \in \mathcal{S}$$

with $\gamma_i^n \geq 0, i = n, ..., m_n$ and $\sum_{i=n}^{m_n} \gamma_i^n = 1$, such that

$$\|G_n - G\|_{L,l,2} \to 0.$$

In particular we have

$$\mathbb{E}\,|G_n - G| \leq \|G_n - G\|_{L,l,2} \to 0.$$

Also, we notice that by i),

$$\mathbb{E}\,|G_n - F| \leq \sum_{i=n}^{m_n} \gamma_i^n \times \mathbb{E}\,|F_i - F| \to 0.$$

So we conclude that $F = G \in \mathcal{H}_l$. Thus, we have

$$\mathbb{E}(|G_n - F|_l^2) + \mathbb{E}(|LG_n - LF|_{l-2}^2) \leq \|G_n - F\|_{L,l,2}^2 \to 0.$$

By passing to a subsequence, we have $|G_n - F|_l + |LG_n - LF|_{l-2} \to 0$ almost surely. Now, for every $\bar{p} \in [1, p)$, we denote $Y_n := |G_n|_l^{\bar{p}} + |LG_n|_{l-2}^{\bar{p}}$ and $Y := |F|_l^{\bar{p}} + |LF|_{l-2}^{\bar{p}}$. Then, $Y_n \to Y$ almost surely, and for any $\tilde{q} \in [\bar{p}, p]$,

$$\begin{aligned} \mathbb{E}|G_n|_l^{\tilde{q}} + \mathbb{E}|LG_n|_{l-2}^{\tilde{q}} &\leq \|G_n\|_{L,l,\tilde{q}}^{\tilde{q}} = \left\| \sum_{i=n}^{m_n} \lambda_i^n \times F_i \right\|_{L,l,\tilde{q}}^{\tilde{q}} \leq (\sum_{i=n}^{m_n} \lambda_i^n \times \|F_i\|_{L,l,\tilde{q}})^{\tilde{q}} \\ &\leq (\sup_i \|F_i\|_{L,l,\tilde{q}} \times \sum_{i=n}^{m_n} \lambda_i^n)^{\tilde{q}} = \sup_i \|F_i\|_{L,l,\tilde{q}}^{\tilde{q}} \leq K_{l,\tilde{q}}^{\tilde{q}}. \end{aligned}$$

So $(Y_n)_{n\in\mathbb{N}}$ is uniformly integrable, and we have

$$\|F\|^{\bar{p}}_{L,l,\bar{p}} = \mathbb{E}(|F|^{\bar{p}}_l) + \mathbb{E}(|LF|^{\bar{p}}_{l-2}) = \mathbb{E}(Y) = \lim_{n\to\infty} \mathbb{E}(Y_n) \leq K^{\bar{p}}_{l,\bar{p}},$$

Proof of (B) We consider for a moment some general $F, G \in \mathcal{D}^d_\infty$. Notice that

$$\langle \sigma(F)\zeta, \zeta\rangle = |\langle DF, \zeta\rangle|^2_{\mathcal{H}},$$

so

$$\lambda(F) = \inf_{|\zeta|=1} |\langle DF, \zeta\rangle|^2_{\mathcal{H}}.$$

Now we check that

$$|\sqrt{\lambda(F)} - \sqrt{\lambda(G)}| \leq |D(F-G)|_{\mathcal{H}}. \tag{65}$$

Indeed, $|\langle DF, \zeta\rangle|_{\mathcal{H}} \leq |\langle DG, \zeta\rangle|_{\mathcal{H}} + |D(F-G)|_{\mathcal{H}}|\zeta|$, so that by taking the infimum, we get $\sqrt{\lambda(F)} \leq \sqrt{\lambda(G)} + |D(F-G)|_{\mathcal{H}}$. And in a similar way, we have the inverse inequality, so (65) is proved. We now come back to our framework. Recalling that $G_n = \sum_{i=n}^{m_n} \gamma^n_i \times F_i$, we observe that

$$\|DG_n - DF_n\|_{L^2(\Omega;\mathcal{H})} \leq \sum_{i=n}^{m_n} \gamma^n_i \|DF_i - DF_n\|_{L^2(\Omega;\mathcal{H})} \to 0.$$

Here we use the fact that $(DF_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega;\mathcal{H})$. Meanwhile, we know from **(A)** that

$$\|DG_n - DF\|_{L^2(\Omega;\mathcal{H})} \to 0.$$

So we conclude that $\|DF - DF_n\|_{L^2(\Omega;\mathcal{H})} \to 0$. Thus, by (65), $\mathbb{E}|\sqrt{\lambda(F)} - \sqrt{\lambda(F_n)}| \to 0$. This gives that there exists a subsequence (also denote by n) such that $\sqrt{\lambda(F_n)}$ converges to $\sqrt{\lambda(F)}$ almost surely, and consequently $|\lambda(F_n)|^{-p}$ converges to $|\lambda(F)|^{-p}$ almost surely. Since we have (63), $(|\lambda(F_n)|^{-p})_{n\in\mathbb{N}}$ is uniformly integrable. It follows that

$$\mathbb{E}(|\lambda(F)|^{-p}) = \lim_{n\to\infty} \mathbb{E}(|\lambda(F_n)|^{-p}) \leq Q_p.$$

Proof of (C) Since the couples $(F, \bar{F})$ and $(F_n, \bar{F}_n)$ verify the hypotheses of **(A)**, we know by the results of **(A)** that we may find a convex combination such that

$$\overline{\lim}_{n\to\infty} \|\sum_{i=n}^{m_n} \gamma^n_i (DF_i, D\bar{F}_i) - (DF, D\bar{F})\|_{L^2(\Omega;\mathcal{H})} = 0.$$

Then it follows by (64) that

$$\begin{aligned}
\|DF - D\bar{F}\|_{L^2(\Omega;\mathcal{H})} &\leq \overline{\lim}_{n\to\infty} \|\sum_{i=n}^{m_n} \gamma^n_i (DF_i - D\bar{F}_i)\|_{L^2(\Omega;\mathcal{H})} \\
&\leq \overline{\lim}_{n\to\infty} \sum_{i=n}^{m_n} \gamma^n_i \|DF_i - D\bar{F}_i\|_{L^2(\Omega;\mathcal{H})} \\
&\leq \bar{\varepsilon}.
\end{aligned}$$

□

3.1.1 Main consequences

We will use the abstract framework presented above for the IbP framework $(\mathcal{D}^d_\infty, D, L)$, with D and L defined in (62). We recall the notations $\|F\|_{L,l,p}$ in (59), $\Sigma_p(F)$ in (56) and σ_F in (55). For any $\eta > 0$, we take $\Upsilon_\eta(x) : (0,\infty) \to \mathbb{R}$ to be a smooth function such that

$$\mathbb{1}_{[\frac{\eta}{2},\infty)} \leq \Upsilon_\eta \leq \mathbb{1}_{[\eta,\infty)}.$$

We remark that σ_F is invertible on the set $\{\Upsilon_\eta(\det \sigma_F) > 0\}$. We first establish an integration by parts formula.

Lemma 3.4. ***(A)*** *Let $F = (F_1, \cdots, F_d) \in \mathcal{D}^d_\infty$. We suppose that the Malliavin covariance matrix σ_F is invertible. We denote*

$$\Gamma_F = (\Gamma_F^{j,i})_{j,i=1,\cdots,d} = \sigma_F^{-1}.$$

We also assume that $\det \sigma_F$ is almost surely invertible and $(\det \sigma_F)^{-1} \in \mathcal{D}_\infty$. Then for every $f \in C^1_p(\mathbb{R}^d)$ and $G \in \mathcal{D}_\infty$,

$$\mathbb{E}(\partial_i f(F) G) = \mathbb{E}(f(F) H_i(F, G)),$$

with

$$H_i(F,G) = \sum_{j=1}^d G(\Gamma_F^{j,i} L F_j - \langle D\Gamma_F^{j,i}, DF_j\rangle_{\mathcal{H}}) - \sum_{j=1}^d \Gamma_F^{j,i} \langle DG, DF_j\rangle_{\mathcal{H}}.$$

Moreover, iterating this relation, for every multi-index β and every $f \in C^{|\beta|}_p(\mathbb{R}^d)$, we get

$$\mathbb{E}(\partial_\beta f(F) G) = \mathbb{E}(f(F) H_\beta(F, G)), \tag{66}$$

where $H_\beta(F,G)$ is obtained by iterations: for $\beta = (\beta_1, \cdots, \beta_m) \in \{1, \cdots, d\}^m$ and $\bar\beta = (\beta_1, \cdots \beta_{m-1})$, we define $H_\beta(F,G) = H_{\beta_m}(F, H_{\bar\beta}(f, G))$.

(B) *Let $F = (F_1, \cdots, F_d) \in \mathcal{D}^d_\infty$. For any $j, i = 1, \cdots, d$ we define*

$$\Gamma^{j,i}_{F,\eta} = (\sigma_F^{-1})^{j,i} \Upsilon_\eta(\det \sigma_F).$$

Then for every $f \in C^1_p(\mathbb{R}^d)$ and $G \in \mathcal{D}_\infty$,

$$\mathbb{E}(\partial_i f(F) G \Upsilon_\eta(\det \sigma_F)) = \mathbb{E}(f(F) H_{\eta,i}(F, G)),$$

with

$$H_{\eta,i}(F,G) = \sum_{j=1}^d G(\Gamma_F^{j,i} L F_j - \langle D\Gamma^{j,i}_{F,\eta}, DF_j\rangle_{\mathcal{H}}) - \sum_{j=1}^d \Gamma^{j,i}_{F,\eta} \langle DG, DF_j\rangle_{\mathcal{H}}.$$

Moreover, iterating this relation, for every multi-index β and every $f \in C^{|\beta|}_p(\mathbb{R}^d)$, we get

$$\mathbb{E}(\partial_\beta f(F) G \Upsilon_\eta(\det \sigma_F)) = \mathbb{E}(f(F) H_{\eta,\beta}(F, G)), \tag{67}$$

where $H_{\eta,\beta}(F,G)$ is obtained by iterations: for $\beta = (\beta_1, \cdots, \beta_m) \in \{1, \cdots, d\}^m$ and $\bar\beta = (\beta_1, \cdots \beta_{m-1})$, we define $H_{\eta,\beta}(F,G) = H_{\eta,\beta_m}(F, H_{\eta,\bar\beta}(f, G))$.

Remark. In **(A)**, we assume the non-degeneracy condition for F, so we have the standard integration by parts formula. However in **(B)**, we do not assume any non-degeneracy condition of F, and we obtain a localized form of integration by parts formula.

Proof. The proof of this lemma is standard, and we refer to [7]. □

As a consequence of the integration by parts formula, we obtain the following proposition based on some estimations of the weights $\mathbb{E}|H_\beta(F,1)|$ and $\mathbb{E}|H_{\eta,\beta}(F,1)|$.

Proposition 3.4.1. *Let $F = (F_1, \cdots, F_d) \in \mathcal{D}^d_\infty$. We fix $q \in \mathbb{N}$.*

***(A)** Suppose that there exists a constant C_q (dependent on q) such that $\|F\|_{L,q+2,8dq} + \Sigma_{4q}(F) \leq C_q$. Then for any multi-index β with $|\beta| = q$ and any function $f \in C^q_b(\mathbb{R}^d)$,*

$$(\boldsymbol{B}_q) \quad |\mathbb{E}(\partial^\beta f(F))| \leq C_q \|f\|_\infty, \quad \forall |\beta| = q.$$

***(B)** Suppose that there exists a constant C'_q (dependent on q) such that $\|F\|_{L,q+2,(4d+1)q} \leq C'_q$. Then for any $\eta > 0$, any multi-index β with $|\beta| = q$ and any function $f \in C^q_b(\mathbb{R}^d)$,*

$$(\boldsymbol{B}'_q) \quad |\mathbb{E}(\partial^\beta f(F) \Upsilon_\eta(\det \sigma_F))| \leq C'_q \|f\|_\infty \times \frac{1}{\eta^{2q}}, \quad \forall |\beta| = q.$$

Remark. In **(A)**, we assume the non-degeneracy condition for F, so we can control the weight H_β in the standard integration by parts formula (66). In **(B)**, we no longer suppose non-degeneracy condition for F, so we apply (67) and obtain a localized form of estimate.

As an immediate application of **Proposition 3.4.1**, we have the regularity of the density.

Corollary 3.4.1. *We fix $p \in \mathbb{N}$. Let $F = (F_1, \cdots, F_d) \in \mathcal{D}^d_\infty$. We assume that $\|F\|_{L,p+d+2,8d(p+d)} + \Sigma_{4(p+d)}(F) \leq \infty$. Then, the law of random variable F is absolutely continuous with respect to the Lebesgue measure and has a density $p_F(x)$ which is $p-$times differentiable. And one has*

$$p_F(x) = \mathbb{E}(\partial^\beta \prod_{j=1}^d \mathbb{1}_{[0,\infty)}(F_j - x_j)), \quad \beta = (1, \cdots, d), \tag{68}$$

with $x = (x_1, \cdots, x_d) \in \mathbb{R}^d$.

Proof. **Proposition 3.4.1** is proved in [7] and **Corollary 3.4.1** follows by standard regularization arguments. □

We consider the $d-$dimensional regularization kernels

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}}, \quad \varphi_\delta(x) = \frac{1}{\delta^d} \varphi(\frac{x}{\delta}), \quad 0 < \delta \leq 1,$$

and we denote

$$f_\delta(x) = f * \varphi_\delta(x) = \int_{\mathbb{R}^d} f(y) \varphi_\delta(x - y) dy.$$

Then we have the following regularization lemma.

Lemma 3.5. ***(A)** $i)$ For a multi index β, we suppose that F satisfies $(\boldsymbol{B}_{2+|\beta|})$. Then for any function $f \in C^{2+|\beta|}_b(\mathbb{R}^d)$,*

$$\left|\mathbb{E}(\partial^\beta f(F)) - \mathbb{E}(\partial^\beta f_\delta(F))\right| \leq d C_{2+|\beta|} \|f\|_\infty \times \delta^2. \tag{69}$$

*$ii)$ **(Romberg)** For a multi-index β, we suppose that F satisfies $(\boldsymbol{B}_{4+|\beta|})$. Then for any function $f \in C^{4+|\beta|}_b(\mathbb{R}^d)$,*

$$\left|\mathbb{E}(\partial^\beta f(F)) + \mathbb{E}(\partial^\beta f_\delta(F)) - 2\mathbb{E}(\partial^\beta f_{\delta/\sqrt{2}}(F))\right| \leq 6d^2 C_{4+|\beta|} \|f\|_\infty \times \delta^4. \tag{70}$$

***(B)** $iii)$ We suppose that F satisfies $(\boldsymbol{B}'_2)$. We fix $\rho > 0$ and we take some $G \in \mathcal{D}^d_\infty$ such that for any $p \in \mathbb{N}$, $\|F\|_{1,p} + \|G\|_{1,p} + \Sigma_\rho(G) < \infty$. For any $\varepsilon_0 > 0$, we denote $q = 2/(1 + \varepsilon_0)$. Then there exists a constant C depending on p, q, ρ and d such that for any $\eta > 0$ and $\delta > 0$, for any function $f \in C^2_b(\mathbb{R}^d)$,*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f_\delta(F))| \leq C \|f\|_\infty \times (\frac{\delta^2}{\eta^4} + (\eta^{-1} \|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q + \eta^\rho). \tag{71}$$

*$iv)$ **(Romberg)** We suppose that F satisfies $(\boldsymbol{B}'_4)$. Under the same hypotheses as $iii)$, for any function $f \in C^4_b(\mathbb{R}^d)$, we have*

$$\left|\mathbb{E}(f(F)) + \mathbb{E}(f_\delta(F)) - 2\mathbb{E}(f_{\delta/\sqrt{2}}(F))\right| \leq C \|f\|_\infty \times (\frac{\delta^4}{\eta^8} + (\eta^{-1} \|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q + \eta^\rho). \tag{72}$$

Remark. We remark that in **(A)**, we assume the non-degeneracy condition for F, and we have the standard regularization lemma (69). While in **(B)**, we do not assume the non-degeneracy condition for F, but we need to assume that we have another random variable G which is non-degenerated (such that DG is close to DF). Then we obtain a variant form of regularization lemma (71). Moreover, applying Romberg method, we have (70) and (72). We also remark that the regularization lemma here is slightly different from the one in [7]. The kernel considered in [7] is the super kernel, but we are not able to simulate the super kernel. So in our paper, we consider the Gaussian kernel φ_δ which allows us to do the simulation.

Proof. Through all this proof we use the notation $g = \partial^\beta f$.

Proof of (A) $i)$ **:** We denote

$$R_q(\delta, x) = \frac{1}{q!} \sum_{|\alpha|=q} \int_0^1 d\lambda (1-\lambda)^q \int_{\mathbb{R}^d} dy \varphi_\delta(y) y^\alpha \partial^\alpha g(x + \lambda y)$$

with $y^\alpha = \prod_{i=1}^q y_{\alpha_i}$ for $\alpha = (\alpha_1, ..., \alpha_q)$. Notice that if F satisfies $(\mathbf{B}_q)$ then (recall that $\partial^\alpha g = \partial^\alpha \partial^\beta f$)

$$|\mathbb{E}(R_q(\delta, F))| \leq C_{q+|\beta|} \|f\|_\infty \int_{\mathbb{R}^d} dy \varphi_\delta(y) |y|^q = C_{q+|\beta|} \int_{\mathbb{R}^d} \varphi(y) |y|^q dy \|f\|_\infty \delta^q. \tag{73}$$

We use a development in Taylor series of order two in order to get

$$\begin{aligned} \partial^\beta f(x) - \partial^\beta f_\delta(x) &= \int_{\mathbb{R}^d} dy \varphi_\delta(x-y)(\partial^\beta f(x) - \partial^\beta f(y)) \\ &= \int_{\mathbb{R}^d} dy \varphi_\delta(x-y)(g(x) - g(y)) \\ &= R_2(\delta, x). \end{aligned}$$

Here we use the fact that $\int_{\mathbb{R}^d} y_j \varphi_\delta(y) dy = 0$. This, together with (73) yields (69).

Proof of (A) $ii)$ **:** Using a development in Taylor series of order 4

$$\partial^\beta f(x) - \partial^\beta f_\delta(x) = \frac{\delta^2}{2} \nabla^2 g(x) + R_4(\delta, x).$$

Here we have used the fact that the third moments of the normal distribution are null and $\int_{\mathbb{R}^d} y_j^2 \varphi_\delta(y) dy = \delta^2$. We fix $a \in (0,1)$ and we use the above equality for $a\delta$:

$$\frac{1}{a^2} \partial^\beta f(x) - \frac{1}{a^2} \partial^\beta f_{a\delta}(x) = \frac{\delta^2}{2} \nabla^2 g(x) + \frac{1}{a^2} R_4(a\delta, x).$$

Subtracting the equality for δ and for $a\delta$, we obtain

$$(\frac{1}{a^2} - 1)\partial^\beta f(x) - (\frac{1}{a^2} \partial^\beta f_{a\delta}(x) - \partial^\beta f_\delta(x)) = \frac{1}{a^2} R_4(a\delta, x) - R_4(\delta, x).$$

Taking $a = 1/\sqrt{2}$ we get

$$\partial^\beta f(x) = 2\partial^\beta f_{\delta/\sqrt{2}}(x) - \partial^\beta f_\delta(x) + 2R_4(\delta/\sqrt{2}, x) - R_4(\delta, x).$$

And using (73) we get (70) (we have also used $\int_{\mathbb{R}^d} \varphi(y) |y|^4 dy \leq 3d^2$).

Proof of (B) $iii)$ **:** We take $|\beta| = 0$. Notice that if F satisfies $(\mathbf{B}'_q)$, then

$$|\mathbb{E}(R_q(\delta, F) \Upsilon_\eta(\det \sigma_F))| \leq C'_q \frac{\|f\|_\infty}{\eta^{2q}} \int_{\mathbb{R}^d} dy \varphi_\delta(y) |y|^q = C'_q \int_{\mathbb{R}^d} \varphi(y) |y|^q dy \|f\|_\infty \frac{\delta^q}{\eta^{2q}}. \tag{74}$$

We use a development in Taylor series of order two in order to get

$$\begin{aligned}\mathbb{E}(f(F)\Upsilon_\eta(\det\sigma_F)) - \mathbb{E}(f_\delta(F)\Upsilon_\eta(\det\sigma_F)) &= \mathbb{E}(\int_{\mathbb{R}^d} dy\varphi_\delta(F-y)(f(F)-f(y))\Upsilon_\eta(\det\sigma_F)) \\ &= \mathbb{E}(R_2(\delta,F)\Upsilon_\eta(\det\sigma_F)).\end{aligned}$$

Here we use the fact that $\int_{\mathbb{R}^d} y_j\varphi_\delta(y)dy = 0$. Using (74) for $q=2$, we have

$$|\mathbb{E}(R_2(\delta,F)\Upsilon_\eta(\det\sigma_F))| \leq C_2' \int_{\mathbb{R}^d} \varphi(y)\,|y|^2\,dy\,\|f\|_\infty \frac{\delta^2}{\eta^4}.$$

So

$$|\mathbb{E}(f(F)\Upsilon_\eta(\det\sigma_F)) - \mathbb{E}(f_\delta(F)\Upsilon_\eta(\det\sigma_F))| \leq C\,\|f\|_\infty \frac{\delta^2}{\eta^4}. \tag{75}$$

On the other hand, we make a small computational trick as follows which is originally from [11] p14. This trick allows us to obtain a better result. We denote

$$R = \frac{\det\sigma_F - \det\sigma_G}{\det\sigma_G}.$$

This is well-defined since G is non-degenerated. For an arbitrary η, we write

$$\mathbb{P}(\det\sigma_F < \eta) \leq \mathbb{P}(\det\sigma_F < \eta, |R| < \frac{1}{4}) + \mathbb{P}(|R| \geq \frac{1}{4}). \tag{76}$$

When $|R| < \frac{1}{4}$, $|\det\sigma_F - \det\sigma_G| < \frac{1}{4}\det\sigma_G$. This implies that $\det\sigma_F > \frac{1}{2}\det\sigma_G$. Recalling that G is non-degenerated and using Markov inequality, for every $\rho \in \mathbb{N}$, it follows that

$$\mathbb{P}(\det\sigma_F < \eta, |R| < \frac{1}{4}) \leq \mathbb{P}(\det\sigma_G < 2\eta) \leq 2^\rho\eta^\rho\mathbb{E}|\det\sigma_G|^{-\rho} \leq C\eta^\rho. \tag{77}$$

For any $\eta > 0$, $\rho \in \mathbb{N}$, with $q = 2/(1+\varepsilon_0)$, we write

$$\begin{aligned}\mathbb{P}(|R| \geq \frac{1}{4}) &= \mathbb{P}(|\det\sigma_F - \det\sigma_G| \geq \frac{1}{4}\det\sigma_G) \\ &\leq \mathbb{P}(\det\sigma_G \leq \eta) + \mathbb{P}(|\det\sigma_F - \det\sigma_G| > \frac{1}{4}\eta) \\ &\leq C(\eta^\rho + \eta^{-q}\mathbb{E}|\det\sigma_F - \det\sigma_G|^q) \\ &\leq C(\eta^\rho + (\eta^{-1}\|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q),\end{aligned} \tag{78}$$

where in the last two steps, we have used the fact that G is non-degenerated, and $\|F\|_{1,p} + \|G\|_{1,p} < \infty$, $\forall p \geq 1$, and Hölder inequality with conjugates $1+\varepsilon_0$ and $\frac{1+\varepsilon_0}{\varepsilon_0}$. Putting together (76), (77) and (78), we obtain

$$\mathbb{P}(\det\sigma_F < \eta) \leq C(\eta^\rho + (\eta^{-1}\|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q). \tag{79}$$

Then we have

$$|\mathbb{E}((1-\Upsilon_\eta(\det\sigma_F))f(F))| \leq \|f\|_\infty\mathbb{P}(\det\sigma_F < \eta) \leq C\|f\|_\infty(\eta^\rho + (\eta^{-1}\|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q). \tag{80}$$

Similarly, we also have

$$|\mathbb{E}((1-\Upsilon_\eta(\det\sigma_F))f_\delta(F))| \leq C\|f\|_\infty(\eta^\rho + (\eta^{-1}\|DF - DG\|_{L^2(\Omega;\mathcal{H})})^q). \tag{81}$$

We conclude by combining (75), (80) and (81).

Proof of (B) $iv)$ **:** The proof is analogous to the proof of $ii)$. Using a development in Taylor series of order 4

$$f(x) - f_\delta(x) = \frac{\delta^2}{2}\nabla^2 f(x) + R_4(\delta, x).$$

We use the above equality for δ and $\frac{\delta}{\sqrt{2}}$, then by subtracting them, we get

$$f(x) = 2f_{\delta/\sqrt{2}}(x) - f_\delta(x) + 2R_4(\delta/\sqrt{2}, x) - R_4(\delta, x).$$

So by (74),

$$\begin{aligned}
&\left|\mathbb{E}(f(F)\Upsilon_\eta(\det\sigma_F)) + \mathbb{E}(f_\delta(F)\Upsilon_\eta(\det\sigma_F)) - 2\mathbb{E}(f_{\delta/\sqrt{2}}(F)\Upsilon_\eta(\det\sigma_F))\right| \\
&\leq 2\mathbb{E}(R_4(\delta/\sqrt{2}, x)\Upsilon_\eta(\det\sigma_F)) - \mathbb{E}(R_4(\delta, x)\Upsilon_\eta(\det\sigma_F)) \\
&\leq C\|f\|_\infty \frac{\delta^4}{\eta^8}. \qquad (82)
\end{aligned}$$

We conclude together with (80) and (81).

□

The regularization lemma (**Lemma 3.5**) implies the following result concerning the approximation of the density function.

Corollary 3.5.1. $i)$ *Suppose that F satisfies $(\boldsymbol{B}_{2+d})$. Then, for every x,*

$$|p_F(x) - \mathbb{E}(\varphi_\delta(F - x))| \leq dC_{2+d} \times \delta^2. \qquad (83)$$

$ii)$ ***(Romberg)*** *Suppose that F satisfies $(\boldsymbol{B}_{4+d})$. Then*

$$\left|p_F(x) + \mathbb{E}(\varphi_\delta(F - x)) - 2\mathbb{E}(\varphi_{\delta/\sqrt{2}}(F - x))\right| \leq 6d^2 C_{4+d} \times \delta^4. \qquad (84)$$

Proof. We take a multi-index $\beta = (1, \cdots, d)$ and

$$f(y) = \prod_{j=1}^{d} H(y_j), \qquad (85)$$

where $H(y) = \mathbb{1}_{[0,\infty)}(y)$ is the Heaviside function. So by (68),

$$p_F(x) = \mathbb{E}(\partial^\beta f(F - x)).$$

Notice that

$$\partial^\beta f_\delta(F - x) = \prod_{j=1}^{d} H'_\delta(F_j - x_j)) = \varphi_\delta(F - x),$$

so that (69) gives

$$\begin{aligned}
|p_F(x) - \mathbb{E}(\varphi_\delta(F - x)| &= \left|\mathbb{E}(\partial^\beta f(F - x)) - \mathbb{E}(\partial^\beta f_\delta(F - x))\right| \\
&\leq dC_{2+d} \times \delta^2.
\end{aligned}$$

In a similar way (70) gives (84). □

In the following, we define the distances between random variables $F, G : \Omega \to \mathbb{R}^d$:

$$d_r(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \sum_{|\beta|=r} \left\|\partial^\beta f\right\|_\infty \leq 1\}$$

For $r = 1$, this is the Wasserstein distance W_1, while for $r = 0$, this is the total variation distance d_{TV}.

Using the Malliavin integration by parts formula (**Lemma 3.4**), one proves in [7] (lemma 3.9) the following results.

Lemma 3.6. *We fix some index l, some $r \in \mathbb{N}$ and some $\varepsilon > 0$. We define $p_1 = 2(r(\frac{1}{\varepsilon} - 1) + 2)$, $p_2 = \max\{4(l+d), 2(\frac{r+l}{\varepsilon} - r + 2)\}$, $q_1 \geq r(\frac{1}{\varepsilon} - 1) + 4$, $q_2 \geq \max\{l + d + 2, \frac{r+l}{\varepsilon} - r + 4\}$. Let $F, G \in \mathcal{D}^d_\infty$. One may find $p \in \mathbb{N}$, $C \in \mathbb{R}_+$ (depending on r, l and ε) such that*

$$i) \quad d_{TV}(F,G) \leq C(1 + \Sigma_{p_1}(F) + \Sigma_{p_1}(G) + \|F\|_{L,q_1,p} + \|G\|_{L,q_1,p}) \times d_r(F,G)^{1-\varepsilon}, \tag{86}$$

and

$$ii) \quad \|p_F - p_G\|_{l,\infty} \leq C(1 + \Sigma_{p_2}(F) + \Sigma_{p_2}(G) + \|F\|_{L,q_2,p} + \|G\|_{L,q_2,p}) \times d_r(F,G)^{1-\varepsilon}, \tag{87}$$

where $p_F(x)$ and $p_G(x)$ denote the density functions of F and G respectively.

Remark. We explain about the significance of this lemma. If we have already obtained an estimate of a "smooth" distance d_r between two random vectors F and G but we would like to control the total variation distance between them, then we employ some integration by parts techniques which are developed in [BCP] and conclude the following. If both F and G are "smooth" in the sense that $\|F\|_{L,q,p} + \|G\|_{L,q,p} < \infty$ for sufficiently large q, p; and both F and G are non-degenerated in the sense that $\Sigma_p(F) + \Sigma_p(G) < \infty$, with p large enough, then (86) asserts that one may control d_{TV} by d_r, and the control is quasi optimal: we loose just a power $\varepsilon > 0$ which we may take as small as we want. And (87) says that we may also control the distance between the derivatives of density functions by d_r.

If we only assume the non-degeneracy condition on F but no non-degeneracy condition for G, then we have a variant of the previous lemma (see [7] proposition 3.11 and remark 3.14).

Proposition 3.6.1. *We fix some $r \in \mathbb{N}$ and some $\varepsilon > 0$. We define $p_1 = 2(\frac{8r}{\varepsilon} + 2)$, $q_1 \geq \frac{8r}{\varepsilon} + 4$. Let $F, G \in \mathcal{D}^d_\infty$. One may find $p \in \mathbb{N}$, $C \in \mathbb{R}_+$ (depending on r and ε) such that*

$$d_{TV}(F,G) \leq C(1 + \Sigma_{p_1}(F) + \|F\|_{L,q_1,p} + \|G\|_{L,q_1,p}) \times (d_r(F,G) + \|DF - DG\|^2_{L^2(\Omega;\mathcal{H})})^{1-\varepsilon}. \tag{88}$$

Remark. The result in **Proposition 3.6.1** is better than proposition 3.11 and remark 3.14 in [7]. We get $\|DF - DG\|^2_{L^2(\Omega;\mathcal{H})}$ instead of $\|DF - DG\|_{L^2(\Omega;\mathcal{H})}$. This is because rather than the estimate (3.29) with $p' = 1$ in [7], we use a sharper estimate (79) with $q = \frac{2}{1+\varepsilon_0}$ and $\varepsilon_0 = \frac{\varepsilon}{2-2\varepsilon}$. The idea of (79) comes from the paper [11] p14. We benefit a lot from this improvement in the paper. It guarantees that we are able to keep the speed of convergence $1 - \varepsilon$ (instead of $\frac{1}{2} - \varepsilon$) in the final results **Theorem 2.1**$\sim$**2.4**.

3.2 Malliavin calculus for the jump equations

In this section, we present the integration by parts framework which will be used when we deal with the jump equations (51), (52) and (40). There are several approaches given in [13], [26], [32], [33], [41], [45] and [52] for example. Here we give a framework analogous to [8].

To begin we define a regularization function.

$$a(y) = 1 - \frac{1}{1 - (4y-1)^2} \quad for \quad y \in [\tfrac{1}{4}, \tfrac{1}{2}), \tag{89}$$

$$\psi(y) = \mathbb{1}_{\{|y| \leq \frac{1}{4}\}} + \mathbb{1}_{\{\frac{1}{4} < |y| \leq \frac{1}{2}\}} e^{a(|y|)}. \tag{90}$$

We notice that $\psi \in C^\infty_b(\mathbb{R})$ and that its support is included in $[-\frac{1}{2}, \frac{1}{2}]$. We denote

$$\Psi_k(y) = \psi(|y| - (k - \tfrac{1}{2})), \ \forall k \in \mathbb{N}. \tag{91}$$

Then for any $l \in \mathbb{N}$, there exists a constant C_l such that

$$\sup_{k \in \mathbb{N}} \|\Psi_k\|_{l,\infty} \leq C_l < \infty. \tag{92}$$

We focus on $x_t^{\mathcal{P},M}$ and x_t^M (solutions of the equations (52) and (51)) which are functions of random variables $T_i^k, W_i^k, Z_i^k, \Delta$ and X_0 (see Section 2.7). Now we introduce the space of simple functionals $\mathcal{S}$. We

take $\mathcal{G}=\sigma(T_i^k,W_i^k,X_0:k,i\in\mathbb{N})$ to be the $\sigma-$algebra associated to the noises which will not be involved in our calculus. In the following, we will do the calculus based on $Z_i^k=(Z_{i,1}^k,\cdots,Z_{i,d}^k),k,i\in\mathbb{N}$ and $\Delta=(\Delta_1,\cdots,\Delta_d)$. We denote by $C_{\mathcal{G},p}$ the space of the functions $f:\Omega\times\mathbb{R}^{m\times m'\times d+d}\to\mathbb{R}$ such that for each ω, the function $(z_{1,1}^1,...,z_{m,d}^{m'},\delta_1,\cdots,\delta_d)\mapsto f(\omega,z_{1,1}^1,...,z_{m,d}^{m'},\delta_1,\cdots,\delta_d)$ belongs to $C_p^\infty(\mathbb{R}^{m\times m'\times d+d})$ (the space of smooth functions which, together with all the derivatives, have polynomial growth), and for each $(z_{1,1}^1,...,z_{m,d}^{m'},\delta_1,\cdots,\delta_d)$, the function $\omega\mapsto f(\omega,z_{1,1}^1,...,z_{m,d}^{m'},\delta_1,\cdots,\delta_d)$ is $\mathcal{G}$-measurable. And we consider the weights

$$\xi_i^k=\Psi_k(Z_i^k).$$

Then we define the space of simple functionals

$$\mathcal{S}=\{F=f(\omega,(Z_i^k)_{\substack{1\le k\le m'\\1\le i\le m}},\Delta):f\in C_{\mathcal{G},p},m,m'\in\mathbb{N}\}.$$

Remark. The simple functional F is actually a function of $(T_i^k)_{k,i\in\mathbb{N}},(W_i^k)_{k,i\in\mathbb{N}},(Z_i^k)_{k,i\in\mathbb{N}},\Delta$ and X_0. By taking $m=J_t^k$ and $m'=M$, we notice that for any $0<t\le T$, $x_t^{\mathcal{P},M}$ and x_t^M (solutions of the equations (52) and (51)) both belong to $\mathcal{S}^d$.

On the space $\mathcal{S}$ we define the derivative operator $DF=(D^ZF,D^\Delta F)$, where

$$\begin{aligned}D^Z_{(\bar k,\bar i,\bar j)}F &= \xi_{\bar i}^{\bar k}\frac{\partial f}{\partial z_{\bar i,\bar j}^{\bar k}}(\omega,(Z_i^k)_{\substack{1\le k\le m'\\1\le i\le m}},\Delta),\quad \bar k,\bar i\in\mathbb{N},\bar j\in\{1,\cdots,d\},\\ D^\Delta_{\tilde j}F &= \frac{\partial f}{\partial\delta_{\tilde j}}(\omega,(Z_i^k)_{\substack{1\le k\le m'\\1\le i\le m}},\Delta),\quad \tilde j\in\{1,\cdots,d\}.\end{aligned}\tag{93}$$

We regard D^ZF as an element of the Hilbert space l_2 (the space of the sequences $h=(h_{k,i,j})_{k,i\in\mathbb{N},j\in\{1,\cdots,d\}}$ with $|h|_{l_2}^2:=\sum_{k=1}^\infty\sum_{i=1}^\infty\sum_{j=1}^d|h_{i,j}|^2<\infty$) and DF as an element of $l_2\times\mathbb{R}^d$, so we have

$$\langle DF,DG\rangle_{l_2\times\mathbb{R}^d}=\sum_{j=1}^d D_j^\Delta F\times D_j^\Delta G+\sum_{k=1}^\infty\sum_{i=1}^\infty\sum_{j=1}^d D^Z_{(k,i,j)}F\times D^Z_{(k,i,j)}G.\tag{94}$$

We also denote $D^1F=DF$, and we define the derivatives of order $q\in\mathbb{N}$ recursively: $D^qF:=DD^{q-1}F$. And we denote $D^{Z,q}$ (respectively $D^{\Delta,q}$) as the derivative D^Z (respectively D^Δ) of order q.

We also define the Ornstein-Uhlenbeck operator $LF=L^ZF+L^\Delta F$ with

$$\begin{aligned}L^ZF &= -\sum_{k=1}^{m'}\sum_{i=1}^m\sum_{j=1}^d(\partial_{z_{i,j}^k}(\xi_i^kD^Z_{(k,i,j)}F)+D^Z_{(k,i,j)}F\times D^Z_{(k,i,j)}\ln[h(Z_i^k)]),\\ L^\Delta F &= \sum_{j=1}^d D_j^\Delta F\times\Delta_j-\sum_{j=1}^d D_j^\Delta D_j^\Delta F.\end{aligned}\tag{95}$$

One can check that the triplet $(\mathcal{S},D,L)$ is consistent with the IbP framework given in Section 3.1. In particular the duality formula (57) holds true. We refer to [10](Appendix 5.3). We say that F is a "Malliavin smooth functional" if $F\in\mathcal{D}_\infty$ (with the definition given in (61)).

We will use the IbP framework defined here for x_t, x_t^M and $x_t^{\mathcal{P},M}$ (solutions of equations (40),(41) and (42)). We recall that they are obtained in Section 2.7 by optimal coupling in $W_{2+\varepsilon_*}$ distance between $X_{\tau(t)-}^{\mathcal{P},M}$ and X_{t-}. Here we give two lemmas, concerning the Sobolev norms and the covariance matrices of x_t, x_t^M and $x_t^{\mathcal{P},M}$.

Lemma 3.7. *Assuming **Hypothesis 2.1** and **Hypothesis 2.4** b), for all $p\ge1,l\ge0$, there exists a constant $C_{l,p}$ depending on l,p,d and T, such that for any $0<t\le T$,*

$$i)\quad\sup_{\mathcal{P}}\sup_M(\|x_t^{\mathcal{P},M}\|_{L,l,p}+\|x_t^M\|_{L,l,p})\le C_{l,p}.$$

Moreover, x_t belongs to $\mathcal{D}_\infty^d$ and

$$ii)\quad\|x_t\|_{L,l,p}\le C_{l,p}.$$

Lemma 3.8. *Assume that **Hypothesis 2.1, 2.2, 2.3** and **2.4** hold true. Then for every $p \geq 1$, $0 < t \leq T$ such that $t > \frac{2dp}{\theta}$ (with θ defined in (18)), we have (recalling by (55) that σ_F denotes the covariance matrix of F)*

$$i) \quad \sup_M \mathbb{E}(1/\det \sigma_{x_t^M})^p \leq C_p, \tag{96}$$

$$ii) \quad \mathbb{E}(1/\det \sigma_{x_t})^p \leq C_p, \tag{97}$$

with C_p a constant depending on p, d, T.

Remark. In the case $\theta = \infty$, the results in **Lemma 3.8** hold for every $0 < t \leq T$.

The proofs of these two lemmas are rather technical and are postponed to Section 4.1 and 4.2.

Before we end this section, we establish an auxiliary result. We recall by (27) that $\varepsilon_M = \int_{\{|z|>M\}} |\bar{c}(z)|^2 \mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z)\mu(dz)|^2)$.

Lemma 3.9. *We assume that **Hypothesis 2.1** and **Hypothesis 2.4** $b)$ hold true. Then for any $\varepsilon_* > 0$, there exists a constant C dependent on T, d, ε_* such that for every $|\mathcal{P}| \leq 1$, every M with $\varepsilon_M \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$,*

$$i) \quad \|Dx_t^{\mathcal{P},M} - Dx_t^M\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \leq C(\varepsilon_M + |\mathcal{P}|)^{\frac{1}{2+\varepsilon_*}},$$

$$ii) \quad \|Dx_t^M - Dx_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \leq C(\varepsilon_M)^{\frac{1}{2+\varepsilon_*}},$$

$$iii) \quad \|Dx_t^{\mathcal{P},M} - Dx_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \leq C(\varepsilon_M + |\mathcal{P}|)^{\frac{1}{2+\varepsilon_*}}.$$

The proof is also technical and we put it in the Appendix.

3.3 Proofs of Theorem 2.1∼2.4

Before the proofs of **Theorem 2.1∼2.4**, we first give the following lemma. We recall $X_t^{\mathcal{P},M}$ in (25) and X_t in (14).

Lemma 3.10. *Assume that the **Hypothesis 2.1** holds true. Then there exists a constant C dependent on T such that for every partition $\mathcal{P}$ and $M \in \mathbb{N}$ we have*

$$W_1(X_t^{\mathcal{P},M}, X_t) \leq C(|\mathcal{P}| + \sqrt{\varepsilon_M}).$$

Proof. We make a coupling argument similar to Section 2.7. We will do optimal coupling between $X_{\tau(t)-}^{\mathcal{P},M}$ and X_{t-} in W_1 distance. This is the same strategy as the optimal coupling between $X_{\tau(t)-}^{\mathcal{P},M}$ and X_{t-} in $W_{2+\varepsilon_*}$ distance in Section 2.7. We take $\widetilde{\Pi}_t^{\mathcal{P},M}(dv_5, dv_6)$ to be the optimal W_1-coupling of $\rho_{\tau(t)-}^{\mathcal{P},M}(dv_5)$ and $\rho_{t-}(dv_6)$, that is

$$W_1(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t-}) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |v_5 - v_6| \widetilde{\Pi}_t^{\mathcal{P},M}(dv_5, dv_6).$$

Then we construct $(\eta_t^5(w), \eta_t^6(w))$ which represents $\widetilde{\Pi}_t^{\mathcal{P},M}$ in the sense of **Lemma 2.5**. So we have

$$\int_0^1 \phi(\eta_t^5(w), \eta_t^6(w))dw = \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(v_5, v_6) \widetilde{\Pi}_t^{\mathcal{P},M}(dv_5, dv_6).$$

We consider the equations (with $\mathcal{N}(dw, dz, dr)$ the Poisson point measure on the state space $[0,1] \times \mathbb{R}^d$ with intensity measure $dw\mu(dz)dr$ defined in Section 2.7):

$$\widetilde{x}_t = X_0 + \int_0^t b(r, \widetilde{x}_r, \rho_r)dr + \int_0^t \int_{[0,1]\times\mathbb{R}^d} c(r, \eta_r^6(w), z, \widetilde{x}_{r-}, \rho_{r-})\mathcal{N}(dw, dz, dr), \tag{98}$$

$$\begin{aligned} \widetilde{x}_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), \widetilde{x}_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})dr \\ &+ \int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(\tau(r), \eta_r^5(w), z, \widetilde{x}_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M})\mathcal{N}(dw, dz, dr), \end{aligned} \tag{99}$$

with $\rho_t^{\mathcal{P},M}$ the law of $X_t^{\mathcal{P},M}$ (see (25)) and ρ_t the law of X_t (see (14)). One can check that $\widetilde{x}_t$ and $\widetilde{x}_t^{\mathcal{P},M}$ have the same law as x_t and $X_t^{\mathcal{P},M}$ respectively. We remark that $\widetilde{x}_t$, and $\widetilde{x}_t^{\mathcal{P},M}$ are different from x_t, and $x_t^{\mathcal{P},M}$ (see (40) and (42)) since we take different couplings and $\eta_r^1(w) \neq \eta_r^5(w)$, $\eta_r^2(w) \neq \eta_r^6(w)$. Then we have

$$W_1(X_t^{\mathcal{P},M}, X_t) = W_1(\widetilde{x}_t^{\mathcal{P},M}, \widetilde{x}_t) \leq \mathbb{E}|\widetilde{x}_t^{\mathcal{P},M} - \widetilde{x}_t| \leq C(|\mathcal{P}| + \sqrt{\varepsilon_M}),$$

where the last inequality is obtained in a standard way (see the proof of **Lemma 2.6**). □

Proofs of Theorem 2.1 and Theorem 2.2:

Proof. We first prove (28). We recall that by the discussion in Section 2.7, $x_t^{\mathcal{P},M}$ has the same law as $X_t^{\mathcal{P},M}$ and x_t has the same law as X_t. Thanks to **Lemma 3.7** and **Lemma 3.8**, using **Proposition 3.6.1**, for any partition $\mathcal{P}$ of the interval $[0,T]$ with $|\mathcal{P}| \leq 1$, every $M \in \mathbb{N}$ with $\varepsilon_M \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$, for $\varepsilon > 0$, when $t > \frac{8d}{\theta}(\frac{4}{\varepsilon} + 1)$ (with θ defined in (18)),

$$\begin{aligned} d_{TV}(X_t^{\mathcal{P},M}, X_t) &= d_{TV}(x_t^{\mathcal{P},M}, x_t) \\ &\leq C[W_1(x_t^{\mathcal{P},M}, x_t) + \|Dx_t^{\mathcal{P},M} - Dx_t\|^2_{L^2(\Omega; l_2 \times \mathbb{R}^d)}]^{1-\varepsilon}. \end{aligned}$$

For any $\bar{\varepsilon} > 0$, we take $\varepsilon, \varepsilon_* > 0$ such that $\varepsilon_* = \frac{\bar{\varepsilon}}{1-\bar{\varepsilon}}$ and $\varepsilon = \frac{\bar{\varepsilon}}{2}$. So $\frac{2}{2+\varepsilon_*}(1-\varepsilon) = 1 - \bar{\varepsilon}$. Then by **Lemma 3.9** and **Lemma 3.10**, when $t > \frac{8d}{\theta}(\frac{8}{\bar{\varepsilon}} + 1)$, we have

$$\begin{aligned} d_{TV}(X_t^{\mathcal{P},M}, X_t) &\leq C[W_1(X_t^{\mathcal{P},M}, X_t) + \|Dx_t^{\mathcal{P},M} - Dx_t\|^2_{L^2(\Omega; l_2 \times \mathbb{R}^d)}]^{1-\varepsilon} \\ &\leq C[|\mathcal{P}| + \sqrt{\varepsilon_M} + (\varepsilon_M + |\mathcal{P}|)^{\frac{2}{2+\varepsilon_*}}]^{1-\varepsilon} \\ &\leq C[\sqrt{\varepsilon_M} + |\mathcal{P}|]^{1-\bar{\varepsilon}} \to 0, \end{aligned}$$

with C a constant depending on $\bar{\varepsilon}, d$ and T. So (28) is proved.

On the other hand, by **Lemma 3.8** and **Corollary 3.4.1**, when $t > \frac{8d(l+d)}{\theta}$, the law of X_t has a $l-$times differentiable density p_t and the density p_t is a function solution of the equation (21). So (22) is proved. We notice that $(\mathcal{S}, D^\Delta, L^\Delta)$ is also an IbP framework. If we only make Malliavin integration by parts on the Gaussian random variable Δ, then standard arguments give that the law of $X_t^{\mathcal{P},M}$ has a smooth density $p_t^{\mathcal{P},M}$.

Now only (23) is left to be proved. The proof is analogous to the proof of (28). The main strategy is as follows (this is similar to Section 2.7 and Section 3.2). We define an intermediate equation $\bar{X}_t^{\mathcal{P},M}$ (see (100) in the following). There is a difficulty appears here: the equations (14) and (19); (38) and (100) are defined with respect to different Poisson point measures (on different probability spaces). To overcome this difficulty, it is convenient to use similar equations driven by the same Poisson point measure. We make a coupling argument to construct $\bar{x}_t^M$, $\bar{x}_t^{\mathcal{P}}$, $\bar{x}_t^{\mathcal{P},M}$ and $\bar{x}_t$ (see (103), (102), (104) and (101) below) which have the same law as X_t^M, $X_t^{\mathcal{P}}$, $\bar{X}_t^{\mathcal{P},M}$ and X_t (see (38), (19), (100) and (14)) respectively but are defined on the same probability space and verify equations driven by the same Poisson point measure. So to estimate the total variation distance between $X_t^{\mathcal{P}}$ and X_t, it is equivalent to estimate the total variation distance between $\bar{x}_t^{\mathcal{P}}$ and $\bar{x}_t$. We will see that $\bar{x}_t^M$ and $\bar{x}_t^{\mathcal{P},M}$ are simple functionals (belong to $\mathcal{S}^d$) in the sense of Section 3.2. We prove below in **Lemma 3.12** that the Malliavin-Sobolev norms of $\bar{x}_t^M$ and $\bar{x}_t^{\mathcal{P},M}$ are bounded (uniformly in $M, \mathcal{P}$) and that the Malliavin covariance matrix of $\bar{x}_t^M$ is non-degenerate (uniformly in M). Passing to the limit $M \to \infty$, we give below in **Lemma 3.11** that $\bar{x}_t^{\mathcal{P},M} \to \bar{x}_t^{\mathcal{P}}$ and $\bar{x}_t^M \to \bar{x}_t$ in L^1 distance. Then by using **Lemma 3.3**, $\bar{x}_t$ and $\bar{x}_t^{\mathcal{P}}$ are "Malliavin smooth functionals" (belong to $\mathcal{D}_\infty^d$), and we prove below in **Lemma 3.13** that the Malliavin-Sobolev norms of $\bar{x}_t$ and $\bar{x}_t^{\mathcal{P}}$ are bounded (uniformly in $\mathcal{P}$) and that the Malliavin covariance matrix of $\bar{x}_t$ is non-degenerate. So applying **Proposition 3.6.1**, the Euler scheme $X_t^{\mathcal{P}}$ converges to X_t in total variation distance.

Now we give the proof of (23). We first introduce an intermediate equation.

$$\begin{aligned}\bar{X}_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), \bar{X}_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P}}) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, \bar{X}_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P}}) N_{\rho_{\tau(r)-}^{\mathcal{P}}}(dv, dz, dr). \end{aligned} \tag{100}$$

We notice that we take $\rho_{\tau(r)}^{\mathcal{P}}$ (the law of $X_{\tau(r)}^{\mathcal{P}}$) instead of $\rho_{\tau(r)}^{\mathcal{P},M}$ (the law of $X_{\tau(r)}^{\mathcal{P},M}$) in the above equation, so (100) a variant of (25).

Now we make a coupling argument similar to Section 2.7. We will do optimal coupling between $X_{\tau(t)-}^{\mathcal{P}}$ and X_{t-} in $W_{2+\varepsilon_*}$ distance. This is the same strategy as the optimal coupling between $X_{\tau(t)-}^{\mathcal{P},M}$ and X_{t-} in $W_{2+\varepsilon_*}$ distance in Section 2.7. For a small $\varepsilon_* > 0$, we take $\Pi_t^{\mathcal{P}}(dv_3, dv_4)$ to be the optimal $W_{2+\varepsilon_*}-$coupling of $\rho_{\tau(t)-}^{\mathcal{P}}(dv_3)$ and $\rho_{t-}(dv_4)$, that is

$$(W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P}}, \rho_{t-}))^{2+\varepsilon_*} = \int_{\mathbb{R}^d \times \mathbb{R}^d} |v_3 - v_4|^{2+\varepsilon_*} \Pi_t^{\mathcal{P}}(dv_3, dv_4).$$

Then we construct $(\eta_t^3(w), \eta_t^4(w))$ which represents $\Pi_t^{\mathcal{P}}$ in the sense of **Lemma 2.5**. So we have

$$\int_0^1 \phi(\eta_t^3(w), \eta_t^4(w)) dw = \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(v_3, v_4) \Pi_t^{\mathcal{P}}(dv_3, dv_4).$$

We consider some auxiliary equations (with $\mathcal{N}(dw, dz, dr)$ the Poisson point measure on the state space $[0,1] \times \mathbb{R}^d$ with intensity measure $dw\mu(dz)dr$ defined in Section 2.7):

$$\bar{x}_t = X_0 + \int_0^t b(r, \bar{x}_r, \rho_r) dr + \int_0^t \int_{[0,1]\times\mathbb{R}^d} c(r, \eta_r^4(w), z, \bar{x}_{r-}, \rho_{r-}) \mathcal{N}(dw, dz, dr), \tag{101}$$

$$\begin{aligned}\bar{x}_t^{\mathcal{P}} &= X_0 + \int_0^t b(\tau(r), \bar{x}_{\tau(r)}^{\mathcal{P}}, \rho_{\tau(r)}^{\mathcal{P}}) dr \\ &+ \int_0^t \int_{[0,1]\times\mathbb{R}^d} c(\tau(r), \eta_r^3(w), z, \bar{x}_{\tau(r)-}^{\mathcal{P}}, \rho_{\tau(r)-}^{\mathcal{P}}) \mathcal{N}(dw, dz, dr). \end{aligned} \tag{102}$$

$$\bar{x}_t^M = X_0 + a_T^M \Delta + \int_0^t b(r, \bar{x}_r^M, \rho_r) dr + \int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(r, \eta_r^4(w), z, \bar{x}_{r-}^M, \rho_{r-}) \mathcal{N}(dw, dz, dr), \tag{103}$$

$$\begin{aligned}\bar{x}_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), \bar{x}_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P}}) dr \\ &+ \int_0^t \int_{[0,1]\times\mathbb{R}^d} c_M(\tau(r), \eta_r^3(w), z, \bar{x}_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P}}) \mathcal{N}(dw, dz, dr). \end{aligned} \tag{104}$$

One can check that $\bar{x}_t$, $\bar{x}_t^{\mathcal{P}}$, $\bar{x}_t^M$, and $\bar{x}_t^{\mathcal{P},M}$ have the same law as X_t, $X_t^{\mathcal{P}}$, X_t^M, and $\bar{X}_t^{\mathcal{P},M}$ (solutions of the equations (14), (19), (38) and (100)) respectively. We stress that $\bar{x}_t$, $\bar{x}_t^M$, and $\bar{x}_t^{\mathcal{P},M}$ are different from x_t, x_t^M, and $x_t^{\mathcal{P},M}$ (see (40), (41) and (42)). This is because we take different couplings so $\eta_r^1(w) \neq \eta_r^3(w)$ and $\eta_r^2(w) \neq \eta_r^4(w)$. We also remark that we take ρ_r instead of ρ_r^M in (103) and take $\rho_t^{\mathcal{P}}$ instead of $\rho_t^{\mathcal{P},M}$ in (104), so that we can obtain the following lemma.

Lemma 3.11. *Assume that the **Hypothesis 2.1** holds true. Then*

$$i) \quad \sup_{\mathcal{P}} \mathbb{E}|\bar{x}_t^{\mathcal{P},M} - \bar{x}_t^{\mathcal{P}}| \to 0,$$

$$ii) \quad \mathbb{E}|\bar{x}_t^M - \bar{x}_t| \to 0,$$

as $M \to \infty$.

Proof. These results are obtained in a standard way (see the proof of **Lemma 2.6**). □

We notice that $\bar{x}_t^{\mathcal{P},M}$ and $\bar{x}_t^M$ are simple functionals (belong to $\mathcal{S}^d$) in the sense of Section 3.2. Then we have

Lemma 3.12. *Assume that **Hypothesis 2.1, 2.2, 2.3** and **2.4** hold true.*
a) For any $p \geq 1, l \geq 0$, there exists a constant $C_{l,p}$ depending on l, p, d and T such that for every $0 < t \leq T$,

$$\sup_{\mathcal{P}} \sup_{M} (\|\bar{x}_t^{\mathcal{P},M}\|_{L,l,p} + \|\bar{x}_t^M\|_{L,l,p}) \leq C_{l,p}.$$

b) For any $p \geq 1, 0 < t \leq T$ such that $t > \frac{2dp}{\theta}$, there exists a constant C_p depending on p, d, T such that

$$\sup_{M} (\mathbb{E}(1/\det \sigma_{\bar{x}_t^M})^p) \leq C_p.$$

c) For any $\varepsilon_* > 0$, there exists a constant C_p depending on ε_*, d, T such that for every $|\mathcal{P}| \leq 1$, we have

$$i) \quad \|D\bar{x}_t^{\mathcal{P},M} - D\bar{x}_t^M\|_{L^2(\Omega;l_2\times\mathbb{R}^d)} \leq C|\mathcal{P}|^{\frac{1}{2+\varepsilon_*}},$$

$$ii) \quad \|D\bar{x}_t^M - D\bar{x}_t\|_{L^2(\Omega;l_2\times\mathbb{R}^d)} \to 0, \quad \text{as } M \to \infty.$$

Proof. We get $a)$ by an analogous argument to the proof of **Lemma 3.7** $i)$. We have $b)$ in a similar way to the proof of **Lemma 3.8** $i)$. we obtain $c)$ $i)$ and $ii)$ by some analogous arguments to the proofs of **Lemma 3.9** $i)$ and $ii)$ respectively. □

Then applying **Lemma 3.3**, by passing to the limit $M \to \infty$, we obtain the following consequence.

Lemma 3.13. *Assume that **Hypothesis 2.1, 2.2, 2.3** and **2.4** hold true.*
a) $\bar{x}_t^{\mathcal{P}}$ and $\bar{x}_t$ both belong to $\mathcal{D}_\infty^d$. For any $p \geq 1, l \geq 0$, there exists a constant $C_{l,p}$ depending on l, p, d and T such that for every $0 < t \leq T$,

$$\sup_{\mathcal{P}} (\|\bar{x}_t^{\mathcal{P}}\|_{L,l,p} + \|\bar{x}_t\|_{L,l,p}) \leq C_{l,p}.$$

b) For any $p \geq 1, 0 < t \leq T$ such that $t > \frac{2dp}{\theta}$, there exists a constant C_p depending on p, d, T such that

$$\mathbb{E}(1/\det \sigma_{\bar{x}_t})^p \leq C_p.$$

c) For any $\varepsilon_* > 0$, there exists a constant C_p depending on ε_*, d, T such that for every $|\mathcal{P}| \leq 1$, we have

$$\|D\bar{x}_t^{\mathcal{P}} - D\bar{x}_t\|_{L^2(\Omega;l_2\times\mathbb{R}^d)} \leq C|\mathcal{P}|^{\frac{1}{2+\varepsilon_*}}.$$

Proof. **Proof of** $a)$: We apply **Lemma 3.3 (A)** with $F_M = (\bar{x}_t^{\mathcal{P},M}, \bar{x}_t^M)$ and $F = (\bar{x}_t^{\mathcal{P}}, \bar{x}_t)$. By **Lemma 3.11** $i)$, $ii)$ and **Lemma 3.12** $a)$, we obtain our results.

Proof of $b)$: We apply **Lemma 3.3 (B)** with $F_M = \bar{x}_t^M$ and $F = \bar{x}_t$. By **Lemma 3.12** $b)$ and **Lemma 3.12** c $ii)$, it follows that $\mathbb{E}(1/\det \sigma_{\bar{x}_t})^p \leq C_p$.

Proof of $c)$: We apply **Lemma 3.3 (C)** with $(\bar{F}_M, F_M) = (\bar{x}_t^{\mathcal{P},M}, \bar{x}_t^M)$ and $(\bar{F}, F) = (\bar{x}_t^{\mathcal{P}}, \bar{x}_t)$. By **Lemma 3.12** $c)$ $i)$, we have $\|D\bar{x}_t^{\mathcal{P}} - D\bar{x}_t\|_{L^2(\Omega;l_2\times\mathbb{R}^d)} \leq C|\mathcal{P}|^{\frac{1}{2+\varepsilon_*}}$. □

Finally, we can give the proof of (23). We recall that $\bar{x}_t$ and $\bar{x}_t^{\mathcal{P}}$ have the same law as X_t and $X_t^{\mathcal{P}}$ respectively. For any $\bar{\varepsilon} > 0$, we take $\varepsilon, \varepsilon_* > 0$ such that $\varepsilon_* = \frac{\bar{\varepsilon}}{1-\bar{\varepsilon}}$ and $\varepsilon = \frac{\bar{\varepsilon}}{2}$. So $\frac{2}{2+\varepsilon_*}(1-\varepsilon) = 1 - \bar{\varepsilon}$. Thanks to **Lemma 3.13** $a), b)$, using **Proposition 3.6.1**, there exists a constant C dependent on $\bar{\varepsilon}, d, T$ such that for any partition $\mathcal{P}$ of the interval $[0, T]$ with $|\mathcal{P}| \leq 1$, when $t > \frac{8d}{\theta}(\frac{8}{\bar{\varepsilon}} + 1)$, we have

$$\begin{aligned}
d_{TV}(X_t^{\mathcal{P}}, X_t) &= d_{TV}(\bar{x}_t^{\mathcal{P}}, \bar{x}_t) \\
&\leq C[W_1(\bar{x}_t^{\mathcal{P}}, \bar{x}_t) + \|D\bar{x}_t^{\mathcal{P}} - D\bar{x}_t\|^2_{L^2(\Omega;l_2\times\mathbb{R}^d)}]^{1-\varepsilon} \\
&= C[W_1(X_t^{\mathcal{P}}, X_t) + \|D\bar{x}_t^{\mathcal{P}} - D\bar{x}_t\|^2_{L^2(\Omega;l_2\times\mathbb{R}^d)}]^{1-\varepsilon} \\
&\leq C[|\mathcal{P}| + |\mathcal{P}|^{\frac{2}{2+\varepsilon_*}}]^{1-\varepsilon} \\
&\leq C|\mathcal{P}|^{1-\bar{\varepsilon}} \to 0,
\end{aligned}$$

where the second last inequality is obtained by **Lemma 3.13** $c)$ and (20). So (23) is proved. □

Proofs of Theorem 2.3 and Theorem 2.4:

Proof. **Proof of Theorem 2.3** i)**:** We recall in Section 2.7 that x_t (solution of (40)) has the same law as X_t (solution of (14)) and by **Theorem 2.1** a), $\mathcal{L}(x_t)(dx) = \mathcal{L}(X_t)(dx) = p_t(x)dx$. When $t > \frac{8d}{\theta}(2+d)$, **Lemma 3.7** ii) and **Lemma 3.8** ii) give that $\|x_t\|_{L,d+4,8d(2+d)} + \Sigma_{4(2+d)}(x_t) < \infty$ (with the notation $\Sigma_p(F)$ given in (56)). Then we apply **Corollary 3.5.1** i) and obtain that

$$|p_t(x) - \mathbb{E}(\varphi_\delta(X_t - x))| = |p_t(x) - \mathbb{E}(\varphi_\delta(x_t - x))| \leq C\delta^2, \tag{105}$$

where C is a constant dependent on d.

We recall by (12) the definition of the Wasserstein distance of order 1. Noticing $\|\nabla\varphi_\delta\|_\infty \leq \frac{1}{\delta^{d+1}}$, we get

$$|\mathbb{E}(\varphi_\delta(X_t - x)) - \mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M} - x))| \leq W_1(X_t^{\mathcal{P},M}, X_t)\frac{1}{\delta^{d+1}}.$$

So together with **Lemma 3.10**, there exists a constant C dependent on d and T such that

$$|\mathbb{E}(\varphi_\delta(X_t - x)) - \mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M} - x))| \leq C(|\mathcal{P}| + \sqrt{\varepsilon_M})\frac{1}{\delta^{d+1}}. \tag{106}$$

Finally, applying the estimate (4.6) in Theorem 4.1 of [3] with $X_n^i = X_t^{\mathcal{P},M,i}$, $\Theta_{s,s_n}^n(\rho)(dx) = \rho_t^{\mathcal{P},M}(dx)$ and $f(x) = \varphi_\delta(x)$, we get

$$|\mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M} - x)) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M,i} - x))| \leq CV_N\frac{1}{\delta^{d+1}}. \tag{107}$$

Combining (105), (106) and (107),

$$|p_t(x) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M,i} - x))| \leq C[(|\mathcal{P}| + \sqrt{\varepsilon_M})\frac{1}{\delta^{d+1}} + V_N\frac{1}{\delta^{d+1}} + \delta^2].$$

Then we optimize over δ and N. We choose

$$\delta = (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{d+3}}$$

such that

$$(|\mathcal{P}| + \sqrt{\varepsilon_M})\frac{1}{\delta^{d+1}} = \delta^2.$$

So

$$|p_t(x) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}(\varphi_\delta(X_t^{\mathcal{P},M,i} - x))| \leq C[(|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{2}{d+3}} + V_N(|\mathcal{P}| + \sqrt{\varepsilon_M})^{-\frac{d+1}{d+3}}].$$

And we choose N such that

$$V_N \leq |\mathcal{P}| + \sqrt{\varepsilon_M},$$

so

$$(|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{2}{d+3}} \geq V_N(|\mathcal{P}| + \sqrt{\varepsilon_M})^{-\frac{d+1}{d+3}}.$$

Hence, eventually we have (34).

Proof of Theorem 2.3 ii)**:** (35) is obtained in a similar way by using **Corollary 3.5.1** ii).

Proof of Theorem 2.4 i)**:** We take $f \in C_b^\infty(\mathbb{R}^d)$.

Step 1: We recall in Section 2.7 that $x_t^{\mathcal{P},M}$ (solution of (42)) has the same law as $X_t^{\mathcal{P},M}$ (solution of (25)) and x_t (solution of (40)) has the same law as X_t (solution of (14)). We notice by **Theorem 2.2** that for any $\varepsilon > 0$, there exists a constant C dependent on ε, d, T such that when $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$,

$$\Big|\int_{\mathbb{R}^d} f(x)\rho_t(dx) - \int_{\mathbb{R}^d} f(x)\rho_t^{\mathcal{P},M}(dx)\Big| \leq C\,\|f\|_\infty \times (|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}, \tag{108}$$

with ρ_t the law of x_t (also of X_t) and $\rho_t^{\mathcal{P},M}$ the law of $x_t^{\mathcal{P},M}$ (also of $X_t^{\mathcal{P},M}$).

Step 2: We apply the regularization lemma **Lemma 3.5 (B)** $iii)$ with $F = x_t^{\mathcal{P},M}$ and $G = x_t$. For any $\bar{\varepsilon} > 0$, we take $\varepsilon, \varepsilon_*, \varepsilon_0 > 0$ such that

$$\varepsilon_0 = \frac{\varepsilon}{2-2\varepsilon}, \quad \text{and} \quad \varepsilon_* = \frac{\bar{\varepsilon}}{1-\bar{\varepsilon}}, \quad \varepsilon = \frac{\bar{\varepsilon}}{2}.$$

So

$$\frac{2}{2+\varepsilon_*}(1-\varepsilon) = 1-\bar{\varepsilon}.$$

Recalling in **Lemma 3.5 (B)** $iii)$ that $q = \frac{2}{1+\varepsilon_0} = \frac{4(1-\varepsilon)}{2-\varepsilon}$, we have

$$\Big| \int_{\mathbb{R}^d} f(x)\rho_t^{\mathcal{P},M}(dx) - \int_{\mathbb{R}^d} f_\delta(x)\rho_t^{\mathcal{P},M}(dx) \Big| \leq C \left\| f \right\|_\infty \times (\frac{\delta^2}{\eta^4} + \eta^{-q}(|\mathcal{P}| + \varepsilon_M)^{\frac{q}{2+\varepsilon_*}} + \eta^\rho). \tag{109}$$

Here we have used the non-degenerated condition of x_t and the fact that the Sobolev norms of $x_t^{\mathcal{P},M}$ and x_t are bounded (uniformly in $\mathcal{P}, M$). We have also taken advantage of **Lemma 3.9** $iii)$.

Then we optimize over δ, η and ρ. In order to keep the notations clear, we denote temporary that

$$\mathcal{E} = (|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}.$$

We take

$$\eta = \mathcal{E}^{\frac{\varepsilon}{4}} \quad \text{and} \quad \delta = \sqrt{\mathcal{E}}$$

such that

$$\frac{\delta^2}{\eta^4} = \eta^{-q}\mathcal{E}^{\frac{q}{2}} = \mathcal{E}^{1-\varepsilon}.$$

We take moreover

$$\rho = \frac{4(1-\varepsilon)}{\varepsilon}$$

such that

$$\eta^\rho = \mathcal{E}^{1-\varepsilon}.$$

So (109) becomes

$$\Big| \int_{\mathbb{R}^d} f(x)\rho_t^{\mathcal{P},M}(dx) - \int_{\mathbb{R}^d} f_\delta(x)\rho_t^{\mathcal{P},M}(dx) \Big| \leq C \left\| f \right\|_\infty \times \mathcal{E}^{1-\varepsilon} = C \left\| f \right\|_\infty \times (|\mathcal{P}| + \varepsilon_M)^{1-\bar{\varepsilon}}, \tag{110}$$

with C a constant depending on $\bar{\varepsilon}, d, T$ and

$$\delta = (|\mathcal{P}| + \varepsilon_M)^{\frac{1-\bar{\varepsilon}}{2-\bar{\varepsilon}}} = (|\mathcal{P}| + \varepsilon_M)^{\frac{1}{2}(1-\frac{\bar{\varepsilon}}{2-\bar{\varepsilon}})}. \tag{111}$$

Step 3: We apply the estimate (4.6) in Theorem 4.1 of [3]. We notice that $\|\nabla f_\delta\|_\infty \leq C\|f\|_\infty \times \frac{1}{\delta^{d+1}}$. Then we obtain

$$\Big| \int_{\mathbb{R}^d} f_\delta(x)\rho_t^{\mathcal{P},M}(dx) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}(f_\delta(X_t^{\mathcal{P},M,i})) \Big| \leq C\|f\|_\infty V_N \times \frac{1}{\delta^{d+1}}. \tag{112}$$

Now we optimize over N. We take N such that

$$V_N \leq (|\mathcal{P}| + \varepsilon_M)^{\frac{(1-\bar{\varepsilon})(d+3-\bar{\varepsilon})}{2-\bar{\varepsilon}}} = (|\mathcal{P}| + \varepsilon_M)^{\frac{d+3}{2}(1-\frac{(d+5)\bar{\varepsilon}-2\bar{\varepsilon}^2}{(2-\bar{\varepsilon})(d+3)})}, \tag{113}$$

so

$$V_N \times \frac{1}{\delta^{d+1}} \leq (|\mathcal{P}| + \varepsilon_M)^{1-\bar{\varepsilon}}.$$

Then we have

$$\Big| \int_{\mathbb{R}^d} f_\delta(x)\rho_t^{\mathcal{P},M}(dx) - \frac{1}{N}\sum_{i=1}^N \mathbb{E}(f_\delta(X_t^{\mathcal{P},M,i}))\Big| \leq C\|f\|_\infty \times (|\mathcal{P}| + \varepsilon_M)^{1-\bar{\varepsilon}}. \tag{114}$$

Combining (108), (110) and (114), for all $f \in C_b^\infty(\mathbb{R}^d)$, with δ and N given in (111) and (113), when $t > \frac{8d}{\theta}(\frac{16}{\bar{\varepsilon}} + 1)$, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} f(x)\rho_t(dx) &= \frac{1}{N}\sum_{i=1}^N \mathbb{E}(f_\delta(X_t^{\mathcal{P},M,i})) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\bar{\varepsilon}}) \\
&= \frac{1}{N}\sum_{i=1}^N \mathbb{E}\int_{\mathbb{R}^d} f(X_t^{\mathcal{P},M,i} + y)\varphi_\delta(y)dy + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\bar{\varepsilon}}) \\
&= \frac{1}{N}\sum_{i=1}^N \mathbb{E}f(X_t^{\mathcal{P},M,i} + \delta\widetilde{\Delta}) + \|f\|_\infty \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\bar{\varepsilon}}),
\end{aligned} \tag{115}$$

where $\widetilde{\Delta}$ is a $d-$dimensional standard Gaussian random variable independent of $X_t^{\mathcal{P},M,i}, i = 1, \cdots, N$, and $\mathcal{O}(\bullet)$ is the Big $\mathcal{O}$ notation.

Since $C_b^\infty(\mathbb{R}^d)$ is dense in $C_b(\mathbb{R}^d)$, (115) holds for $f \in C_b(\mathbb{R}^d)$. Finally, by Lusin theorem, (115) also holds for any measurable and bounded function f.

Proof of Theorem 2.4 $ii)$**:** (37) is obtained in the same way as **Theorem 2.4** $i)$ by using **Lemma 3.5 (B)** $iv)$ in *Step 2*.

□

4 Proofs

4.1 Sobolev norms

In this section, we give the proof of **Lemma 3.7**. We explain our strategy of the proof. We will first prove that $\sup_{\mathcal{P}}\sup_{M}\|x_t^{\mathcal{P},M}\|_{L,l,p} \leq C_{l,p}$, then by an analogous argument, we also have $\sup_{M}\|x_t^{M}\|_{L,l,p} \leq C_{l,p}$. Afterwards, recalling $\mathbb{E}|x_t^M - x_t| \to 0$ in **Lemma 2.6** $i)$, and applying **Lemma 3.3** with $F_M = x_t^M$ and $F = x_t$, we get that x_t belongs to $\mathcal{D}_\infty^d$ and $\|x_t\|_{L,l,p} \leq C_{l,p}$.

So now we only need to prove the following lemma.

Lemma 4.1. *Under the* ***Hypothesis 2.1*** *and* ***Hypothesis 2.4*** $b)$*, for all $p \geq 2, l \geq 0$, there exists a constant $C_{l,p}$ depending on l, p, d and T, such that*

$$a) \quad \sup_{\mathcal{P}}\sup_{M}\mathbb{E}\sup_{0<t\leq T}|x_t^{\mathcal{P},M}|_l^p \leq C_{l,p}, \tag{116}$$

and

$$b) \quad \sup_{\mathcal{P}}\sup_{M}\mathbb{E}\sup_{0<t\leq T}|Lx_t^{\mathcal{P},M}|_l^p \leq C_{l,p}. \tag{117}$$

Before we prove this lemma, we give some pre-estimations concerning the Sobolev norms of Z_i^k.

Lemma 4.2. *Under the* ***Hypothesis 2.1, Hypothesis 2.4*** b*), for every* $l \geq 0$*, there exists a constant* C_l *dependent on* l, d *such that*

$$i) \quad \sup_{k,i\in\mathbb{N}} |Z_i^k|_{1,l} \leq C_l, \tag{118}$$

$$ii) \quad \sup_{k,i\in\mathbb{N}} |LZ_i^k|_l \leq C_l. \tag{119}$$

Proof. $i)$ We notice by the definition (93) that $D^Z_{(k,i,j)} Z^k_{i,j} = \xi^k_i, \quad D^Z_{(k',i',j')} Z^k_{i,j} = 0$, for $k' \neq k$, $i' \neq i$ or $j' \neq j$, $\quad D^\Delta Z^k_i = 0$. We recall that $\xi^k_i = \Psi_k(Z^k_i)$ in Section 3.2. We observe that using **Lemma 3.1** a), for any $k, i \in \mathbb{N}$ we have

$$|Z_i^k|_{1,l} \leq |\xi_i^k|_{l-1} = |\Psi_k(Z_i^k)|_{l-1} \leq 1 + C_l(|Z_i^k|_{1,l-1} + |Z_i^k|^{l-1}_{1,l-2}).$$

Since $|Z_i^k|_{1,1} = |\xi_i^k| \leq 1$, there is a constant C_l such that $|Z_i^k|_{1,l} \leq C_l$.

$ii)$ We notice by the definition (95) that

$$LZ^k_{i,j} = -\partial_{z^k_{i,j}}(\xi^k_i)^2 - \xi^k_i D^Z_{(k,i,j)} \ln[h(Z^k_i)].$$

We observe that $|\xi^k_i| \leq 1$, and we have $|\partial_{z^k_{i,j}}(\xi^k_i)^2| = 2\Psi_k(Z^k_i)\partial_{z^k_{i,j}}\Psi_k(Z^k_i)$ is bounded by a universal constant (see (92)). These lead to

$$|LZ^k_{i,j}|_l \leq C_l(1 + |D^Z_{(k,i,j)} \ln[h(Z^k_i)]|_l).$$

We recall by **Hypothesis 2.4** b) that h is infinitely differentiable and $\ln h$ has bounded derivatives of any order. Applying **Lemma 3.1** a) and using (118),

$$\begin{aligned} |D^Z_{(k,i,j)} \ln[h(Z^k_i)]|_l &\leq |\ln[h(Z^k_i)]|_{l+1} \\ &\leq C_l + |\nabla \ln h(Z^k_i)||Z^k_i|_{1,l+1} + C_l \sup_{2\leq|\beta|\leq l+1} |\partial^\beta \ln h(Z^k_i)||Z^k_i|^{l+1}_{1,l} \\ &\leq C_l. \end{aligned}$$

Then for any $k, i \in \mathbb{N}$, we obtain that $|LZ^k_i|_l \leq C_l$. □

Now we give the proof of **Lemma 4.1**.

Proof. **Proof of** a)**:** We first prove (116). We will prove by recurrence on l. One can easily check by (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = c_M(\tau(r), \eta^1_r(w), z, X^{\mathcal{P},M}_{\tau(r)-}, \rho^{\mathcal{P},M}_{\tau(r)-})$$

and by **Hypothesis 2.1** that for $l = 0$, $\mathbb{E} \sup_{0<t\leq T} |x^{\mathcal{P},M}_t|^p \leq C_{0,p}$. Then we assume that (116) holds for $l-1$ with $l \geq 1$ and for every $p \geq 2$. We will show that (116) also holds for l and for every $p \geq 2$. We notice by the definitions (93) that $D^\Delta_j \Delta = \boldsymbol{e_j}$, where $\boldsymbol{e_j} = (0, \cdots, 0, 1, 0, \cdots, 0)$ with value 1 at the $j-$th component, $D^{\Delta,q}\Delta = 0$ with $q \geq 2$ and $D^Z\Delta = 0$. Recalling the equation (52), we write $\mathbb{E} \sup_{0<t\leq T} |x^{\mathcal{P},M}_t|^p_{1,l} \leq C_{l,p}(1 + A_1 + A_2)$, with

$$A_1 = \mathbb{E}\int_0^T |b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^p_{1,l} dr,$$

$$A_2 = \mathbb{E}(\sum_{k=1}^M \sum_{i=1}^{J^k_T} |c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau(T^k_i)-})|_{1,l})^p.$$

Using **Lemma 3.1** a), **Hypothesis 2.1** and the recurrence hypothesis,

$$\begin{aligned} A_1 &\leq C_{l,p}[\mathbb{E}\int_0^T |\nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^p |x^{\mathcal{P},M}_{\tau(r)}|^p_{1,l} dr \\ &+ \mathbb{E}\int_0^T \sup_{2\leq|\beta|\leq l} |\partial^\beta_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^p |x^{\mathcal{P},M}_{\tau(r)}|^{lp}_{1,l-1} dr] \\ &\leq C_{l,p}[\int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)}|^p_{1,l} dr + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)}|^{lp}_{1,l-1} dr] \\ &\leq C_{l,p}[1 + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)}|^p_{1,l} dr]. \end{aligned} \tag{120}$$

Next, we estimate A_2. By **Lemma 3.1** a), **Hypothesis 2.1** and **Lemma 4.2**, for any $k, i \in \mathbb{N}$,

$$\begin{aligned} &|c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau(T^k_i)-})|_{1,l} \\ &\leq (|\nabla_z c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau(T^k_i)-})| + |\nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau(T^k_i)-})|) \\ &\times(|Z^k_i|_{1,l} + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|_{1,l}) \\ &+C_l \sup_{2\leq|\beta_1+\beta_2|\leq l} (|\partial^{\beta_1}_z \partial^{\beta_2}_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau(T^k_i)-})|) \times (|Z^k_i|^l_{1,l-1} + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|^l_{1,l-1}) \\ &\leq C_l \bar{c}(Z^k_i)(1 + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|_{1,l} + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|^l_{1,l-1}). \end{aligned}$$

Hence, using (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = \bar{c}(z)(1 + |x^{\mathcal{P},M}_{\tau(r)-}|_{1,l} + |x^{\mathcal{P},M}_{\tau(r)-}|^l_{1,l-1}),$$

Hypothesis 2.1 and the recurrence hypothesis, it follows that

$$\begin{aligned} A_2 &\leq C_{l,p}\mathbb{E}|\sum_{k=1}^M \sum_{i=1}^{J^k_T} \bar{c}(Z^k_i)(1 + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|_{1,l} + |x^{\mathcal{P},M}_{\tau(T^k_i)-}|^l_{1,l-1})|^p \\ &\leq C_{l,p}\mathbb{E}|\int_0^T \int_{[0,1]\times\mathbb{R}^d} \bar{c}(z)(1 + |x^{\mathcal{P},M}_{\tau(r)-}|_{1,l} + |x^{\mathcal{P},M}_{\tau(r)-}|^l_{1,l-1})\mathcal{N}(dw, dz, dr)|^p \\ &\leq C_{l,p}[1 + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)-}|^p_{1,l} dr + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)-}|^{lp}_{1,l-1} dr] \\ &\leq C_{l,p}[1 + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)}|^p_{1,l} dr]. \end{aligned} \tag{121}$$

Combining (120) and (121), one has

$$\mathbb{E}\sup_{0<t\leq T} |x^{\mathcal{P},M}_t|^p_{1,l} \leq C_{l,p}[1 + \int_0^T \mathbb{E}|x^{\mathcal{P},M}_{\tau(r)}|^p_{1,l} dr].$$

So we conclude by Gronwall lemma that

$$\sup_{\mathcal{P},M} \mathbb{E}\sup_{0<t\leq T} |x^{\mathcal{P},M}_t|^p_{1,l} \leq C_{l,p}. \tag{122}$$

Proof of b)**:** Now we pass to the proof of (117). We also prove it by recurrence on l.

Step 1: We take first $l = 0$. We notice by the definition (95) that $L\Delta = \Delta$. So having in mind that Δ has finite moments of any order, and recalling the equation (52), we write $\mathbb{E} \sup_{0<t\leq T} |Lx_t^{\mathcal{P},M}|^p \leq C_{0,p}(1 + S_1 + S_2)$, with

$$S_1 = \mathbb{E}\int_0^T |Lb(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P}}_{\tau(r)})|^p dr,$$
$$S_2 = \mathbb{E}(\sum_{k=1}^{M}\sum_{i=1}^{J_T^k} |Lc(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-})|)^p.$$

Using **Lemma 3.1** $c)$, **Hypothesis 2.1** and (116),

$$\begin{aligned} S_1 &\leq \mathbb{E}\int_0^T |\nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^p |Lx^{\mathcal{P},M}_{\tau(r)}|^p dr \\ &+ \mathbb{E}\int_0^T \sup_{|\beta|=2} |\partial_x^\beta b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^p |x^{\mathcal{P},M}_{\tau(r)}|^{2p}_{1,1} dr \\ &\leq C_{0,p}[1 + \int_0^T \mathbb{E}|Lx^{\mathcal{P},M}_{\tau(r)}|^p dr]. \end{aligned} \tag{123}$$

For S_2, we observe that using **Lemma 3.1** $c)$, **Lemma 4.2**, and **Hypothesis 2.1**, for any $k, i \in \mathbb{N}$,

$$\begin{aligned} &|Lc(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-})| \\ &\leq (|\nabla_z c(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-})| + |\nabla_x c(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-})|) \\ &\times(|LZ_i^k| + |Lx^{\mathcal{P},M}_{\tau(T_i^k)-}|) \\ &+ \sup_{|\beta_1+\beta_2|=2} (|\partial_z^{\beta_1}\partial_x^{\beta_2} c(\tau(T_i^k), \eta^1_{T_i^k}(W_i^k), Z_i^k, x^{\mathcal{P},M}_{\tau(T_i^k)-}, \rho^{\mathcal{P},M}_{\tau(T_i^k)-})|) \times (|Z_i^k|^2_{1,1} + |x^{\mathcal{P},M}_{\tau(T_i^k)-}|^2_{1,1}) \\ &\leq C\bar{c}(Z_i^k)(1 + |Lx^{\mathcal{P},M}_{\tau(T_i^k)-}| + |x^{\mathcal{P},M}_{\tau(T_i^k)-}|^2_{1,1}). \end{aligned}$$

Therefore, using (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = \bar{c}(z)(1 + |Lx^{\mathcal{P},M}_{\tau(r)-}| + |x^{\mathcal{P},M}_{\tau(r)-}|^2_{1,1}),$$

using **Hypothesis 2.1** and (116), it follows that

$$\begin{aligned} S_2 &\leq C_{0,p}\mathbb{E}(\sum_{k=1}^{M}\sum_{i=1}^{J_T^k} \bar{c}(Z_i^k)(1 + |Lx^{\mathcal{P},M}_{\tau(T_i^k)-}| + |x^{\mathcal{P},M}_{\tau(T_i^k)-}|^2_{1,1}))^p \\ &= C_{0,p}\mathbb{E}|\int_0^T \int_{[0,1]\times\mathbb{R}^d} \bar{c}(z)(1 + |Lx^{\mathcal{P},M}_{\tau(r)-}| + |x^{\mathcal{P},M}_{\tau(r)-}|^2_{1,1})\mathcal{N}(dw, dz, dr)|^p \\ &\leq C_{0,p}[1 + \int_0^T \mathbb{E}|Lx^{\mathcal{P},M}_{\tau(r)-}|^p dr]. \end{aligned} \tag{124}$$

Combining (123) and (124), one has

$$\mathbb{E}\sup_{0<t\leq T} |Lx_t^{\mathcal{P},M}|^p \leq C_{0,p}[1 + \int_0^T \mathbb{E}|Lx^{\mathcal{P},M}_{\tau(r)}|^p dr].$$

Applying Gronwall lemma, we obtain

$$\sup_{\mathcal{P},M} \mathbb{E}\sup_{0<t\leq T} |Lx_t^{\mathcal{P},M}|^p \leq C_{0,p}.$$

Step 2: Now we assume that (117) holds for $l-1$ with $l \geq 1$ and for every $p \geq 2$. We will show that (117) also holds for l and for every $p \geq 2$. We recall the equation (52) and that $L\Delta = \Delta$, $D_j^{\Delta}\Delta = \boldsymbol{e_j}$, where $\boldsymbol{e_j} = (0, \cdots, 0, 1, 0, \cdots, 0)$ with value 1 at the $j-$th component, $D^{\Delta,q}\Delta = 0$ with $q \geq 2$ and $D^{Z}\Delta = 0$. Having in mind that Δ has finite moments of any order, we write $\mathbb{E} \sup_{0<t\leq T} |Lx_t^{\mathcal{P},M}|_l^p \leq C_{l,p}(1 + B_1 + B_2)$, with

$$B_1 = \mathbb{E}\int_0^T |Lb(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P}})|_l^p dr,$$
$$B_2 = \mathbb{E}(\sum_{k=1}^{M}\sum_{i=1}^{J_T^k} |Lc(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M})|_l)^p.$$

Using **Lemma 3.1** b), **Hypothesis 2.1**, (116) and the recurrence hypothesis,

$$\begin{aligned}
B_1 &\leq C_{l,p}[\mathbb{E}\int_0^T |\nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})|^p |Lx_{\tau(r)}^{\mathcal{P},M}|_l^p dr \\
&+ \mathbb{E}\int_0^T \sup_{2\leq|\beta|\leq l+2} |\partial_x^\beta b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})|^p (1 + |x_{\tau(r)}^{\mathcal{P},M}|_{l+1}^{(l+2)p})(1 + |Lx_{\tau(r)}^{\mathcal{P},M}|_{l-1}^p) dr] \\
&\leq C_{l,p}[1 + \int_0^T \mathbb{E}|Lx_{\tau(r)}^{\mathcal{P},M}|_l^p dr]. \qquad (125)
\end{aligned}$$

Next, we estimate B_2. We observe that using **Lemma 3.1** b), **Lemma 4.2**, and **Hypothesis 2.1**, for any $k, i \in \mathbb{N}$,

$$\begin{aligned}
&|Lc(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M})|_l \\
&\leq (|\nabla_z c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M})| + |\nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M})|) \\
&\times(|LZ_i^k|_l + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_l) \\
&+C_l \sup_{2\leq|\beta_1+\beta_2|\leq l+2} (|\partial_z^{\beta_1}\partial_x^{\beta_2} c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau(T_i^k)-}^{\mathcal{P},M})|) \\
&\times(1 + |Z_i^k|_{l+1}^{l+2} + |x_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l+1}^{l+2})(1 + |LZ_i^k|_{l-1} + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l-1}) \\
&\leq C_l \bar{c}(Z_i^k)(1 + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_l + (1 + |x_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l+1}^{l+2})(1 + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l-1})).
\end{aligned}$$

Hence, using (45) with

$$\bar{\Phi}(r, w, z, \omega, \rho) = \bar{c}(z)(1 + |Lx_{\tau(r)-}^{\mathcal{P},M}|_l + (1 + |x_{\tau(r)-}^{\mathcal{P},M}|_{l+1}^{l+2})(1 + |Lx_{\tau(r)-}^{\mathcal{P},M}|_{l-1})),$$

using **Hypothesis 2.1**, (116), and the recurrence hypothesis, it follows that

$$\begin{aligned}
B_2 &\leq C_{l,p}\mathbb{E}(\sum_{k=1}^{M}\sum_{i=1}^{J_T^k} \bar{c}(Z_i^k)(1 + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_l + (1 + |x_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l+1}^{l+2})(1 + |Lx_{\tau(T_i^k)-}^{\mathcal{P},M}|_{l-1})))^p \\
&\leq C_{l,p}\mathbb{E}|\int_0^T \int_{[0,1]\times\mathbb{R}^d} \bar{c}(z)(1 + |Lx_{\tau(r)-}^{\mathcal{P},M}|_l + (1 + |x_{\tau(r)-}^{\mathcal{P},M}|_{l+1}^{l+2})(1 + |Lx_{\tau(r)-}^{\mathcal{P},M}|_{l-1}))\mathcal{N}(dw, dz, dr)|^p \\
&\leq C_{l,p}[1 + \int_0^T \mathbb{E}|Lx_{\tau(r)-}^{\mathcal{P},M}|_l^p dr]. \qquad (126)
\end{aligned}$$

Combining (125) and (126), one has

$$\mathbb{E}\sup_{0<t\leq T} |Lx_t^{\mathcal{P},M}|_l^p \leq C_{l,p}[1 + \int_0^T \mathbb{E}|Lx_{\tau(r)}^{\mathcal{P},M}|_l^p dr].$$

Then we conclude by Gronwall lemma that

$$\sup_{\mathcal{P},M}\mathbb{E}\sup_{0<t\leq T}|Lx_t^{\mathcal{P},M}|_l^p\leq C_{l,p}. \tag{127}$$

□

So now **Lemma 4.1** is proved. Then by an analogous argument, we also have $\sup_M \|x_t^M\|_{L,l,p}\leq C_{l,p}$. Finally, recalling that $\mathbb{E}|x_t^M-x_t|\to 0$ in **Lemma 2.6** i), and applying **Lemma 3.3** with $F_M=x_t^M$ and $F=x_t$, we get that x_t belongs to $\mathcal{D}_\infty^d$ and $\|x_t\|_{L,l,p}\leq C_{l,p}$.

4.2 Covariance matrices

In this section, we give the proof of **Lemma 3.8**.

Proof of i)**:** We proceed in 4 steps.

Step 1 We notice by the definitions (93) and the equation (51) that for any $M\in\mathbb{N}$, any $k_0,i_0\in\mathbb{N},j\in\{1,\cdots,d\}$,

$$\begin{aligned}
D^Z_{(k_0,i_0,j)}x_t^M&=\int_{T^{k_0}_{i_0}}^t\nabla_x b(r,x_r^M,\rho_r)D^Z_{(k_0,i_0,j)}x_r^M dr\\
&+\mathbb{1}_{\{0<T^{k_0}_{i_0}\leq t\}}\mathbb{1}_{\{1\leq k_0\leq M\}}\xi^{k_0}_{i_0}\partial_{z_j}c(T^{k_0}_{i_0},\eta^2_{T^{k_0}_{i_0}}(W^{k_0}_{i_0}),Z^{k_0}_{i_0},x^M_{T^{k_0}_{i_0}-},\rho_{T^{k_0}_{i_0}-})\\
&+\sum_{k=1}^M\sum_{T^{k_0}_{i_0}<T^k_i\leq t}\nabla_x c(T^k_i,\eta^2_{T^k_i}(W^k_i),Z^k_i,x^M_{T^k_i-},\rho_{T^k_i-})D^Z_{(k_0,i_0,j)}x^M_{T^k_i-},
\end{aligned}\tag{128}$$

$$D^\Delta_j x_t^M=a_T^M\boldsymbol{e_j}+\int_0^t\nabla_x b(r,x_r^M,\rho_r)D^\Delta_j x_r^M dr+\sum_{k=1}^M\sum_{i=1}^{J_t^k}\nabla_x c(T^k_i,\eta^2_{T^k_i}(W^k_i),Z^k_i,x^M_{T^k_i-},\rho_{T^k_i-})D^\Delta_j x^M_{T^k_i-}, \tag{129}$$

where $\boldsymbol{e_j}=(0,\cdots,0,1,0,\cdots,0)$ with value 1 at the $j-$th component.

Now we introduce $(Y_t^M)_{t\in[0,T]}$ (this is a variant of the tangent flow and for simplicity of the expression, we still call it the tangent flow) which is the matrix solution of the linear equation

$$Y_t^M=I_d\quad+\quad\int_0^t\nabla_x b(r,x_r^M,\rho_r)Y_r^M dr+\sum_{k=1}^M\sum_{i=1}^{J_t^k}\nabla_x c(T^k_i,\eta^2_{T^k_i}(W^k_i),Z^k_i,x^M_{T^k_i-},\rho_{T^k_i-})Y^M_{T^k_i-}.$$

And using Itô's formula, the inverse matrix $\widetilde{Y}_t^M=(Y_t^M)^{-1}$ verifies the equation

$$\widetilde{Y}_t^M=I_d\quad-\quad\int_o^t\widetilde{Y}_r^M\nabla_x b(r,x_r^M,\rho_r)dr-\sum_{k=1}^M\sum_{i=1}^{J_t^k}\widetilde{Y}^M_{T^k_i-}\nabla_x c(I_d+\nabla_x c)^{-1}(T^k_i,\eta^2_{T^k_i}(W^k_i),Z^k_i,x^M_{T^k_i-},\rho_{T^k_i-}).$$

Applying **Hypothesis 2.1** and **Hypothesis 2.2**, with C_p a constant not dependent on M, one also has

$$\mathbb{E}(\sup_{0<t\leq T}(\left\|Y_t^M\right\|^p+\left\|\widetilde{Y}_t^M\right\|^p))\leq C_p<\infty. \tag{130}$$

Then using the uniqueness of solution to the equation (128) and (129), one obtains

$$D^Z_{(k,i,j)}x_t^M=\mathbb{1}_{\{0<T^k_i\leq t\}}\mathbb{1}_{\{1\leq k\leq M\}}\xi^k_i Y_t^M\widetilde{Y}^M_{T^k_i}\partial_{z_j}c(T^k_i,\eta^2_{T^k_i}(W^k_i),Z^k_i,x^M_{T^k_i-},\rho_{T^k_i-}), \tag{131}$$

and $D_j^{\Delta} x_t^M = a_T^M Y_t^M \boldsymbol{e_j}$.

In the following, we denote the lowest eigenvalue of the Malliavin covariance matrix $\sigma_{x_t^M}$ by λ_t^M. Then we have (recalling the definitions (55) and (94))

$$\lambda_t^M = \inf_{|\zeta|=1} \langle \sigma_{x_t^M} \zeta, \zeta \rangle \geq \inf_{|\zeta|=1} \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} \sum_{j=1}^{d} \langle D^Z_{(k,i,j)} x_t^M, \zeta \rangle^2 + \inf_{|\zeta|=1} \sum_{j=1}^{d} \langle D_j^{\Delta} x_t^M, \zeta \rangle^2.$$

By (131),

$$\lambda_t^M \quad \geq \quad \inf_{|\zeta|=1} \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} \sum_{j=1}^{d} |\xi_i^k|^2 \langle \partial_{z_j} c(T_i^k, \eta^2_{T_i^k}(W_i^k), Z_i^k, x^M_{T_i^k -}, \rho_{T_i^k -}), (Y_t^M \widetilde{Y}^M_{T_i^k})^* \zeta \rangle^2 + \inf_{|\zeta|=1} \sum_{j=1}^{d} |a_T^M|^2 \langle \boldsymbol{e_j}, (Y_t^M)^* \zeta \rangle^2,$$

where Y^* denotes the transposition of a matrix Y.

We recall the ellipticity hypothesis (**Hypothesis 2.3**): there exists a non-negative function $\underline{c}(z)$ such that

$$\sum_{j=1}^{d} \langle \partial_{z_j} c(r, v, z, x, \rho), \zeta \rangle^2 \geq \underline{c}(z) |\zeta|^2.$$

So we deduce that

$$\begin{aligned}
\lambda_t^M \quad &\geq \quad \inf_{|\zeta|=1} \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} |\xi_i^k|^2 \underline{c}(Z_i^k) |(Y_t^M \widetilde{Y}^M_{T_i^k})^* \zeta|^2 + |a_T^M|^2 \inf_{|\zeta|=1} |(Y_t^M)^* \zeta|^2 \\
&\geq \quad \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} |\xi_i^k|^2 \underline{c}(Z_i^k) \|\widetilde{Y}_t^M\|^{-2} \|Y^M_{T_i^k}\|^{-2} + |a_T^M|^2 \|\widetilde{Y}_t^M\|^{-2} \\
&\geq \quad (\inf_{0<t\leq T} \|\widetilde{Y}_t^M\|^{-2} \|Y_t^M\|^{-2}) (\sum_{k=1}^{M} \sum_{i=1}^{J_t^k} |\xi_i^k|^2 \underline{c}(Z_i^k) + |a_T^M|^2).
\end{aligned}$$

We denote

$$\chi_t^M = \sum_{k=1}^{M} \sum_{i=1}^{J_t^k} |\xi_i^k|^2 \underline{c}(Z_i^k). \tag{132}$$

By (130), $(\mathbb{E} \sup_{0\leq t\leq T} \|\widetilde{Y}_t\|^{4dp} \|Y_t\|^{4dp})^{1/2} \leq C_{d,p} < \infty$, so that using Schwartz inequality, we have

$$\mathbb{E} |\frac{1}{\det \sigma_{x_t^M}}|^p \leq \mathbb{E}(|\lambda_t^M|^{-dp}) \leq C(\mathbb{E}(|\chi_t^M + |a_T^M|^2|^{-2dp}))^{\frac{1}{2}}. \tag{133}$$

Step 2 Since it is not easy to compute $\mathbb{E}(|\chi_t^M + |a_T^M|^2|^{-2dp}))$ directly, we make the following argument where the idea comes originally from [13]. Let $\Gamma(p) = \int_0^\infty s^{p-1} e^{-s} ds$ be the Gamma function. By a change of variables, we have the numerical equality

$$\frac{1}{|\chi_t^M + |a_T^M|^2|^{2dp}} = \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} e^{-s(\chi_t^M + |a_T^M|^2)} ds,$$

which, by taking expectation, gives

$$\mathbb{E}(\frac{1}{|\chi_t^M + |a_T^M|^2|^{2dp}}) = \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \mathbb{E}(e^{-s(\chi_t^M + |a_T^M|^2)}) ds. \tag{134}$$

Step 3 Now we compute $\mathbb{E}(e^{-s(\chi_t^M+|a_T^M|^2)})$ for any $s>0$. We recall that $I_1=B_1$, $I_k=B_k-B_{k-1}, k\geq 2$ (given in Section 2.4), and $\xi_i^k=\Psi_k(Z_i^k)$ (given in Section 3.2). We take $\Lambda_k(dz,dr)$ to be a Poisson point measure with intensity

$$\widehat{\Lambda}_k(dz,dr):=\mathbb{1}_{I_k}(z)\mu(dz)dr.$$

Since for different $k\in\mathbb{N}$, I_k are disjoint, the Poisson point measures $\Lambda_k, k\in\mathbb{N}$ are independent. And we put $\Theta_M(dz,dr)=\sum\limits_{k=1}^{M}\Lambda_k(dz,dr)$. Then

$$\chi_t^M = \sum_{k=1}^{M}\int_0^t\int_{I_k}|\Psi_k(z)|^2\underline{c}(z)\Lambda_k(dz,dr)=\int_0^t\int_{B_M}\Psi(z)\underline{c}(z)\Theta_M(dz,dr),$$

with $\Psi(z)=\sum\limits_{k=1}^{\infty}|\Psi_k(z)|^2\mathbb{1}_{I_k}(z)$. Using Itô formula,

$$\begin{aligned}\mathbb{E}(e^{-s\chi_t^M}) &= 1+\mathbb{E}\int_0^t\int_{B_M}(e^{-s(\chi_{r-}^M+\Psi(z)\underline{c}(z))}-e^{-s\chi_{r-}^M})\widehat{\Theta}_M(dz,dr)\\ &= 1-\int_0^t\mathbb{E}(e^{-s\chi_r^M})dr\int_{B_M}(1-e^{-s\Psi(z)\underline{c}(z)})\sum_{k=1}^{M}\mathbb{1}_{I_k}(z)\mu(dz).\end{aligned}$$

Solving the above equation we obtain

$$\begin{aligned}\mathbb{E}(e^{-s\chi_t^M}) &= \exp(-t\int_{B_M}(1-e^{-s\Psi(z)\underline{c}(z)})\sum_{k=1}^{M}\mathbb{1}_{I_k}(z)\mu(dz))\\ &= \exp(-t\sum_{k=1}^{M}\int_{I_k}(1-e^{-s|\Psi_k(z)|^2\underline{c}(z)})\mu(dz))\\ &\leq \exp(-t\sum_{k=1}^{M}\int_{I_k}(1-e^{-s\mathbb{1}_{[k-\frac{3}{4},k-\frac{1}{4}]}(|z|)\underline{c}(z)})\mu(dz))\\ &= \exp(-t\sum_{k=1}^{M}\int_{I_k}(1-e^{-s\underline{c}(z)})\mathbb{1}_{[k-\frac{3}{4},k-\frac{1}{4}]}(|z|)\mu(dz))\\ &= \exp(-t\int_{B_M}(1-e^{-s\underline{c}(z)})\nu(dz)),\end{aligned}$$

with

$$\nu(dz)=\sum_{k=1}^{\infty}\mathbb{1}_{[k-\frac{3}{4},k-\frac{1}{4}]}(|z|)\mu(dz).$$

On the other hand, we denote

$$\bar{\chi}_t^M=\int_0^t\int_{B_M^c}\Psi(z)\underline{c}(z)\Theta(dz,dr),$$

where B_M^c denote the complementary set of B_M and Θ is a Poisson point measure with intensity $\mu(dz)dr$. Then in the same way,

$$\mathbb{E}(e^{-s\bar{\chi}_t^M})\leq\exp(-t\int_{B_M^c}(1-e^{-s\underline{c}(z)})\nu(dz)).$$

We recall by (24) that $a_T^M=\sqrt{T\int_{\{|z|>M\}}\underline{c}(z)\mu(dz)}\geq\sqrt{\mathbb{E}\bar{\chi}_t^M}$. Notice that using Jensen inequality for the convex function $f(x)=e^{-sx}$, $s,x>0$, we have

$$e^{-s|a_T^M|^2}\leq e^{-s\mathbb{E}\bar{\chi}_t^M}\leq\mathbb{E}(e^{-s\bar{\chi}_t^M})\leq\exp(-t\int_{B_M^c}(1-e^{-s\underline{c}(z)})\nu(dz)).$$

So for every $M \in \mathbb{N}$, we deduce that

$$\begin{aligned}
&\mathbb{E}(e^{-s(\chi_t^M + |a_T^M|^2)}) = \mathbb{E}(e^{-s\chi_t^M}) \times e^{-s|a_T^M|^2} \\
&\leq \exp(-t\int_{B_M}(1-e^{-s\underline{c}(z)})\nu(dz)) \times \exp(-t\int_{B_M^c}(1-e^{-s\underline{c}(z)})\nu(dz)) \\
&= \exp(-t\int_{\mathbb{R}^d}(1-e^{-s\underline{c}(z)})\nu(dz)), \qquad (135)
\end{aligned}$$

and the last term does not depend on M.

Step 4 Now we use the Lemma 14 from [8], which states the following.

Lemma 4.3. *We consider an abstract measurable space B, a σ-finite measure $\mathcal{M}$ on this space and a non-negative measurable function $f : B \to \mathbb{R}_+$ such that $\int_B f d\mathcal{M} < \infty$. For $t > 0$ and $p \geq 1$, we note*

$$\beta_f(s) = \int_B (1-e^{-sf(x)})\mathcal{M}(dx) \quad \textit{and} \quad I_t^p(f) = \int_0^\infty s^{p-1}e^{-t\beta_f(s)}ds.$$

We suppose that for some $t > 0$ and $p \geq 1$,

$$\underline{\lim}_{u\to\infty} \frac{1}{\ln u}\mathcal{M}(f \geq \frac{1}{u}) > \frac{p}{t}, \qquad (136)$$

then $I_t^p(f) < \infty$.

We will use the above lemma for $\mathcal{M}(dz) = \nu(dz)$, $f(z) = \underline{c}(z)$ and $B = \mathbb{R}^d$. Thanks to (18) in **Hypothesis 2.4**,

$$\underline{\lim}_{u\to\infty} \frac{1}{\ln u}\nu(\underline{c} \geq \frac{1}{u}) = \theta > 0. \qquad (137)$$

Then for every $p \geq 1, 0 < t \leq T$, when $\theta > \frac{2dp}{t}$ (i.e. $t > \frac{2dp}{\theta}$), we deduce from (133),(134),(135) and **Lemma 4.3** that

$$\begin{aligned}
\sup_M \mathbb{E}|\frac{1}{\det \sigma_{x_t^M}}|^p &\leq \sup_M \mathbb{E}(|\lambda_t^M|^{-dp}) \leq C \sup_M (\mathbb{E}(|\chi_t^M + |a_T^M|^2|^{-2dp}))^{\frac{1}{2}} \\
&\leq C(\frac{1}{\Gamma(2dp)}\int_0^\infty s^{2dp-1}\mathbb{E}(e^{-s(\chi_t^M+|a_T^M|^2)})ds)^{\frac{1}{2}} \\
&\leq C(\frac{1}{\Gamma(2dp)}\int_0^\infty s^{2dp-1}\exp(-t\int_{\mathbb{R}^d}(1-e^{-s\underline{c}(z)})\nu(dz)ds)^{\frac{1}{2}} < \infty. \qquad (138)
\end{aligned}$$

Proof of $ii)$**:** We recall in **Lemma 2.6** $i)$ that $\mathbb{E}|x_t^M - x_t| \to 0$, and in **Lemma 3.7** that $\|x_t^M\|_{L,l,p} \leq C_{l,p}$. Moreover, by **Lemma 3.9** $ii)$, we know that $(Dx_t^M)_{M\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega; l^2 \times \mathbb{R}^d)$. Then applying **Lemma 3.3 (B)** with $F_M = x_t^M$ and $F = x_t$, by (96), we obtain (97).

□

5 Appendix

In the Appendix, we give the proof of **Lemma 3.9**.

Proof. **Proof of** i**):** We notice by the definitions (93) and the equations (52), (51) that for any partition $\mathcal{P} = \{0 = r_0 < r_1 < \cdots < r_{n-1} < r_n = T\}$, $M \in \mathbb{N}$, any $k_0, i_0 \in \mathbb{N}, j \in \{1, \cdots, d\}$,

$$
\begin{aligned}
D^Z_{(k_0,i_0,j)} x^{\mathcal{P},M}_t &= \int_0^t \nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)}) D^Z_{(k_0,i_0,j)} x^{\mathcal{P},M}_{\tau(r)} dr \\
&+ \mathbb{1}_{\{0<T^{k_0}_{i_0}\le t\}} \mathbb{1}_{\{1\le k_0\le M\}} \xi^{k_0}_{i_0} \partial_{z_j} c(\tau(T^{k_0}_{i_0}), \eta^1_{T^{k_0}_{i_0}}(W^{k_0}_{i_0}), Z^{k_0}_{i_0}, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^{k_0}_{i_0})-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^{k_0}_{i_0})-}) \\
&+ \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}) D^Z_{(k_0,i_0,j)} x^{\mathcal{P},M}_{\tau(T^k_i)-},
\end{aligned} \tag{139}
$$

$$
\begin{aligned}
D^\Delta_j x^{\mathcal{P},M}_t &= a^M_T \boldsymbol{e_j} + \int_0^t \nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)}) D^\Delta_j x^{\mathcal{P},M}_{\tau(r)} dr \\
&+ \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}) D^\Delta_j x^{\mathcal{P},M}_{\tau(T^k_i)-},
\end{aligned} \tag{140}
$$

where $\boldsymbol{e_j} = (0, \cdots, 0, 1, 0, \cdots, 0)$ with value 1 at the $j-$th component. And

$$
\begin{aligned}
D^Z_{(k_0,i_0,j)} x^M_t &= \int_0^t \nabla_x b(r, x^M_r, \rho_r) D^Z_{(k_0,i_0,j)} x^M_r dr \\
&+ \mathbb{1}_{\{0<T^{k_0}_{i_0}\le t\}} \mathbb{1}_{\{1\le k_0\le M\}} \xi^{k_0}_{i_0} \partial_{z_j} c(T^{k_0}_{i_0}, \eta^2_{T^{k_0}_{i_0}}(W^{k_0}_{i_0}), Z^{k_0}_{i_0}, x^M_{T^{k_0}_{i_0}-}, \rho_{T^{k_0}_{i_0}-}) \\
&+ \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^M_{T^k_i-}, \rho_{T^k_i-}) D^Z_{(k_0,i_0,j)} x^M_{T^k_i-},
\end{aligned} \tag{141}
$$

$$
D^\Delta_j x^M_t = a^M_T \boldsymbol{e_j} + \int_0^t \nabla_x b(r, x^M_r, \rho_r) D^\Delta_j x^M_r dr + \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^M_{T^k_i-}, \rho_{T^k_i-}) D^\Delta_j x^M_{T^k_i-}. \tag{142}
$$

For $h \in l_2$, we will use the notation $|h|^2_{l_2} = |h_{(\bullet,\circ,\diamond)}|^2_{l_2} = \sum_{k=1}^\infty \sum_{i=1}^\infty \sum_{j=1}^d |h_{(k,i,j)}|^2$. We write $\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_t - D^Z_{(\bullet,\circ,\diamond)} x^M_t|^2_{l_2} \le C[H_1 + H_2 + H_3]$, with

$$
H_1 = \mathbb{E}| \int_0^t \nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)}) D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(r)} dr - \int_0^t \nabla_x b(r, x^M_r, \rho_r) D^Z_{(\bullet,\circ,\diamond)} x^M_r dr|^2_{l_2},
$$

$$
\begin{aligned}
H_2 &= \mathbb{E}| \mathbb{1}_{\{0<T^\bullet_\circ\le t\}} \mathbb{1}_{\{1\le \bullet\le M\}} (\partial_{z_\diamond} c(\tau(T^\bullet_\circ), \eta^1_{T^\bullet_\circ}(W^\bullet_\circ), Z^\bullet_\circ, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^\bullet_\circ)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^\bullet_\circ)-}) \\
&- \partial_{z_\diamond} c(T^\bullet_\circ, \eta^2_{T^\bullet_\circ}(W^\bullet_\circ), Z^\bullet_\circ, x^M_{T^\bullet_\circ-}, \rho_{T^\bullet_\circ-}))|^2_{l_2},
\end{aligned}
$$

$$
\begin{aligned}
H_3 &= \mathbb{E}| \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}) D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(T^k_i)-} \\
&- \sum_{k=1}^M \sum_{i=1}^{J^k_t} \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^M_{T^k_i-}, \rho_{T^k_i-}) D^Z_{(\bullet,\circ,\diamond)} x^M_{T^k_i-}|^2_{l_2}.
\end{aligned}
$$

We take a small $\varepsilon_* > 0$. We recall ε_M in (27). Firstly, using **Hypothesis 2.1**, we get

$$\begin{aligned} H_1 &\leq C[\mathbb{E}\int_0^t |\nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)}) - \nabla_x b(r, x^M_r, \rho_r)|^2 |D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr \\ &+ \mathbb{E}\int_0^t |\nabla_x b(\tau(r), x^{\mathcal{P},M}_{\tau(r)}, \rho^{\mathcal{P},M}_{\tau(r)})|^2 |D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr] \\ &\leq C[\mathbb{E}\int_0^t [|\mathcal{P}|^2 + |x^{\mathcal{P},M}_{\tau(r)} - x^M_r|^2 + (W_1(\rho^{\mathcal{P},M}_{\tau(r)}, \rho_r))^2] |D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr \\ &+ \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr]. \end{aligned}$$

Then by **Lemma 3.7**, using Hölder inequality with conjugates $1 + \frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, by **Lemma 2.6** and (50), we have

$$\begin{aligned} H_1 &\leq C[|\mathcal{P}|^2 + \int_0^t (\mathbb{E}|x^{\mathcal{P},M}_{\tau(r)} - x^M_r|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (W_{2+\varepsilon_*}(\rho^{\mathcal{P},M}_{\tau(r)}, \rho_r))^2 dr \\ &+ \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr] \\ &\leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)} x^M_r|^2_{l_2} dr]. \end{aligned} \tag{143}$$

Secondly, using **Hypothesis 2.1** and the isometry of the Poisson point measure $\mathcal{N}$, we get

$$\begin{aligned} H_2 &= \mathbb{E}\sum_{k=1}^M \sum_{i=1}^{J^k_t} \sum_{j=1}^d |\partial_{z_j} c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}) - \partial_{z_j} c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^M_{T^k_i-}, \rho_{T^k_i-})|^2 \\ &\leq C\mathbb{E}\sum_{k=1}^M \sum_{i=1}^{J^k_t} |\bar{c}(Z^k_i)|^2 [|\tau^{\mathcal{P}}(T^k_i) - T^k_i| + |\eta^1_{T^k_i}(W^k_i) - \eta^2_{T^k_i}(W^k_i)| + |x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-} - x^M_{T^k_i-}| + W_1(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho_{T^k_i-})]^2 \\ &\leq C\mathbb{E}\int_0^t \int_{[0,1]\times\mathbb{R}^d} |\bar{c}(z)|^2 [|\tau^{\mathcal{P}}(r) - r|^2 + |\eta^1_r(w) - \eta^2_r(w)|^2 \\ &+ |x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-} - x^M_{r-}|^2 + (W_1(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-}, \rho_{r-}))^2] \mathcal{N}(dw, dz, dr) \\ &= C\mathbb{E}\int_0^t \int_{[0,1]\times\mathbb{R}^d} |\bar{c}(z)|^2 [|\tau^{\mathcal{P}}(r) - r|^2 + |\eta^1_r(w) - \eta^2_r(w)|^2 \\ &+ |x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-} - x^M_{r-}|^2 + (W_1(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-}, \rho_{r-}))^2] dw \mu(dz) dr. \end{aligned}$$

Then by (39), (43), **Lemma 2.6**, (50), and Hölder inequality with conjugates $1 + \frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, we have

$$\begin{aligned} H_2 &\leq C[|\mathcal{P}|^2 + \int_0^t (\mathbb{E}|x^{\mathcal{P},M}_{\tau(r)} - x_r|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (\mathbb{E}|x^{\mathcal{P},M}_{\tau(r)} - x^M_r|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (W_{2+\varepsilon_*}(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)}, \rho_r))^2 dr] \\ &\leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}. \end{aligned} \tag{144}$$

Thirdly, we write $H_3 \leq C[H_{3,1} + H_{3,2}]$, where

$$\begin{aligned} H_{3,1} &= \mathbb{E}(\sum_{k=1}^M \sum_{i=1}^{J^k_t} |\nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}) \\ &- \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^M_{T^k_i-}, \rho_{T^k_i-})| |D^Z_{(\bullet,\circ,\diamond)} x^M_{T^k_i-}|_{l_2})^2, \end{aligned}$$

$$H_{3,2} = \mathbb{E}(\sum_{k=1}^M \sum_{i=1}^{J^k_t} |\nabla_x c(\tau(T^k_i), \eta^1_{T^k_i}(W^k_i), Z^k_i, x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-}, \rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(T^k_i)-})| |D^Z_{(\bullet,\circ,\diamond)} x^{\mathcal{P},M}_{\tau(T^k_i)-} - D^Z_{(\bullet,\circ,\diamond)} x^M_{T^k_i-}|_{l_2})^2.$$

Using **Hypothesis 2.1** and (45) with

$$\bar{\Phi}(r,w,z,\omega,\rho) = |\nabla_x c(\tau(r),\eta_r^1(w),z,x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-}) - \nabla_x c(r,\eta_r^2(w),z,x^M_{r-},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|_{l_2},$$

we get

$$\begin{aligned} H_{3,1} &\leq \mathbb{E}(\int_0^t \int_{[0,1]\times\mathbb{R}^d} |\nabla_x c(\tau(r),\eta_r^1(w),z,x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-}) \\ &- \nabla_x c(r,\eta_r^2(w),z,x^M_{r-},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|_{l_2}\mathcal{N}(dw,dz,dr))^2 \\ &\leq C\mathbb{E}\int_0^t\int_{[0,1]}[|\tau^{\mathcal{P}}(r)-r|^2 + |\eta_r^1(w)-\eta_r^2(w)|^2 \\ &+ |x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-} - x^M_{r-}|^2 + (W_1(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho_{r-}))^2]|D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|^2_{l_2}dwdr. \end{aligned}$$

Then using (39), (43), **Lemma 3.7**, and Hölder inequality with conjugates $1+\frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, we have

$$\begin{aligned} H_{3,1} &\leq C[|\mathcal{P}|^2 + \int_0^t (\mathbb{E}|x^{\mathcal{P},M}_{\tau(r)-} - x_{r-}|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}}dr \\ &+ \int_0^t (\mathbb{E}|x^{\mathcal{P},M}_{\tau(r)-} - x^M_{r-}|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}}dr + \int_0^t (W_{2+\varepsilon_*}(\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho_{r-}))^2 dr] \\ &\leq C(|\mathcal{P}|+\varepsilon_M)^{\frac{2}{2+\varepsilon_*}}, \end{aligned}$$

where the last inequality is a consequence of **Lemma 2.6** and (50).
Using **Hypothesis 2.1**, (45) with

$$\bar{\Phi}(r,w,z,\omega,\rho) = |\nabla_x c(\tau(r),\eta_r^1(w),z,x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-})||D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(r)-} - D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|_{l_2},$$

we have

$$\begin{aligned} H_{3,2} &\leq \mathbb{E}(\int_0^t\int_{[0,1]\times\mathbb{R}^d} |\nabla_x c(\tau(r),\eta_r^1(w),z,x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-})||D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(r)-} - D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|_{l_2}\mathcal{N}(dw,dz,dr))^2 \\ &\leq C\int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(r)-} - D^Z_{(\bullet,\circ,\diamond)}x^M_{r-}|^2_{l_2}dr. \end{aligned}$$

Therefore,

$$H_3 \leq C[H_{3,1}+H_{3,2}] \leq C[(|\mathcal{P}|+\varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)}x^M_r|^2_{l_2}dr]. \quad (145)$$

Combining (143), (144) and (145),

$$\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_t - D^Z_{(\bullet,\circ,\diamond)}x^M_t|^2_{l_2} \leq C[(|\mathcal{P}|+\varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(r)} - D^Z_{(\bullet,\circ,\diamond)}x^M_r|^2_{l_2}dr].$$

In a similar way, we notice by (116), the isometry of the Poisson point measure N, and (45) with

$$\bar{\Phi}(r,w,z,\omega,\rho) = \nabla_x c(\tau(r),\eta_r^1(w),z,x^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-},\rho^{\mathcal{P},M}_{\tau^{\mathcal{P}}(r)-})D^Z_{(k_0,i_0,j)}x^{\mathcal{P},M}_{\tau(r)-}$$

that

$$\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_{\tau(t)} - D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_t|^2_{l_2} \leq C|\mathcal{P}|, \quad (146)$$

so

$$\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_t - D^Z_{(\bullet,\circ,\diamond)}x^M_t|^2_{l_2} \leq C[(|\mathcal{P}|+\varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_r - D^Z_{(\bullet,\circ,\diamond)}x^M_r|^2_{l_2}dr].$$

We conclude by Gronwall lemma that $\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{\mathcal{P},M}_t - D^Z_{(\bullet,\circ,\diamond)}x^M_t|^2_{l_2} \leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}$. Finally, by a similar argument, $\mathbb{E}|D^\Delta_{(\diamond)}x^{\mathcal{P},M}_t - D^\Delta_{(\diamond)}x^M_t|^2_{\mathbb{R}^d} \leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}$, and we obtain what we need.

Proof of ii**):** We only need to prove that for any $M_1, M_2 \in \mathbb{N}$ with $\varepsilon_{M_1 \wedge M_2} \leq 1$ and $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M_1\wedge M_2\}} \leq 1$, we have

$$\|Dx^{M_1}_t - Dx^{M_2}_t\|_{L^2(\Omega;l^2\times\mathbb{R}^d)} \leq (\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{1}{2+\varepsilon_*}}. \tag{147}$$

In fact, if $(Dx^M_t)_{M\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega; l^2 \times \mathbb{R}^d)$, then it has a limit Y in $L^2(\Omega; l^2 \times \mathbb{R}^d)$. But when we apply **Lemma 3.3 (A)** with $F_M = X^M_t$ and $F = X_t$, we know that there exists a convex combination $\sum\limits_{M'=M}^{m_M} \gamma^M_{M'} \times F^{M'}_t$, with $\gamma^M_{M'} \geq 0, M' = M, ..., m_M$ and $\sum\limits_{M'=M}^{m_M} \gamma^M_{M'} = 1$, such that

$$\left\| \sum_{M'=M}^{m_M} \gamma^M_{M'} \times Dx^{M'}_t - Dx_t \right\|_{L^2(\Omega;l^2\times\mathbb{R}^d)} \to 0,$$

as $M \to \infty$. Meanwhile, we have

$$\left\| \sum_{M'=M}^{m_M} \gamma^M_{M'} \times DF^{M'}_t - Y \right\|_{L^2(\Omega;l^2\times\mathbb{R}^d)} \leq \sum_{M'=M}^{m_M} \gamma^M_{M'} \left\| Dx^{M'}_t - Y \right\|_{L^2(\Omega;l^2\times\mathbb{R}^d)} \to 0.$$

So $Y = Dx_t$ and we conclude by passing to the limit $M_2 \to \infty$ in (147).

Now we prove (147). We recall the equation (141) and we write $\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_t - D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_t|^2_{l_2} \leq C[O_1 + O_2 + O_3]$, with

$$O_1 = \mathbb{E}| \int_0^t \nabla_x b(r, x^{M_1}_r, \rho_r) D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_r dr - \int_0^t \nabla_x b(r, x^{M_2}_r, \rho_r) D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r dr|^2_{l_2},$$

$$\begin{aligned} O_2 &= \mathbb{E}|\mathbb{1}_{\{0<T^\bullet_\circ\leq t\}}(\mathbb{1}_{\{1\leq\bullet\leq M_1\}}\partial_{z_\diamond} c(T^\bullet_\circ, \eta^2_{T^\bullet_\circ}(W^\bullet_\circ), Z^\bullet_\circ, x^{M_1}_{T^\bullet_\circ-}, \rho_{T^\bullet_\circ-}) \\ &- \mathbb{1}_{\{1\leq\bullet\leq M_2\}}\partial_{z_\diamond} c(T^\bullet_\circ, \eta^2_{T^\bullet_\circ}(W^\bullet_\circ), Z^\bullet_\circ, x^{M_2}_{T^\bullet_\circ-}, \rho_{T^\bullet_\circ-}))|^2_{l_2}, \end{aligned}$$

$$\begin{aligned} O_3 &= \mathbb{E}| \sum_{k=1}^{M_1} \sum_{i=1}^{J^k_t} \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^{M_1}_{T^k_i-}, \rho_{T^k_i-}) D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_{T^k_i-} \\ &- \sum_{k=1}^{M_2} \sum_{i=1}^{J^k_t} \nabla_x c(T^k_i, \eta^2_{T^k_i}(W^k_i), Z^k_i, x^{M_2}_{T^k_i-}, \rho_{T^k_i-}) D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_{T^k_i-}|^2_{l_2}. \end{aligned}$$

Firstly, using **Hypothesis 2.1**, we have

$$\begin{aligned} O_1 &\leq C[\mathbb{E}\int_0^t |\nabla_x b(r, x^{M_1}_r, \rho_r) - \nabla_x b(r, x^{M_2}_r, \rho_r)|^2 |D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr \\ &+ \mathbb{E}\int_0^t |\nabla_x b(r, x^{M_1}_r, \rho_r)|^2 |D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_r - D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr] \\ &\leq C[\mathbb{E}\int_0^t |x^{M_1}_r - x^{M_2}_r|^2 |D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_r - D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr]. \end{aligned}$$

Then by **Lemma 3.7**, Hölder inequality with conjugates $1 + \frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, and **Lemma 2.6**, we obtain

$$\begin{aligned} O_1 &\leq C[\int_0^t (\mathbb{E}|x^{M_1}_r - x^{M_2}_r|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_r - D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr] \\ &\leq C[(\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_r - D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_r|^2_{l_2} dr]. \end{aligned} \tag{148}$$

Secondly, using **Hypothesis 2.1**, the isometry of the Poisson point measure $\mathcal{N}$, we have

$$\begin{aligned}
O_2 &\leq C[\mathbb{E}\sum_{k=1}^{M_1}\sum_{i=1}^{J_t^k}\sum_{j=1}^{d}|\partial_{z_j}c(T_i^k,\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_1}_{T_i^k-},\rho_{T_i^k-})-\partial_{z_j}c(T_i^k,\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_2}_{T_i^k-},\rho_{T_i^k-})|^2\\
&+ \mathbb{E}\sum_{k=M_1\wedge M_2}^{M_1\vee M_2}\sum_{i=1}^{J_t^k}\sum_{j=1}^{d}|\partial_{z_j}c(T_i^k,\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_2}_{T_i^k-},\rho_{T_i^k-})|^2]\\
&\leq C[\mathbb{E}\sum_{k=1}^{M}\sum_{i=1}^{J_t^k}|\bar{c}(Z_i^k)|^2|x^{M_1}_{T_i^k-}-x^{M_2}_{T_i^k-}|^2+\mathbb{E}\sum_{k=M_1\wedge M_2}^{M_1\vee M_2}\sum_{i=1}^{J_t^k}|\bar{c}(Z_i^k)|^2]\\
&\leq C[\mathbb{E}\int_0^t\int_{[0,1]\times\mathbb{R}^d}|\bar{c}(z)|^2|x^{M_1}_{r-}-x^{M_2}_{r-}|^2\mathcal{N}(dw,dz,dr)+\mathbb{E}\int_0^t\int_{[0,1]}\int_{\{|z|>M_1\wedge M_2\}}|\bar{c}(z)|^2\mathcal{N}(dw,dz,dr)]\\
&= C[\mathbb{E}\int_0^t\int_{[0,1]\times\mathbb{R}^d}|\bar{c}(z)|^2|x^{M_1}_{r-}-x^{M_2}_{r-}|^2dw\mu(dz)dr+\mathbb{E}\int_0^t\int_{[0,1]}\int_{\{|z|>M_1\wedge M_2\}}|\bar{c}(z)|^2dw\mu(dz)dr].
\end{aligned}$$

Then by Hölder inequality with conjugates $1+\frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, **Hypothesis 2.1** and **Lemma 2.6**,

$$\begin{aligned}
O_2 &\leq C[\int_0^t(\mathbb{E}|x^{M_1}_r-x^{M_2}_r|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}}dr+\varepsilon_{M_1\wedge M_2}]\\
&\leq C(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}. \qquad (149)
\end{aligned}$$

Thirdly, we write $O_3\leq C[O_{3,1}+O_{3,2}+O_{3,3}]$, where

$$O_{3,1}=\mathbb{E}(\sum_{k=M_1\wedge M_2}^{M_1\vee M_2}\sum_{i=1}^{J_t^k}|\nabla_x c(\tau(T_i^k),\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_1\vee M_2}_{T_i^k-},\rho_{T_i^k-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_1\vee M_2}_{T_i^k-}|_{l_2})^2,$$

$$\begin{aligned}
O_{3,2} &= \mathbb{E}(\sum_{k=1}^{M_1\wedge M_2}\sum_{i=1}^{J_t^k}|\nabla_x c(T_i^k,\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_1}_{T_i^k-},\rho_{T_i^k-})\\
&- \nabla_x c(T_i^k,\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_2}_{T_i^k-},\rho_{T_i^k-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_{T_i^k-}|_{l_2})^2,
\end{aligned}$$

$$O_{3,3}=\mathbb{E}(\sum_{k=1}^{M_1\wedge M_2}\sum_{i=1}^{J_t^k}|\nabla_x c(\tau(T_i^k),\eta^2_{T_i^k}(W_i^k),Z_i^k,x^{M_1}_{T_i^k-},\rho_{T_i^k-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_1}_{T_i^k-}-D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_{T_i^k-}|_{l_2})^2.$$

Using **Hypothesis 2.1**, **Lemma 3.7**, (44) with

$$\bar{\Phi}(r,w,z,\omega,\rho)=|\nabla_x c(\tau(r),\eta^2_r(w),z,x^{M_1\vee M_2}_{r-},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_1\vee M_2}_{r-}|_{l_2},$$

we get

$$\begin{aligned}
O_{3,1} &\leq \mathbb{E}(\int_0^t\int_{[0,1]}\int_{\{|z|>M_1\wedge M_2\}}|\nabla_x c(\tau(r),\eta^2_r(w),z,x^{M_1\vee M_2}_{r-},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_1\vee M_2}_{r-}|_{l_2}\mathcal{N}(dw,dz,dr))^2\\
&\leq C[(\int_{\{|z|>M_1\wedge M_2\}}\bar{c}(z)\mu(dz))^2+\int_{\{|z|>M_1\wedge M_2\}}|\bar{c}(z)|^2\mu(dz)]\\
&= C\varepsilon_{M_1\wedge M_2}.
\end{aligned}$$

Using **Hypothesis 2.1**, **Lemma 3.7**, (45) with

$$\bar{\Phi}(r,w,z,\omega,\rho)=|\nabla_x c(r,\eta^2_r(w),z,x^{M_1}_{r-},\rho_{r-})-\nabla_x c(r,\eta^2_r(w),z,x^{M_2}_{r-},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)}x^{M_2}_{r-}|_{l_2},$$

by **Lemma 2.6**, and Hölder inequality with conjugates $1+\frac{\varepsilon_*}{2}$ and $\frac{2+\varepsilon_*}{\varepsilon_*}$, we have

$$\begin{aligned}
O_{3,2} &\leq \mathbb{E}(\int_0^t \int_{[0,1]\times\mathbb{R}^d} |\nabla_x c(r,\eta_r^2(w),z,x_{r-}^{M_1},\rho_{r-}) - \nabla_x c(r,\eta_r^2(w),z,x_{r-}^{M_2},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_2}|_{l_2} \mathcal{N}(dw,dz,dr))^2 \\
&\leq C\mathbb{E}\int_0^t \int_{[0,1]} |x_{r-}^{M_1} - x_{r-}^{M_2}|^2 |D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_2}|^2_{l_2} dwdr \\
&\leq C\int_0^t (\mathbb{E}|x_{r-}^{M_1} - x_{r-}^{M_2}|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr \\
&\leq C(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}.
\end{aligned}$$

Using **Hypothesis 2.1**, (45) with

$$\bar{\Phi}(r,w,z,\omega,\rho) = |\nabla_x c(r,\eta_r^2(w),z,x_{r-}^{M_1},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_2}|_{l_2},$$

we have

$$\begin{aligned}
O_{3,3} &\leq \mathbb{E}(\int_0^t \int_{[0,1]\times\mathbb{R}^d} |\nabla_x c(r,\eta_r^2(w),z,x_{r-}^{M_1},\rho_{r-})||D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_2}|_{l_2} \mathcal{N}(dw,dz,dr))^2 \\
&\leq C\int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_{r-}^{M_2}|^2_{l_2} dr.
\end{aligned}$$

Therefore,

$$O_3 \leq C[O_{3,1}+O_{3,2}+O_{3,3}] \leq C[(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x_r^{M_1} dr - D^Z_{(\bullet,\circ,\diamond)} x_r^{M_2}|^2_{l_2} dr]. \quad (150)$$

Combining (148), (149) and (150),

$$\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x_t^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_t^{M_2}|^2_{l_2} \leq C[(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x_r^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_r^{M_2}|^2_{l_2} dr].$$

So we conclude by Gronwall lemma that $\mathbb{E}|D^Z_{(\bullet,\circ,\diamond)} x_t^{M_1} - D^Z_{(\bullet,\circ,\diamond)} x_t^{M_2}|^2_{l_2} \leq C(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}$.

Finally, We recall by (24) that $a_T^M = \sqrt{T\int_{\{|z|>M\}} \underline{c}(z)\mu(dz)}$ and by **Hypothesis 2.3** that $\underline{c}(z) \leq |\bar{c}(z)|^2$. We notice that

$$\mathbb{E}|a_T^{M_1}\boldsymbol{e}_\diamond - a_T^{M_2}\boldsymbol{e}_\diamond|^2_{\mathbb{R}^d} \leq C\mathbb{E}|a_T^{M_1} - a_T^{M_2}|^2 \leq C\int_{\{|z|>M_1\wedge M_2\}} \underline{c}(z)\mu(dz) \leq \varepsilon_{M_1\wedge M_2}.$$

Then by a similar argument as above, $\mathbb{E}|D^\Delta_{(\diamond)} x_t^{M_1} - D^\Delta_{(\diamond)} x_t^{M_2}|^2_{\mathbb{R}^d} \leq C(\varepsilon_{M_1}+\varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}$, and we obtain (147).

Proof of $iii)$**:** $iii)$ is an immediate consequence of $i)$ and $ii)$.

□

Acknowledgment The author thanks Prof. Vlad Bally for his kind and patient guidance and many useful suggestions.

References

[1] S. Albeverio, B. Rüdiger, and P. Sundar: The Enskog process. *J. Stat. Phys.* **167**(1):90–122 (2017).

[2] R. Alexandre: A review of Boltzmann equation with singular kernels. *Kinet. Relat. Models*, **2**(4), 551-646 (2009).

[3] A. Alfonsi, V. Bally : Construction of Boltzmann and McKean Vlasov type flows (the sewing lemma approach). arXiv:2105.12677 [math.PR](2021).

[4] A. Alfonsi, B. Jourdain, A. Kohatsu-Higa: Optimal transport bounds between the time-marginals of a multidimensional diffusion and its Euler scheme. *Electronic Journal of Probability* **20**, 1-31 06 (2015).

[5] F. Antonelli, A. Kohatsu-Higa: Rate of convergence of a particle method to the solution of the McKean-Vlasov equation. *The Annals of Applied Probability*, **12**(2): 423-476 (2002).

[6] D. Applebaum: *Lévy Processes and Stochastic Calculus* (2nd ed., Cambridge Studies in Advanced Mathematics). Cambridge: Cambridge University Press. (2009) doi:10.1017/CBO9780511809781.

[7] V. Bally, L. Caramellino, G. Poly: Regularization lemmas and convergence in total variation. *Electron. J. Probab.* **25** 1-20. (2020).

[8] V. Bally, E. Clément: Integration by parts formula and applications to equations with jumps. *Probab. Th. Rel. Fields*, **151**, 613-657 (2011).

[9] V. Bally, N. Fournier: Regularization properties of the 2D homogeneous Boltzmann equation without cutoff. *Probab. Theory Related Fields*, **151**(3-4), 659-704 (2011).

[10] V. Bally, Y. Qin: Total variation distance between a jump-equation and its Gaussian approximation. *Stoch PDE: Anal Comp*. (2022).

[11] V. Bally, D. Talay: The Law of the Euler scheme for stochastic differential equations: I. convergence rate of the distribution function. [Research Report] RR-2244, INRIA. (1994). ⟨inria-00074427⟩

[12] J. Bao, C. Reisinger, P. Ren, et al: First-order convergence of Milstein schemes for McKean–Vlasov equations and interacting particle systems. *Proceedings of the Royal Society A*, **477**(2245): 20200258 (2021).

[13] K. Bichteler, J. B. Gravereaux, J. Jacod: *Malliavin calculus for processes with jumps*. Gordon and Breach, (1987).

[14] R. Carmona, F. Delarue: *Probability theory of mean field games with applications*. Springer *Probability Theory and Stochastic Modelling* Vol.**83** (2018).

[15] C. Cercignani: *The Boltzmann equation and its applications*. Springer-Verlag *Applied Mathematical Sciences*, Vol.**67** (1988).

[16] R. Cont, P. Tankov: *Finacial modelling with jump processes*. Chapman & Hall/CRC (2004).

[17] L. Desvillettes, C. Graham, S. Méléard: Probabilistic interpretation and numerical approximation of a Kac equation without cutoff. *Stochastic Process. Appl.*, **84**(1), 115-135 (1999).

[18] N. Fournier, A. Guillin: From a Kac-like particle system to the Landau equation for hard potentials and Maxwell molecules. *Ann. Sci. Éc. Norm. Supér.* (4), **50**(1), 157-199 (2017).

[19] N. Fournier, S. Mischler: Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Ann. Proba.*, **44**(1), 589-627 (2016).

[20] M. Friesen, B. Rüdiger, and P. Sundar: The Enskog process for hard and soft potentials. *NoDEA Nonlinear Differential Equations Appl.* **26**(3):Paper No. 20, 42 (2019).

[21] M. Friesen, B. R*ü*diger, and P. Sundar: On uniqueness and stability for the Enskog equation (2020).

[22] E. Gobet, G. Pag*è*s, H. Pham and J. Printems: Discretization and simulation for a class of SPDEs with applications to Zakai and McKean-Vlasov equations. (2005). ⟨hal-00003917v1⟩

[23] C. Graham: Mckean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications* **40** 69-82 (1992).

[24] A. L. Haji-Ali, R. Tempone: Multilevel and Multi-index Monte Carlo methods for the McKean–Vlasov equation. *Statistics and Computing*, **28**(4): 923-935 (2018).

[25] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes*. 2nd ed. Amsterdam, Netherlands, North Holland, (1989).

[26] Y. Ishikawa: *Stochastic Calculus of Variations for Jump Processes*, Berlin, Boston: De Gruyter. (2013).

[27] B. Jourdain and A. Kohatsu-Higa: A review of recent results on approximation of solutions of stochastic differential equations. Proceedings of the Workshop on Stochastic Analysis with Financial Applications: Hong Kong (2009). Birkhauser (2011).

[28] A. Kohatsu-Higa: The Euler approximation for stochastic differential equations with boundary conditions. Proceedings of the Workshop on Turbulent Diffusion and Related Problems in Stochastic Numerics. The Institute of Statistical Mathematics, Tokyo (1996).

[29] A. Kohatsu-Higa, S. Ogawa: Monte Carlo methods weak rate of convergence for an Euler scheme of nonlinear SDE's. *Monte Carlo Methods and Its Applications*, vol **3**, 327-345 (1997).

[30] A. Kohatsu-Higa, P. Protter: The Euler scheme for 9SDE's driven by semimartingales. In Stochastic analysis on infinite dimensional spaces. H. Kunita and H.Kuo (Eds.), 141-151, Pitman Research Notes in Mathematics Series ,vol. **310** (1994).

[31] A. Kohatsu-Higa and P. Tankov: Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*, Vol. **120**, 2258-2285 (2010).

[32] A. M. Kulik: Malliavin calculus for Lévy processes with arbitrary Lévy measures. *Theor. Probability ad Math. Statist.* No.**72**, 75-92 (2006).

[33] A. M. Kulik: Stochastic calculus of variations for general Lévy processes and its applications to jump-type SDE's with non-degenerated drift. arXiv:math/0606427 (2007).

[34] H. Kunita: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: Rao, MM,ed. *Real and stochastic analysis*. Boston, USA, Birkhaäuser, 305-373 (2004).

[35] H. Kunita: *Stochastic flows and jump-diffusions*. Springer,(2019).

[36] B. Lapeyre, *É*. Pardoux and R. Sentis: *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*. Springer-Verlag, Berlin, (1998).

[37] E. Mariucci, M. Reiß: Wasserstein and total variation distance between marginals of Lévy processes. *Electronic Journal of Statistics* **12**, 2482-2514 (2018).

[38] S. Méléard: Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. *Probabilistic models for nonlinear partial differential equations* volume **1627** of Lecture Notes in Math., pages 42–95. Springer, Berlin (1996).

[39] M. A. Mezerdi: On the convergence of carathéodory numerical scheme for Mckean-Vlasov equations. *Stochastic Analysis and Applications*, **39**(5): 804-818 (2021).

[40] R. Naito, T. Yamada: A higher order weak approximation of McKean–Vlasov type SDEs. *BIT Computer Science and Numerical Mathematics* United States (2022).

[41] D. Nualart: *The Malliavin calculus and related topics*. Springer-Verlag, (2006).

[42] P. Protter and D. Talay: The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.* Vol **25**, No.1, pg 393-423 (1997).

[43] K. Sato. Lévy processes and infinitely divisible distributions. *Cambridge University press*, Cambridge (1999).

[44] K. Sato. Basic results on Lévy processes. In O.E. Barndorff-Nielsen, T. Mikosch, and S.I. Resnick, editors, *Lévy processes. Theory and applications*, pages 3–37. Birkhäuser (2001).

[45] Y. Song and X. Zhang: Regularity of density for SDEs driven by degenerate Lévy noises. arXiv:1401.4624 (2014).

[46] A.-S. Sznitman: Équations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete*, **66**(4):559–592 (1984).

[47] L. Szpruch, S. Tan, A. Tse: Iterative multilevel particle approximation for Mckean–Vlasov SDEs. *The Annals of Applied Probability* Vol. **29**, No. 4, 2230–2265 (2019).

[48] H. Tanaka: Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete.* **46**(1), 67-105 (1978).

[49] H. Tanaka: Stochastic differential equation corresponding to the spatially homogeneous Boltzmann equation of Maxwellian and noncutoff type. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **34**(2), 351-369 (1987).

[50] C. Villani: On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations. *Arch. Rational Mech. Anal.*, **143**(3), 273-307 (1998).

[51] C. Villani: *Optimal Transport* Springer-Verlag, (2009).

[52] X. Zhang: Densities for SDEs driven by degenerate α-stable processes. *Ann. Probab.* Vol **42**, No.5, 1885-1910 (2014).