



# Approximation of the stochastic differential equation with jumps and convergence in total variation distance

Yifeng Qin

## ► To cite this version:

Yifeng Qin. Approximation of the stochastic differential equation with jumps and convergence in total variation distance. Analysis of PDEs [math.AP]. Université Gustave Eiffel, 2023. English. NNT : 2023UEFL2028 . tel-04231403v2

**HAL Id: tel-04231403**

**<https://hal.science/tel-04231403v2>**

Submitted on 8 Nov 2023

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

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

**UNIVERSITÉ GUSTAVE EIFFEL**

**ÉCOLE DOCTORALE MSTIC**

**Thèse de doctorat**

**Mathématique-Probabilité**

**QIN Yifeng**

Approximation de l'équation différentielle stochastique avec sauts et convergence en distance  
de la variation totale

*Thèse dirigée par ...Professeur Vlad BALLY...Professeur Dan GOREAC.....*

Soutenue le ...23/06/2023.....

**Jury :**

M. Vlad BALLY, Université Gustave Eiffel

M. Dan GOREAC, Université Gustave Eiffel; Shandong University, Weihai

Mme. Emmanuel CLÉMENT, Université Gustave Eiffel

M. Rainer BUCKDAHN (Rapporteur), Université de Bretagne Occidentale, Brest

M. Dan CRISAN (Rapporteur), Imperial College London, South Kensington Campus

Mme. Juan LI, Shandong University, Weihai

# Approximation de l'équation différentielle stochastique avec sauts et convergence en distance de la variation totale

Université Gustave Eiffel  
LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA  
Cité Descartes, 5 Boulevard Descartes  
Champs-sur-Marne 77454 Marne-la-Vallée Cedex 2  
FRANCE

## Resumé

Dans cette thèse, nous considérons l'approximation de la solution de l'équation différentielle stochastique avec des sauts et nous nous concentrons sur la convergence en distance de la variation totale. La principale méthode que nous utilisons est la technique d'intégration par parties dans le calcul de Malliavin. Cette thèse contient trois parties. Dans la première partie, nous voulons obtenir un schéma d'approximation précis pour l'équation de saut. Suivant l'idée de [1], nous remplaçons les "petits sauts" par un mouvement brownien. Nous prouvons que pour chaque temps fixé  $t$ , la variable aléatoire approchée  $X_t^\varepsilon$  converge vers la variable aléatoire originale  $X_t$  en distance de variation totale et nous estimons l'erreur. Nous donnons également une estimation de la distance entre les densités des lois des deux variables aléatoires.

Dans la seconde partie, nous traitons des équations à sauts de type McKean-Vlasov et Boltzmann. Cela signifie que les coefficients de l'équation dépendent de la loi de la solution et que l'équation est dirigée par une mesure ponctuelle de Poisson avec une mesure d'intensité qui dépend également de la loi de la solution. Dans [2], Alfonsi et Bally ont prouvé que sous certaines conditions convenables, la solution  $X_t$  d'une telle équation existe et est unique. Ils prouvent également que  $X_t$  est l'interprétation probabiliste d'une équation faible analytique. De plus, étant donné une partition  $\mathcal{P}$  de l'intervalle de temps, ils définissent  $X_t^{\mathcal{P}}$  comme étant le schéma d'Euler associé à  $\mathcal{P}$ , et prouvent que  $X_t^{\mathcal{P}}$  converge vers  $X_t$  en distance de Wasserstein. Dans cette thèse, sous des hypothèses plus restreintes, nous montrons que le schéma d'Euler  $X_t^{\mathcal{P}}$  converge vers  $X_t$  en distance de variation totale et  $X_t$  a une densité "smooth" (qui est une fonction solution de l'équation faible analytique). D'autre part, en vue de la simulation, nous utilisons un schéma d'Euler tronqué  $X_t^{\mathcal{P},M}$  qui a un nombre fini de sauts dans tout intervalle compact. Nous prouvons que  $X_t^{\mathcal{P},M}$  converge également vers  $X_t$  en distance de variation totale. Enfin, nous donnons un algorithme basé sur un système de particules associé à  $X_t^{\mathcal{P},M}$  afin d'approximer la densité de la loi de  $X_t$ . Des estimations complètes de l'erreur sont obtenues.

Dans la troisième partie, nous établissons un cadre abstrait pour l'approximation de la mesure de probabilité invariante d'un semi-groupe de Markov. Suivant l'approche de [4], nous utilisons le schéma d'Euler avec étapes décroissantes (qui est appelé l'algorithme de Langevin non-ajusté dans les littératures de Monte Carlo) pour faire la simulation. Sous certaines "propriétés de Lipschitz exponentielle" et propriétés de régularisation, nous donnons une estimation de l'erreur en distance de la variation totale. Les principaux résultats dans [4] et [3] sont des cas particuliers de notre cadre. Nous appliquons également ce cadre aux processus de saut et obtenons une estimation de l'approximation de la mesure de probabilité invariante en distance de variation totale.

**Mots clés:** Equations différentielles stochastiques avec sauts, Calcul de Malliavin, Intégration par parties, Distance de la variation totale, Distance de Wasserstein, Equation de McKean-Vlasov, Equation de Boltzmann, Système de particules, Mesure de probabilité invariante

## References

- [1] S. Asmussen and J. Rosinski: Approximations of small jumps of Lévy processes with a view towards simulation, *Journal of Applied Probability* **38** , 482-493 (2001).
- [2] A. Alfonsi, V. Bally : Construction of Boltzmann and McKean Vlasov type flows (the sewing lemma approach). arXiv:2105.12677 [math.PR](2021).
- [3] P. Chen, C. Deng, René L. Schilling, et. al.: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. arXiv:2205.01342v2 [math.PR] (2023).
- [4] G. Pagès, F. Panloup: Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. (hal-03082311v3) (2022).

# Approximation of the stochastic differential equation with jumps and convergence in total variation distance

## Abstract

In this thesis, we consider the approximation of the solution of the stochastic differential equation with jumps and we focus on the convergence in total variation distance. The main method we use is the integration by parts technique in Malliavin calculus. This thesis contains three parts. In the first part, we want to obtain an accurate approximation scheme for the jump equation. Following the idea of [1], we replace the "small jumps" by a Brownian motion. We prove that for every fixed time  $t$ , the approximate random variable  $X_t^\varepsilon$  converges to the original random variable  $X_t$  in total variation distance and we estimate the error. We also give an estimate of the distance between the densities of the laws of the two random variables.

In the second part, we deal with McKean-Vlasov and Boltzmann type jump equations. This means that the coefficients of the 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 [2], Alfonsi and Bally have proved that under some suitable conditions, the solution  $X_t$  of such equation exists and is unique. They also prove that  $X_t$  is the probabilistic interpretation of an analytical weak equation. Moreover, given a partition  $\mathcal{P}$  of the time interval, they define  $X_t^{\mathcal{P}}$  to be the Euler scheme associated to  $\mathcal{P}$ , and prove that  $X_t^{\mathcal{P}}$  converges to  $X_t$  in Wasserstein distance. In this thesis, 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.

In the third part, we establish an abstract framework for the approximation for the invariant probability measure of a Markov semigroup. Following the approach from [4], we use the Euler scheme with decreasing steps (which is called the unadjusted Langevin algorithm in the Monte Carlo literature) to do the simulation. Under some "exponential Lipschitz property" and regularization properties, we give an estimate of the error in total variation distance. The main results in [4] and [3] are particular cases of our framework. We also apply this framework to jump processes and obtain an estimate of the approximation for the invariant probability measure in total variation distance.

**Key words:** Stochastic differential equations with jumps, Malliavin calculus, Integration by parts, Total variation distance, Wasserstein distance, McKean-Vlasov equation, Boltzmann equation, Particle system, Invariant probability measure

## References

- [1] S. Asmussen and J. Rosinski: Approximations of small jumps of Lévy processes with a view towards simulation, *Journal of Applied Probability* **38** , 482-493 (2001).

- [2] A. Alfonsi, V. Bally : Construction of Boltzmann and McKean Vlasov type flows (the sewing lemma approach). arXiv:2105.12677 [math.PR](2021).
- [3] P. Chen, C. Deng, René L. Schilling, et. al.: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. arXiv:2205.01342v2 [math.PR] (2023).
- [4] G. Pagès, F. Panloup: Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. (hal-03082311v3) (2022).

# Acknowledgment

I would like to express my deepest gratitude to my advisors Prof. Vlad Bally and Prof. Dan Goreac for their kind and patient guidance, their dedication and many useful suggestions. I am very grateful for the nice and free atmosphere in the laboratory LAMA.

I would also like to thank my family and friends for their supports in these 3 years, and thank Mathinnov, Fondation Sciences Mathématiques de Paris for their financial support.

## Résultats de la Partie I

Le premier chapitre concerne l'approximation d'équations différentielles à sauts de type

$$(1) \quad X_t = x + \int_0^t \int_{(0,1]} c(s, z, X_{s-}) N_\mu(ds, dz),$$

où  $x$  est la valeur initiale, avec  $T > 0$  l'horizon de temps,  $\mu$  est une mesure  $\sigma$ -finie à support  $(0, 1]$ ,  $N_\mu$  est une mesure ponctuelle de Poisson sur  $(0, 1]$  dont le compensateur est  $\mu(dz)ds$  et  $c$  est un coefficient qui vérifiera des hypothèses de régularité. Les sauts de faible amplitude sont agrégés dans un drift  $b_\varepsilon(s, x) = \int_{(0,\varepsilon]} c(s, z, x) \mu(dz)$  et un terme diffusif gouverné par un mouvement Brownien  $W_\mu$  espace-temps :

$$(2) \quad \begin{aligned} X_t^\varepsilon &= x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, X_{s-}^\varepsilon) N_\mu(ds, dz) \\ &\quad + \int_0^t b_\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \int_{(0,\varepsilon]} c(s, z, X_s^\varepsilon) W_\mu(ds, dz). \end{aligned}$$

Nous nous concentrons sur l'estimation de l'erreur commise en distance de la variation totale  $d_{TV}$ . Le résultat principal que nous mettons en évidence est que, pour tout paramètre de temps fixé  $t$ , la variable aléatoire approximante  $X_t^\varepsilon$  converge vers la variable aléatoire d'origine  $X_t$ , que cette convergence a lieu en variation totale et, résultat quantitatif fondamental, nous sommes en mesure de fournir une estimation de l'erreur commise. En plus, nous obtenons une estimation de la distance entre les densités des lois associées aux variables aléatoires susmentionnées.

Nous commencerons par une courte liste de notations.

- $T > 0$  est l'horizon de temps ;
- $C_b^l(\mathbb{R})$  est l'espace des fonctions dérivable  $l$ -fois et dont les dérivées sont bornées ; la norme associée est  $\|f\|_{l,\infty} := \sum_{i \leq l} \|f^{(i)}\|_\infty$  ;
- $C_p^\infty$  sera l'espace des fonctions régulières présentant, avec leur dérivées, une croissance polynômiale ;
- $P_t$  et  $P_t^\varepsilon$  dénotent les semi-groupes de  $X$  et  $X^\varepsilon$ , tandis que les générateur infinitésimaux sont  $L_t$ ,  $L_t^\varepsilon$  ;
- $d_{TV}$  est la distance en variation totale regardée comme cas particulier de

$$d_k(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \sum_{0 \leq i \leq k} \|f^{(i)}\|_\infty \leq 1\},$$

lorsque  $F, G$  sont des v.a. réelle sur  $\Omega$ .

Les résultats classiques de type Trotter-Kato conduisent à

$$(3) \quad \sup_{s \leq T} \|(P_t - P_t^\varepsilon)f\|_\infty \leq C \|f\|_{3,\infty} \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz),$$

avec  $\hat{c}(z) := \sup_{s \leq T} \sup_{x \in \mathbb{R}} |c(s, z, x)|$ . Nous améliorons ceci afin d'être en mesure d'appliquer la distance pour des fonctions  $f$  ayant une moindre régularité.

Les étapes de notre travail sont :



1. Nous présentons un cadre abstrait d'intégration par parties dans **Section 3, Partie I**. Notre cadre unifié inclut, au delà du calcul de Malliavin standard, de différentes autres versions, entre autres le calcul basée sur la méthode « splitting » utilisée dans [3] [4] [5], le  $\Gamma$ -calcul de [2], etc. Ce cadre Malliavin-Sobolev nous permet d'obtenir **Lemma 3.4, Partie I** : pour tout  $\delta > 0$ , on a

$$(4) \quad i) \quad d_{TV}(F, G) \leq C \times d_r(F, G)^{1-\delta}.$$

Lorsque les v.a.  $F, G$  ont des densités régulières  $p_F, p_G$ , pour  $l \in \mathbb{N}$ , on a

$$(5) \quad ii) \quad \|p_F - p_G\|_{l, \infty} \leq C \times d_r(F, G)^{1-\delta}.$$

2. Une raison de technique nous emmène à effectuer le changement de variable  $z \mapsto \frac{1}{z}$ . Soit  $\theta : (0, 1] \rightarrow [1, \infty)$  une fonction telle que  $\theta(z) = \frac{1}{z}$ . Nous travaillons avec une équation régie par une mesure finie

$$(6) \quad \begin{aligned} \hat{X}_t^M &= x + \int_0^t \int_{[1, M)} \tilde{c}(s, z, \hat{X}_{s-}^M) N_\nu(ds, dz) \\ &+ \int_0^t b_M(s, \hat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \tilde{c}(s, z, \hat{X}_s^M) W_\nu(ds, dz), \end{aligned}$$

où  $M = \frac{1}{\varepsilon}$ ,  $\nu(dz) = \mu \circ \theta^{-1}(dz)$ ,  $\tilde{c}(s, z, x) = c(s, \frac{1}{z}, x)$ ,

$$(7) \quad b_M(s, x) = \int_{\{z \geq M\}} \tilde{c}(s, z, x) \nu(dz).$$

3. Une hypothèse sectorielle nous permet de travailler avec une représentation via un processus de Poisson composé localisé sur  $z \in I_k := [k, k+1)$

$$\begin{aligned} \hat{X}_t^M &= x + \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \tilde{c}(T_i^k, Z_i^k, \hat{X}_{T_i^k-}^M) \\ &+ \int_0^t b_M(s, \hat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \tilde{c}(s, z, \hat{X}_s^M) W_\nu(ds, dz). \end{aligned}$$

Ici,  $T_i^k$  sont les temps des sauts du processus de Poisson  $J^k$  de paramètre  $\nu(I_k)$  et  $Z_i^k$  sont des variables aléatoires indépendantes de loi  $\mathbf{1}_{I_k} \frac{\nu(dz)}{\nu(I_k)}$  et indépendantes aussi de  $J^k$ . Les techniques de type Malliavin concerneront à la fois  $W$  et  $Z$  et sont expliquées dans la **Section 4 de la Partie I**.

4. Avec ces outils, la contribution principale de cette partie est de montrer que, pour tout  $\delta > 0$ , on a

$$(8) \quad d_{TV}(X_t, X_t^\varepsilon) \leq C_\delta \left( \int_{(0, \varepsilon]} |\hat{c}(z)|^3 \mu(dz) \right)^{1-\delta}.$$

En plus, lorsque  $X_t$  et  $X_t^\varepsilon$  ont de densités  $p_t(x, y)dy$ , respectivement  $p^\varepsilon(x, y)dy$  suffisamment régulières en  $y$ , on a

$$(9) \quad \left\| \partial_y^k p_t - \partial_y^k p_t^\varepsilon \right\|_\infty \leq C_{k, \delta} \left( \int_{(0, \varepsilon]} |\hat{c}(z)|^3 \mu(dz) \right)^{1-\delta},$$

pour tout  $k \in \mathbb{N}$  et tout  $\delta > 0$ .

Concernant la stratégie de preuve et les hypothèses nécessaires, mentionnons les aspects suivants.

1. Nous demandons une régularité du coefficient  $c$  et de ses dérivées jusqu'à un ordre  $q^* \in \mathbb{N}$  dans **Hypothesis 2.1, Partie I** ainsi que l'ellipticité **Hypothesis 2.3, Partie I** ;
2. Comme nous travaillerons avec des représentations de type processus de Poisson composé, nous avons besoin d'une hypothèse sectorielle **Hypothesis 2.4, Partie I** ;
3. Une difficulté supplémentaire concerne la loi des  $Z_i^k$  (voir ci-dessus) supposée à densité dont le logarithme est régulier. Afin d'arriver à maîtriser ces lois, nous utilisons une hypothèse de type Doeblin et la méthode dite du "splitting" (voir **Section 4.1 dans la Partie I**).

## Résultats de la Partie II

Le deuxième chapitre se penche sur des équations à sauts avec des composantes McKean-Vlasov ou dans l'esprit de l'équation de Boltzmann. En d'autres termes, nous considérons des équations dont les coefficients dépendent de la loi même de la solution cherchée. Ces équations sont régies par des mesures aléatoires ponctuelles de Poisson dont l'intensité a une forme de dépendance trajectorielle de la solution. Notre point de départ est l'article [1] dans lequel les auteurs prouvent un résultat d'existence et unicité de telles solutions (notées  $X_t$  de manière générique à ce niveau). Cette solution fournit une interprétation probabiliste de la solution à une équation analytique faiblement posée. Sous des hypothèses plus fortes que celle d'origine, nous prouvons que le schéma d'Euler (notée  $X^{\mathcal{P}}$  associée à la solution converge en variation totale vers la solution initiale laquelle possède une densité régulière. D'autre part, nous considérons une deuxième approximation  $X^{\mathcal{P},M}$  obtenue en introduisant la finitude de sauts dans les intervalles compacts. Cette deuxième approximation converge elle-même vers la solution initiale et toujours en variation totale. Nous présentons un algorithme de type particules permettant d'approcher la densité de la loi de la solution initiale et nous proposons une analyse de l'erreur.

Nous commencerons par une courte liste de notations

- $T > 0$  est l'horizon de temps ;
- lorsque  $\beta$  est un multi-indice, sa longueur est notée par  $|\beta|$  ;
- $C_b^l(\mathbb{R}^d)$  est l'espace des fonctions dérivable  $l$ -fois et dont les dérivées sont bornées ; la norme associée est  $\|f\|_{l,\infty} := \sum_{|\beta| \leq l} \|\partial^\beta f\|_\infty$  ;
- $\mathcal{P}_l(\mathbb{R}^d)$  est l'espace des mesures de probabilités ayant un moment d'ordre  $l$  fini ; la distance de Wasserstein pour  $l = 1$  est notée par  $W_1$  ; cette distance sera écrite sur des v.a.  $F, G \in L^1(\Omega)$  en ayant en tête les lois  $\mathcal{L}(F)$  and  $\mathcal{L}(G)$  ;
- $\mathcal{P} = \{0 = r_0 < r_1 < \dots < r_{n-1} < r_n = T\}$  est une partition de l'intervalle de temps  $[0, T]$  et on définit  $\tau(r) = r_k$  quand  $r \in [r_k, r_{k+1})$ .

L'équation qui fera l'objet de notre travail est

$$(10) \quad X_t = X_0 + \int_0^t b(r, X_r, \rho_r) dr + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(r, v, z, X_{r-}, \rho_{r-}) N_{\rho_{r-}}(dv, dz, dr),$$

où  $\rho_t(dv) = \mathbb{P}(X_t \in dv)$  et  $N_{\rho_t}$  est une mesure ponctuelle de Poisson dont l'intensité est  $\rho_t(dv)\mu(dz)dr$ . La mesure  $\mu$  est positive,  $\sigma$ -finie sur  $\mathbb{R}^d$ .  $b$  et  $c$  sont des fonctions qui vérifient certaines conditions

de régularité et d'ellipticité (voir **Hypothèses 2.1~2.4** dans la Section 2.2, Partie II). En particulier, on suppose que pour tout multi-indices  $\beta_1, \beta_2$ , il existe une fonction positive  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  tel que pour chaque  $z \in \mathbb{R}^d$ , on a

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

avec  $\int_{\mathbb{R}^d} |\bar{c}(z)|^p \mu(dz) < \infty$ ,  $\forall p \geq 1$ . De plus, on suppose qu'il existe une fonction positive  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  telle que pour tout  $r \in [0, T], v, x \in \mathbb{R}^d, z \in \mathbb{R}^d, \rho \in \mathcal{P}_1(\mathbb{R}^d), \zeta \in \mathbb{R}^d$ , on a

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

Les étapes de notre travail sont :

### 1. Le premier pas

- (a) remplacer  $c$  avec une mesure  $c_M$  localisée sur  $B_M := \{z \in \mathbb{R}^d : |z| \leq M\}$  ;
- (b) les sauts  $|z| > M$  seront remplacés par une v.a. Gaussienne et aura une contribution  $a_T^M \Delta := \Delta \sqrt{T \int_{|z| > M} \underline{c}(z) \mu(dz)}$  ;

$$\begin{aligned} X_t^{\mathcal{P}, M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M}) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P}, M}, \rho_{\tau(r)-}^{\mathcal{P}, M}) N_{\rho_{\tau(r)-}^{\mathcal{P}, M}}(dv, dz, dr), \end{aligned} \quad (11)$$

où  $\rho_t^{\mathcal{P}, M}(dv) = \mathbb{P}(X_t^{\mathcal{P}, M} \in dv)$ .

- (c) dans **Theorem 2.2, Partie II**, le schéma résultant (dont la solution est notée par  $X_t^{\mathcal{P}, M}$ ) converge vers  $X_t$  en  $d_{TV}$ , i.e. , pour tout  $\varepsilon > 0$ , il existe une constante  $C > 0$  telle que

$$d_{TV}(X_t^{\mathcal{P}, M}, X_t) \leq C(\sqrt{\varepsilon_M} + |\mathcal{P}|)^{1-\varepsilon} \rightarrow 0, \quad (12)$$

lorsque  $(|\mathcal{P}|, M) \rightarrow (0, \infty)$ . La quantité  $\varepsilon_M$  est donnée explicitement par

$$\varepsilon_M := \int_{\{|z| > M\}} |\bar{c}(z)|^2 \mu(dz) + \left| \int_{\{|z| > M\}} \bar{c}(z) \mu(dz) \right|^2. \quad (13)$$

### 2. Le deuxième pas : fournir un schéma pour $\rho^{\mathcal{P}, M}$

- (a) Ce pas est basé sur un système classique de particules  $X_t^{\mathcal{P}, M, i}$  et l'utilisation de la moyenne empirique  $\hat{\rho}$  pour remplacer  $\rho$  ;
- (b) comme dans la première partie, le schéma est complété par l'utilisation d'un processus de Poisson composé adéquate.

### 3. Régularité et approximation

- (a) **Théorème 2.1, Partie II** donne l'existence de la densité  $p_t(x)$  de la loi de  $X_t$  ;
- (b) **Théorème 2.3, Partie II** donne des résultats d'approximation de ladite densité pour  $\delta \in \left\{ (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{d+3}}, (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{d+5}} \right\}$  et  $N$  suffisamment grand. Dans le premier cas on obtient

$$p_t(x) = \frac{1}{N} \sum_{i=1}^N \mathbb{E} \varphi_\delta(X_t^{\mathcal{P}, M, i} - x) + \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{2}{d+3}}). \quad (14)$$

Dans le deuxième cas, on obtient

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

Les quantités  $\varphi$  et  $\varphi_{\delta}$  correspondent au noyau Gaussien standard en dimension  $d$ , respectivement une modification avec  $\delta$ -échelle. La deuxième approximation est basée sur la méthode de Romberg. Comme toujours, nous avons le phénomène dit "curse of dimensionality".

4. Lorsque l'on considère une perturbation avec une v.a. Gaussienne  $d$ -dimensionnelle supplémentaire  $\tilde{\Delta}$ , nous obtenons, pour  $\varepsilon > 0$

$$(15) \quad \int_{\mathbb{R}^d} f(x) p_t(x) dx = \frac{1}{N} \sum_{i=1}^N \mathbb{E} f(X_t^{\mathcal{P},M,i} + \delta \tilde{\Delta}) + \|f\|_{\infty} \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}),$$

pour toute fonction mesurable bornée  $f$ . Le paramètre  $\delta$  peut être choisi comme  $\delta = (|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{1}{2}(1-\frac{\varepsilon}{2-\varepsilon})}$  et le résultat est valable pour  $N$  assez grand.

Concernant la stratégie de preuve et les hypothèses nécessaires, mentionnons les aspects suivants.

1. Afin d'utiliser une méthode d'intégration par parties, nous avons besoin (voir **Hypothesis 2.4, Partie II**) que  $\mu$  est absolument continue par rapport à la mesure de Lebesgue, que la dérivée de Radon-Nikodym est infiniment dérivable et que son logarithme l'est aussi et possède des dérivées bornées.
2. Nous employons des techniques de couplage adéquate pour la distance  $W_{2+\varepsilon_*}$ , avec  $\varepsilon_* > 0$  suffisamment petit. Des résultats de représentabilité sont fournis dans **Lemme 2.5, Partie II** et employés dans les arguments.
3. **Section 3 de la Partie II** décrit les techniques d'intégration par parties de type Malliavin afin d'obtenir les résultats pour la distance en variation totale  $d_{TV}$ . Par rapport à la Partie I, la non-dégénérescence de la covariance de Malliavin n'est pas garantie pour les deux v.a. Nous proposons une extension dans la **Proposition 3.6.1 de la Partie II**.
4. La régularisation nécessaire pour proposer l'algorithme de type particules conduit aux résultats dans **Lemme 3.5** et **Corolaire 3.5.1, Partie II**.

### Résultats de la Partie III

La troisième et dernière contribution vise à établir, dans un cadre relativement général et abstrait, une méthode d'approximation de la loi invariante d'un processus. Cette contribution se base sur le schéma d'Euler à pas décroissant (algorithme de type Langevin non-ajusté dans la littérature de type méthodes de Monte-Carlo). Sous une condition de Lipschitz exponentielle et sous condition de propriété de régularisation, nous obtenons une estimation de l'erreur en variation totale. Ceci généralise les articles [7] et [6]. La méthode trouve ses applications aux processus à sauts avec, comme but, d'obtenir l'approximation de la mesure invariante et quantifier la précision de l'approximation en variation totale.

Nous nous proposons de fournir un cadre abstrait pour l'approximation de la mesure de probabilité invariante.

Nous commencerons par une courte liste de notations.

- lorsque  $\beta$  est un multi-indice, sa longueur est notée par  $|\beta|$  ;

- $C_b^l(\mathbb{R}^d)$  est l'espace des fonctions dérivable  $l$ -fois et dont les dérivées sont bornées ; la norme associée est  $\|f\|_{l,\infty} := \sum_{|\beta| \leq l} \|\partial^\beta f\|_\infty$  ;
- sur l'espace de fonctions mesurables bornées  $\mathcal{M}_b(\mathbb{R}^d)$ , nous considérons un semi-groupe  $P_t$  possédant (au moins) une mesure de probabilité invariante  $\nu$  ;
- $\phi$  sera un noyau régularisant dans l'espace de Schwartz tel que

$$\int_{\mathbb{R}^d} \phi(x) dx = 1, \quad \int_{\mathbb{R}^d} y^{\beta_1} \phi(y) dy = 0 \text{ pour } |\beta_1| \geq 1, \quad \int_{\mathbb{R}^d} |y|^m |\partial^{\beta_2} \phi(y)| dy < \infty \text{ pour } m \in \mathbb{N}, |\beta_2| \geq 1$$

et, pour  $\delta > 0$ ,  $\phi_\delta(\cdot) := \frac{1}{\delta^d} \phi(\frac{\cdot}{\delta})$  et la convolution  $f_\delta := f \star \phi_\delta$ .

Les étapes de notre travail sont :

1. Nous considérons un schéma de type Euler à pas décroissant. Pour cela, on considère une famille décroissante ver 0 notée par  $(\gamma_n)_{n \geq 1}$  et  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . On fait l'hypothèse

$$(\Gamma) \quad \sum_{i=1}^{\infty} \gamma_i = \lim_{n \rightarrow \infty} \Gamma_n = \infty$$

et notons par  $\{\Gamma\} = \{\Gamma_n : n \in \mathbb{N}\}$ . Pour  $s, t \in \{\Gamma\}$ , le schéma d'Euler est

$$(16) \quad \bar{P}_{s,t} = \prod_{i=N(s)}^{N(t)-1} \bar{P}_{\gamma_i},$$

où  $\bar{P}_\gamma : C_b^\infty \rightarrow C_b^\infty$  est une opérateur tel que  $\|\bar{P}_\gamma \varphi\|_\infty \leq \|\varphi\|_\infty$  qui approxime  $P$  dans le sens

$$(17) \quad A(k_0, \alpha) \quad \|(P_\gamma - \bar{P}_\gamma)\varphi\|_\infty \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma^{1+\alpha}.$$

Ici,  $\alpha > 0$ ,  $k_0 \in \mathbb{N}$  ;

2. Le résultat principal de la partie **Proposition 2.1.1 dans la Partie III** est valide sous l'hypothèse que pour tous  $q, \kappa \in \mathbb{N}$  il existe une constante  $C_{q,\kappa,p}$  telle que

$$\begin{aligned} R_{\bar{P}}(p, \beta) & \quad \|\bar{P}_{t-1,t} P_{t,r} f - \bar{P}_{t-1,t} P_{t,r} f_\delta\|_\infty + \|\bar{P}_{t-1,t} \bar{P}_{t,r} f - \bar{P}_{t-1,t} \bar{P}_{t,r} f_\delta\|_\infty \\ & \leq C_{q,\kappa,p} \times \|f\|_\infty \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-p} \gamma_{N(t)}^{p\beta} + \eta^\kappa \right), \end{aligned}$$

pour tous  $\delta, \eta > 0$ , tous  $1 < t < r < t + 2$  et toute fonction  $f \in \mathcal{M}_b(\mathbb{R}^d)$ .

Si  $P_t$  satisfait une hypothèse de type Lipschitz exponentiel, une hypothèse de régularité et possède une mesure invariante  $\nu$  et si  $(R_{\bar{P}}(p, \beta))$  est valable, alors  $\nu$  est unique et, pour tout  $\varepsilon > 0$  et pour tout  $x \in \mathbb{R}^d$ , nous avons

$$(18) \quad d_{TV}(\bar{P}_{0,\Gamma_n}(x, \cdot), \nu) \leq C_\varepsilon (\gamma_n^{(p\beta) \wedge \alpha} - \varepsilon) + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\rho \Gamma_n}.$$

3. Dans la **Section 4 de la Partie III**, les résultats abstraits sont appliqués pour l'équation

$$(19) \quad X_t = x + \int_0^t b(X_r) dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{r-}) N(dz, dr).$$

- (a) Pour une partition  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{n-1} < \Gamma_n < \dots\}$  à pas de temps décroissant nous considérons le schéma d'Euler

$$X_t^{\mathcal{P}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}}) dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{\tau(r)-}^{\mathcal{P}}) N(dz, dr),$$

avec  $\tau(r) = \Gamma_k$  quand  $r \in [\Gamma_k, \Gamma_{k+1})$ .

- (b) Pour tout  $\gamma > 0$ , on définit la fonction de troncature  $M(\gamma) \in \mathbb{N}$  telle que  $\varepsilon_{M(\gamma)} \leq \gamma^2$  avec l'erreur  $\varepsilon_m$  le même comme (13). Pour  $\Gamma_n < t \leq \Gamma_{n+1}$ , on note  $M_{\mathcal{P}}(t) = M(\Gamma_{n+1} - \Gamma_n)$ . Nous réduisons l'argument au cas tronqué sur  $B_{M_{\mathcal{P}}(t)} := \{z \in \mathbb{R}^d : |z| \leq M_{\mathcal{P}}(t)\}$  :

$$(20) \quad X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \int_0^t \int_{B_{M_{\mathcal{P}}(r)}} c(z, X_{\tau(r)-}^{\mathcal{P}, M_{\mathcal{P}}}) N(dz, dr).$$

Cette équation peut être représentée par un processus de Poisson composé.

- (c) Le résultat théorique conduit au **Théorème 4.1, Partie III** donnant la distance suivante

$$d_{TV}(\mathcal{L}(X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}), \nu) \leq C_\varepsilon(\gamma_n^{1-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\frac{\theta}{2}\Gamma_n}),$$

avec  $\gamma_n = \Gamma_n - \Gamma_{n-1}$ . Ce résultat est à mettre en rapport avec [6] présentant la même vitesse mais une distance de Wasserstein, tandis que nous obtenons une distance en variation totale.

Concernant la stratégie de preuve et les hypothèses nécessaires, mentionnons les aspects suivants.

1. Afin d'utiliser une méthode d'intégration par parties, nous avons besoin (voir **Hypothesis 2.4, Partie III**) que  $\mu$  est absolument continue par rapport à la mesure de Lebesgue, que la dérivée de Radon-Nikodym est infiniment dérivable et que son logarithme l'est aussi et possède des dérivées bornées.
2. Des hypothèses de régularité et non-dégénérescence déjà mentionnées sont demandées dans **Hypothesis 2.1~2.3 Partie III**.
3. Nous travaillons sous une hypothèse de type Lyapunov  $LV \leq \bar{\beta} - \bar{\alpha}V$ , pour  $\bar{\alpha}; \bar{\beta} > 0$  pour le générateur infinitésimal de l'équation ainsi qu'une hypothèse coercitive relativement classique pour les coefficients (voir **Hypothesis 2.5 de la Partie III**). Ceci garantit l'existence de la mesure invariante  $\nu$ .

## References

- [1] A. Alfonsi, V. Bally : Construction of Boltzmann and McKean Vlasov type flows (the sewing lemma approach). arXiv:2105.12677 [math.PR](2021).
- [2] D. Bakry, I. Gentil, M. Ledoux: Analysis and geometry of Markov diffusion operators. *Grundlehren der Mathematischen Wissenschaften* **348**, Springer, MR-3155209 (2014).
- [3] V. Bally and L. Caramellino: On the distances between probability density functions. *Electron. J. Probab.* **19**, no. 110, 1-33. MR-3296526 (2014).
- [4] V. Bally and L. Caramellino: Asymptotic development for the CLT in total variation distance. *Bernoulli* **22**, 2442–2485. MR-3498034 (2016).
- [5] V. Bally, L. Caramellino, G. Poly: Non universality for the variance of the number of real roots of random trigonometric polynomials. *Probab. Th. Rel. Fields.* **174**, 887-927. MR-3980307 (2019).
- [6] P. Chen, C. Deng, René L. Schilling, et. al.: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. arXiv:2205.01342v2 [math.PR] (2023).
- [7] G. Pagès, F. Panloup: Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. <hal-03082311v3> (2022).

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Part I . . . . .	1
1.2	Part II . . . . .	6
1.3	Part III . . . . .	14
<b>2</b>	<b>Part I: Total variation distance between a jump-equation and its Gaussian approximation</b>	<b>23</b>
<b>3</b>	<b>Part II: Approximation schemes for McKean-Vlasov and Boltzmann type equations (error analysis in total variation distance)</b>	<b>59</b>
<b>4</b>	<b>Part III: Approximation for the invariant measure with applications for jump processes (convergence in total variation distance)</b>	<b>111</b>

# 1 Introduction

In this thesis, we consider the approximation of the solution of the stochastic differential equation with jumps and we focus on the convergence in total variation distance. The main method we use is the integration by parts technique in Malliavin calculus. This thesis contains three parts. In the following, we give an introduction of each of these three parts.

## 1.1 Part I

In the first part, we consider the one dimensional stochastic differential equation with jumps

$$X_t = x + \int_0^t \int_{(0,1]} c(s, z, X_{s-}) N_\mu(ds, dz) \quad (1)$$

where  $x \in \mathbb{R}$  is the initial value,  $t \in [0, T]$  with  $T$  the time horizon,  $\mu$  is a  $\sigma$ -finite measure on  $(0, 1]$ ,  $N_\mu$  is a Poisson point measure on  $(0, 1]$  with compensator  $\mu(dz)ds$  and  $c$  is a coefficient which verifies strong regularity hypotheses (see **Hypotheses 2.1-2.4** in Section 2.1 of Part I). We mention that we assume

$$\int_{\mathbb{R}^d} \sup_{s \in [0, T]} \sup_{x \in \mathbb{R}} |c(s, z, x)| \mu(dz) < \infty,$$

so we are in the finite variation case. The typical example that we have in mind is

$$\mu(dz) = \frac{dz}{z^{1+\rho}} 1_{\{z \in (0, 1]\}},$$

with  $\rho \in [0, 1)$ , so this is a truncated stable process - however, throughout the paper, we keep the general framework in which  $\mu$  is a measure which has infinite mass around zero. Our aim is to replace the "small jumps" (on the set  $\{z \leq \varepsilon\}$ ) by a stochastic integral with respect to a space-time Brownian motion: for any small  $\varepsilon > 0$ , we construct

$$\begin{aligned} X_t^\varepsilon &= x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, X_{s-}^\varepsilon) N_\mu(ds, dz) \\ &\quad + \int_0^t b_\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \int_{(0, \varepsilon]} c(s, z, X_s^\varepsilon) W_\mu(ds, dz), \end{aligned} \quad (2)$$

where  $W_\mu(ds, dz)$  is a space-time Brownian motion (in the sense of Walsh [21]) with covariance measure  $\mu(dz)ds$ , and the coefficient  $b_\varepsilon$  is defined by

$$b_\varepsilon(s, x) = \int_{(0, \varepsilon]} c(s, z, x) \mu(dz).$$

The interest of such approximation appears in various frameworks.

Our main motivation comes from numerical computations. If  $\mu(E) < \infty$ , then there are a finite number of jumps in any compact interval of time, so  $X_t$  may be represented by means of a compound Poisson process which may be explicitly simulated. But if  $\mu(E) = \infty$ , this is not possible anymore (except in very particular situations - see Talay and Protter [19] for example), and the "small jumps" should be truncated to revert to the case of a finite measure. This procedure is rather rough and gives large errors. In order to improve the approximation scheme, one may replace the "small jumps", namely those smaller than  $\varepsilon$ , by a stochastic integral with respect to  $W_\mu(ds, dz)$ . Note that the Poisson point measure  $dN_\mu$  is not compensated, which is why the drift corresponding to  $b_\varepsilon$  appears. This idea goes back to Asmussen and Rosinski [3]. In the case of stochastic differential equations driven by a Lévy process, Fournier [12] gives a precise estimate of the error and compares the approximation obtained just by truncating the small jumps



to the one obtained by adding a Gaussian noise as in (2). An enlightening discussion on the complexity of the two methods is also given. However, in that paper, the strong error is considered, while in our paper we discuss the weak error.

A second motivation comes from modelization problems in chemistry and biology. We are concerned by reactions which are naturally modelled by means of jump processes containing two regimes: one is very rapid but the jumps are small, and another is much slower and the jumps are larger – see for example [1], [2], [4], [10], [16], [17]. In this case, the regime corresponding to the rapid scale with small jumps may be modelled by a stochastic integral with respect to a Gaussian process and the slow regime with large jumps by a compound Poisson process. It may also be reasonable to consider an intermediary regime and this would be modelled by a drift term. This kind of model is exactly the equation (2).

A third motivation is given by a class of statistical problems (see [9], [11] and references therein), where a stochastic process is observed at various times and it should be decided whether its increments are due to small jumps or to a Gaussian component. In this framework, it is important to estimate the error in the sense of total variation distance. The authors explain that, if the error in total variation distance between the laws of  $X_t$  and of  $X_t^\varepsilon$  goes to zero, then there is no way to construct a statistical hypothesis test which decides if the noise comes from small jumps or from the Brownian motion. So, asymptotically, the two models contain the same information. This is a significant reason why we deal with the total variation distance.

Let us now discuss briefly our results and the relation to previously available estimates. If  $L_t$  (respectively  $L_t^\varepsilon$ ) represents the infinitesimal operator of  $X_t$  (respectively of  $X_t^\varepsilon$ ), then a development in Taylor series of order two gives

$$\|(L_t - L_t^\varepsilon)f\|_\infty \leq C \|f\|_{3,\infty} \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz), \quad (3)$$

where  $\hat{c}(z) := \sup_{s \leq T} \sup_{x \in \mathbb{R}} |c(s, z, x)|$  and  $\|f\|_{l,\infty} := \sum_{0 \leq i \leq l} \|f^{(i)}\|_\infty$ . Then a Trotter-Kato type argument (or Lindeberg method) yields

$$\sup_{s \leq T} \|(P_t - P_t^\varepsilon)f\|_\infty \leq C \|f\|_{3,\infty} \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz) \quad (4)$$

where  $P_t$  (respectively of  $P_t^\varepsilon$ ) represents the semigroup of  $X_t$  (respectively of  $X_t^\varepsilon$ ).

The drawback of the above estimate is that the bound on the error involves  $\|f\|_{3,\infty}$ , so it only applies to smooth test functions. The main contribution of our paper is to replace  $\|f\|_{3,\infty}$  by  $\|f\|_\infty$ , so as to prove the convergence in total variation distance: for any  $\delta > 0$ , we have

$$d_{TV}(X_t, X_t^\varepsilon) \leq C_\delta \left( \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz) \right)^{1-\delta}, \quad (5)$$

where  $d_{TV}$  denotes the total variation distance between two random variables  $F, G : \Omega \rightarrow \mathbb{R}$ :

$$d_{TV}(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \|f\|_\infty \leq 1\}.$$

Moreover, under these hypotheses, we prove that  $\mathbb{P}(X_t(x) \in dy) = p_t(x, y)dy$  and  $\mathbb{P}(X_t^\varepsilon(x) \in dy) = p_t^\varepsilon(x, y)dy$  with smooth densities  $y \mapsto p_t(x, y)$  and  $y \mapsto p_t^\varepsilon(x, y)$ . And, for every  $k$  and every  $\delta > 0$ , we obtain

$$\|\partial_y^k p_t - \partial_y^k p_t^\varepsilon\|_\infty \leq C_{k,\delta} \left( \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz) \right)^{1-\delta}. \quad (6)$$

This proves that  $p_t^\varepsilon$  converges to  $p_t$  in distribution norms as  $\varepsilon \rightarrow 0$ . We stress that all these estimates are non asymptotic (hold for every fixed  $\varepsilon > 0$ ). We also mention that the estimation technique that we use has a drawback: we loose systematically some  $\delta > 0$  in the speed of convergence.

One may also consider an approximate equation obtained just by discarding the small jumps:

$$\widetilde{X}_t^\varepsilon = x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, \widetilde{X}_{s-}^\varepsilon) N_\mu(ds, dz). \quad (7)$$

Then  $\|(L_t - \widetilde{L}_t^\varepsilon)f\|_\infty \leq \|f\|_{1,\infty} \int_{(0,\varepsilon]} |\hat{c}(z)| \mu(dz)$ , where  $\widetilde{L}_t^\varepsilon$  is the infinitesimal operator of  $\widetilde{X}_t^\varepsilon$ . So the Trotter-Kato method gives

$$\|P_t f - \widetilde{P}_t^\varepsilon f\|_\infty \leq C \|f\|_{1,\infty} \times \int_{(0,\varepsilon]} |\hat{c}(z)| \mu(dz). \quad (8)$$

The gain in (5) is that we have  $\int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz)$  instead of  $\int_{(0,\varepsilon]} |\hat{c}(z)| \mu(dz)$  in (8), which means that we have a faster speed of convergence. So (2) is a better approximation scheme than (7).

We provide an example to illustrate our results. For  $t \in [0, T]$ , we consider the following stochastic differential equation driven by a Lévy process:

$$X_t = x + \int_0^t \sigma(X_{s-}) dZ_s, \quad (9)$$

where  $(Z_t)_{t \in [0, T]}$  is a pure jump Lévy process with intensity measure

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

We approximate (9) by

$$X_t^\varepsilon = x + \int_0^t \sigma(X_{s-}^\varepsilon) dZ_s^\varepsilon + b(\varepsilon) \int_0^t \sigma(X_s^\varepsilon) ds + c(\varepsilon) \int_0^t \sigma(X_s^\varepsilon) dB_s,$$

where  $(Z_t^\varepsilon)_{t \in [0, T]}$  is a pure jump Lévy process with intensity measure  $\mathbb{1}_{(\varepsilon,1]}(z) \mu(dz)$ ,  $(B_t)_{t \in [0, T]}$  is a standard Brownian motion independent of  $(Z_t^\varepsilon)_{t \in [0, T]}$ , and

$$b(\varepsilon) = \int_{(0,\varepsilon]} z \mu(dz), \quad c(\varepsilon) = \sqrt{\int_{(0,\varepsilon]} z^2 \mu(dz)}.$$

We denote  $C_b^\infty(\mathbb{R})$  to be the space of smooth functions with bounded derivatives of any order. We assume that  $\sigma \in C_b^\infty(\mathbb{R})$ ,  $\inf_{x \in \mathbb{R}} \sigma(x) > 0$  and  $\inf_{x \in \mathbb{R}} \sigma'(x) > -1$ , where  $\sigma'$  is the differential of  $\sigma$  in  $x$ . Then for any  $\delta > 0$ , there is a constant  $C > 0$  such that for any  $t \in [0, T]$ ,

$$d_{TV}(X_t, X_t^\varepsilon) \leq C \varepsilon^{3-\rho-\delta}.$$

Moreover, the laws of  $X_t$  and  $X_t^\varepsilon$  have smooth densities  $p_{X_t}(x)$  and  $p_{X_t^\varepsilon}(x)$  respectively. And for any index  $l$  and any  $\delta > 0$ , there exists a constant  $C > 0$  such that

$$\|p_{X_t} - p_{X_t^\varepsilon}\|_{l,\infty} \leq C \varepsilon^{3-\rho-\delta}.$$

Now we explain the method we use in order to obtain our main results (5) and (6). Our approach uses a strategy based on the integration by parts technique (an abstract Malliavin calculus) developed in [6]. We introduce an abstract integration by parts framework (in Section 3 of Part I) built on a particular case of the Dirichlet form theory (see [65] and [6]). This technique is very similar to the standard Malliavin calculus but is presented in a more general framework which goes far beyond the sole case of the Wiener space. In particular, we aim at providing a minimalist setting leading to our estimates in total variation distance. Our unified framework includes the standard Malliavin calculus and different known versions: the calculus based on the splitting method developed and used in [66], [67], [68] as well as the  $\Gamma$ -calculus in [65]. We also mention that our approach applies in the case of the Malliavin calculus for jump processes

as settled by Bichteler, Gravereaux, Jacod [8] and in the "lent particle" approach for Poisson point measures developed by [69].

One main consequence of the abstract integration by parts technique is as follows. We define the distances between random variables  $F, G : \Omega \rightarrow \mathbb{R}$ :

$$d_k(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \sum_{0 \leq i \leq k} \|f^{(i)}\|_\infty \leq 1\}.$$

For  $k = 1$ , this is the Fortet Mourier distance (which is a variant of the Wasserstein distance), while for  $k = 0$ , this is the total variation distance and we denote it by  $d_{TV}$ . We use Malliavin calculus for some random variables  $F$  and  $G$ . We denote  $C_p^\infty(\mathbb{R}^d)$  to be the space of smooth functions which, together with all the derivatives, have polynomial growth. We associate a derivative operator  $D$  with value in a Hilbert space  $\mathcal{H}$  satisfying the chain rule: for every  $\phi \in C_p^\infty(\mathbb{R}^d)$  and  $F = (F_1, \dots, F_d)$ , we have

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

And we associate an Ornstein-Uhlenbeck operator  $L$  verifying the duality formula:

$$\mathbb{E}\langle DF, DG \rangle_{\mathcal{H}} = \mathbb{E}[FL(G)] = \mathbb{E}[GL(F)]. \quad (10)$$

We define by iteration the derivative operators of higher orders  $D^q$  and we define the Malliavin-Sobolev norms by

$$|F|_l = |F| + \sum_{q=1}^l |D^q F|_{\mathcal{H}^{\otimes q}}, \quad \|F\|_{L,l,p} = (\mathbb{E}|F|_l^p)^{1/p} + (\mathbb{E}|LF|_{l-2}^p)^{1/p}. \quad (11)$$

For  $F = (F^1, \dots, F^d)$  we denote the Malliavin covariance matrix  $\sigma_F = (\sigma_F^{i,j})_{i,j=1,\dots,d}$  with  $\sigma_F^{i,j} = \langle DF_i, DF_j \rangle_{\mathcal{H}}$ . And we denote

$$\Sigma_p(F) = \mathbb{E}(1/\det \sigma_F)^p. \quad (12)$$

We say that  $F$  is non-degenerated if  $\Sigma_p(F) < \infty$ ,  $\forall p \geq 1$ . Now we suppose that the Malliavin-Sobolev norms of  $F$  and  $G$  are bounded (i.e.  $\|F\|_{L,l,p} + \|G\|_{L,l,p} < \infty$ ,  $\forall l, p \geq 1$ ), and the Malliavin covariance matrix of  $F$  and  $G$  are non-degenerated (i.e.  $\Sigma_p(F) + \Sigma_p(G) < \infty$ ,  $\forall p \geq 1$ ). We state in **Lemma 3.4** of Part I the followings. We take some  $r \in \mathbb{N}$ . For any  $\delta > 0$ , we have

$$i) \quad d_{TV}(F, G) \leq C \times d_r(F, G)^{1-\delta}. \quad (13)$$

Moreover, the law of  $F$  (resp.  $G$ ) has a smooth density  $p_F(x)$  (resp.  $p_G(x)$ ), and for any  $l \in \mathbb{N}$  we have

$$ii) \quad \|p_F - p_G\|_{l,\infty} \leq C \times d_r(F, G)^{1-\delta}. \quad (14)$$

This is the key step in order to obtain our main results (5) and (6). The significance is the following. Suppose that one has already obtained an estimate of a "smooth" distance  $d_r$  between two random vectors  $F$  and  $G$  (in our case  $r = 3$  in (4)). But we would like to control the total variation distance between them (to obtain (5) in our case). In order to do this, one can employ (13) and (14) and conclude the following. We need to assume that the Malliavin-Sobolev norms of  $F$  and  $G$  are bounded ( $\|F\|_{L,l,p} + \|G\|_{L,l,p} < \infty$ ), and the Malliavin covariance matrix of  $F$  and  $G$  are non-degenerated ( $\Sigma_p(F) + \Sigma_p(G) < \infty$ ). Then (13) asserts that one may control  $d_{TV}$  by  $d_r$ , and the control is quasi optimal: we loose just a power  $\delta > 0$  which we may take as small as we want. And (14) says that we may also control the distance between the derivatives of density functions by  $d_r$ .

Now we need to apply (13) and (14) to jump processes so that we can obtain (5) and (6). So we will use Malliavin calculus for jump processes and define the operators  $D$  and  $L$ . For some technical reasons, we

make the change of variable  $z \mapsto \frac{1}{z}$ . Let  $\theta : (0, 1] \rightarrow [1, \infty)$  be a function such that  $\theta(z) = \frac{1}{z}$ . By a change of variables, instead of dealing with equation (2), it is equivalent to consider the following equation.

$$\begin{aligned}\widehat{X}_t^M &= x + \int_0^t \int_{[1, M)} \widetilde{c}(s, z, \widehat{X}_{s-}^M) N_\nu(ds, dz) \\ &\quad + \int_0^t b_M(s, \widehat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \widetilde{c}(s, z, \widehat{X}_s^M) W_\nu(ds, dz),\end{aligned}\tag{15}$$

where  $M = \frac{1}{\varepsilon}$ ,  $\nu(dz) = \mu \circ \theta^{-1}(dz)$ ,  $\widetilde{c}(s, z, x) = c(s, \frac{1}{z}, x)$ ,

$$b_M(s, x) = \int_{\{z \geq M\}} \widetilde{c}(s, z, x) \nu(dz),\tag{16}$$

$N_\nu$  is a Poisson point measure with intensity measure  $\nu(dz)ds$  and  $W_\nu$  is the space-time Brownian motion with covariance measure  $\nu(dz)ds$ . One can check that  $\widehat{X}_t^M$  has the same law as  $X_t^\varepsilon$ . We notice that in (15) the intensity measure is  $\mathbb{1}_{[1, M)}(z)\nu(dz)ds$  and this is a finite measure. Then the corresponding Poisson Point measure  $N_\nu$  may be represented by means of a compound Poisson process. We produce the representation on each set  $\{z \in I_k = [k, k+1)\}$ ,  $k \in \mathbb{N}$ , so the equation (15) reads

$$\begin{aligned}\widehat{X}_t^M &= x + \int_0^t \sum_{k=1}^{M-1} \int_{\{z \in I_k\}} \widetilde{c}(s, z, \widehat{X}_{s-}^M) N_\nu(ds, dz) \\ &\quad + \int_0^t b_M(s, \widehat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \widetilde{c}(s, z, \widehat{X}_s^M) W_\nu(ds, dz) \\ &= x + \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \widetilde{c}(T_i^k, Z_i^k, \widehat{X}_{T_i^k-}^M) \\ &\quad + \int_0^t b_M(s, \widehat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \widetilde{c}(s, z, \widehat{X}_s^M) W_\nu(ds, dz).\end{aligned}$$

Here  $T_i^k, k, i \in \mathbb{N}$  are the jump times of the Poisson process  $(J_t^k)_{t \in [0, T]}$  of parameter  $\nu(I_k)$ , and  $Z_i^k, k, i \in \mathbb{N}$  are independent random variables of law  $\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{\nu(I_k)}$ , which are independent of  $J_T^k$  as well. We remark that we produce the representation by compound Poisson process on each set  $I_k$  instead of on the whole interval  $[1, M)$ . This is necessary in order to give an accurate representation of the intensity measure  $\nu(dz)$ . And this is important in order to prove that we have sufficient noise which gives the non-degeneracy condition of the Malliavin covariance matrix (see Section 5.2 of Part I for details).

We will work conditionally to the jump times  $T_i^k, k, i \in \mathbb{N}$ , so the randomness in the system comes from  $W_\nu$  on one hand and from  $Z_i^k, k, i \in \mathbb{N}$  on the other hand. Concerning  $W_\nu$  we will use the standard Malliavin calculus (which fits in the abstract framework presented in Section 3 of Part I). But we will also use this integration by parts calculus with respect to the amplitude of the jumps given by  $Z_i^k, k, i \in \mathbb{N}$ . We present this kind of calculus now. Suppose for a moment (just for simplicity) the law of  $Z_i^k$  is absolutely continuous with respect to the Lebesgue measure and has a smooth density  $h_k(z)$  which has compact support. We also assume that the logarithm of the density  $\ln h_k$  is a smooth function. Then we look to  $\widehat{X}_t^M$  as to a functional  $F(Z_1^1, \dots, Z_{J_t^k}^{M-1})$  and we define the derivative operator

$$D_{k,i}^Z F = \frac{\partial}{\partial z_i^k} F(Z_1^1, \dots, Z_{J_t^k}^{M-1})$$

and the Ornstein-Uhlenbeck operator

$$L^Z F = - \sum_{k,i} D_{k,i}^Z D_{k,i}^Z F + D_{k,i}^Z F \times \partial_z \ln h_k(Z_i^k)$$

which verify the duality formula (10). Since we want to use integration by parts with respect to both  $W_\nu$  and  $Z_i^k, k, i \in \mathbb{N}$ , we will consider the derivative operator  $D = (D^W, D^Z)$  and the operator  $L = (L^W, L^Z)$

where  $D^W$  and  $L^W$  are the derivative and Ornstein Uhlenbeck operators from the standard Malliavin calculus for Gaussian random variables (see [18]). With these operators at hand, we are able to apply (13) and (14) with the Sobolev norms associated to the operators  $D = (D^W, D^Z)$  and  $L = (L^W, L^Z)$  and the covariance matrix associated to the operator  $D = (D^W, D^Z)$ .

Roughly speaking this is our strategy. But there is one more point: the hypotheses we raise for the law of  $Z_i^k$  that it has a smooth density with compact support and has a smooth logarithm density, is rather strong and we want to weaken it. This is the aim of the "splitting method" that we present now. The law of  $Z_i^k$ , which is  $\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{\nu(I_k)}$ , could be very irregular and it is not possible to make integration by parts based on it. In order to overcome this difficulty, we suppose that the law of  $Z_i^k$  is lower bounded by the Lebesgue measure (the Doeblin condition): for every  $k \in \mathbb{N}$ , there exists  $\varepsilon_k > 0$  such that

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{\nu(I_k)} \geq \mathbb{1}_{I_k}(z) \varepsilon_k \times dz.$$

We denote  $C_c^\infty(\mathbb{R})$  the space of smooth functions with compact support. We define a regularization function  $\psi \in C_c^\infty(\mathbb{R})$  with support included in  $[-\frac{1}{2}, \frac{1}{2}]$  such that the logarithm of  $\phi$  is also smooth, and we denote  $m(\psi) = \int_{-1/2}^{1/2} \psi(y) dy$ . We choose  $\varepsilon_k < 1/m(\psi)$  so that  $1 - \varepsilon_k m(\psi) > 0$ . Then we produce three independent random variables  $V_i^k, U_i^k$  and  $\xi_i^k$  such that  $Z_i^k$  has the same law as  $\xi_i^k V_i^k + (1 - \xi_i^k) U_i^k$ :

$$Z_i^k \sim \xi_i^k V_i^k + (1 - \xi_i^k) U_i^k,$$

where  $\xi_i^k$  is a Bernoulli random variable with law

$$\mathbb{P}(\xi_i^k = 1) = \varepsilon_k m(\psi), \quad \mathbb{P}(\xi_i^k = 0) = 1 - \varepsilon_k m(\psi),$$

$V_i^k$  is a random variable with law

$$\mathbb{P}(V_i^k \in dz) = \frac{1}{m(\psi)} \psi(z - (k + \frac{1}{2})) dz$$

which has good regularization properties, and  $U_i^k$  is a random variable with law

$$\mathbb{P}(U_i^k \in dz) = \frac{1}{1 - \varepsilon_k m(\psi)} (\mathbb{P}(Z^k \in dz) - \varepsilon_k \psi(z - (k + \frac{1}{2})) dz).$$

So we split  $Z_i^k$  into two parts,  $V_i^k$  and  $U_i^k$ . We notice that we do it in such a way that the law of  $V_i^k$  has a smooth density with compact support and has a smooth logarithm density. Then we perform the Malliavin calculus with respect to  $V_i^k$  instead of  $Z_i^k$  and we work conditionally to  $\xi_i^k$  and  $U_i^k$  which appear as constants. This is the so-called splitting method. We refer to Section 4.1 of Part I for more details.

The splitting method presented here is analogous to the one in [8]. Therein, Bichteler, Gravereau and Jacod deal with 2 kinds of independent Poisson point measures. One is very regular, and smooth enough to make Malliavin calculus on it (in our paper,  $V_i^k$  play the same role). The other one can be arbitrary, and it may be very irregular (in our paper, it corresponds to  $U_i^k$ ). But the difference is that instead of splitting the Poisson point measure, we split the random variables, and so this method can also be applied in a large class of different problems. For example, Bally, Caramellino and Poly use the splitting method to show the convergence in total variation distance in the central limit theorem in [5]. Other possible approaches to the Malliavin calculus for jump processes are given in the papers [7], [14], [15], [20], [22] and the book [13] for example.

## 1.2 Part II

In the second part, we deal with the McKean-Vlasov and Boltzmann jump equation. To begin, we introduce some notations. We give a time horizon  $T > 0$  and let  $0 < t \leq T$ . For a multi-index  $\beta$ , we denote

$|\beta|$  to be the length of  $\beta$ . 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 denote the norm  $\|f\|_{l,\infty} := \sum_{|\beta| \leq l} \|\partial^\beta f\|_\infty$  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 (of order 1)  $W_1$  by

$$W_1(\rho_1, \rho_2) = \sup_{Lip(f) \leq 1} \left| \int_{\mathbb{R}^d} f(x) \rho_1(dx) - \int_{\mathbb{R}^d} f(x) \rho_2(dx) \right|, \quad (17)$$

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

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

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  $G$ ).

We consider a  $d$ -dimensional McKean-Vlasov and Boltzmann type jump equation as follows.

$$X_t = X_0 + \int_0^t b(r, X_r, \rho_r) dr + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(r, v, z, X_{r-}, \rho_{r-}) N_{\rho_{r-}}(dv, dz, dr), \quad (18)$$

where

$$\rho_t(dv) = \mathbb{P}(X_t \in dv)$$

is the law of  $X_t$ ,  $N_{\rho_t}$  is a Poisson point measure on the state space  $\mathbb{R}^d \times \mathbb{R}^d$  with intensity measure

$$\rho_t(dv) \mu(dz) dr,$$

$\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ ,  $X_0$  is the initial random variable independent of the Poisson point measure  $N_{\rho_t}$ , and  $b, c$  are functions which verify some regularity and ellipticity conditions (see **Hypotheses 2.1~2.4** in Section 2.2 of Part II for precise statements). In particular, we assume that for every multi-indices  $\beta_1, \beta_2$ , there exists a non-negative function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that for every  $z \in \mathbb{R}^d$ , we have

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

with  $\int_{\mathbb{R}^d} |\bar{c}(z)|^p \mu(dz) < \infty$ ,  $\forall p \geq 1$ . We also assume that for any  $r \in [0, T]$ ,  $v_1, v_2, x \in \mathbb{R}^d$ ,  $z \in \mathbb{R}^d$ ,  $\rho_1, \rho_2 \in \mathcal{P}_1(\mathbb{R}^d)$ , we have the following Lipschitz property:

$$|c(r, v_1, z, x, \rho_1) - c(r, v_2, z, x, \rho_2)| \leq \bar{c}(z)(|v_1 - v_2| + W_1(\rho_1, \rho_2)). \quad (19)$$

Moreover, we suppose that there exists a non-negative function  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  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

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

We remark that we use the notations from [43] and we refer to [28], [34], [43], [49], [50], [56] and [57] 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 which appears in (1) of Part I for example) 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 [41] for a mathematical approach to this kind of equation and see [27], [31], [40], [42], [54], [55] and [59] for the approximation schemes of a McKean-Vlasov equation. Moreover, the intensity of the Poisson point measure  $N_{\rho_t}$  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 [60], [61],

and followed by many others in [29], [35], [36], [37], [51], [53] and [58] for example. One can also see [24] and [62] for the analytical Boltzmann equation and [33] for the physical background. Recently, there is also some work on inhomogeneous Boltzmann equations (see for instance [23], [38] and [39]). We have to mention however that our equation (18) 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  $\gamma(z, x)\mu(dz)$  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 [36] for example).

Now we construct the Euler scheme. For any partition  $\mathcal{P} = \{0 = r_0 < r_1 < \dots < r_{n-1} < r_n = T\}$  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

$$\begin{aligned} X_t^{\mathcal{P}} &= X_0 + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P}}, \rho_{\tau(r)}^{\mathcal{P}}) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P}}, \rho_{\tau(r)-}^{\mathcal{P}}) N_{\rho_{\tau(r)-}^{\mathcal{P}}} (dv, dz, dr), \end{aligned} \quad (20)$$

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 [30], [26], [44], [45], [46], [47], [48], [19] and the references therein. We denote  $C_b^1(\mathbb{R}^d)$  to be the space of differentiable and bounded functions with bounded derivatives. For the equation (18), Alfonsi and Bally [25] have proved recently that under some regularity conditions on the coefficients  $b$  and  $c$ , the solution of the equation (18) exists and is unique, and  $X_t$  is the probabilistic interpretation of the following analytical weak equation.

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

Moreover, [25] has proved that the Euler scheme  $X_t^{\mathcal{P}}$  (see (20)) converges to  $X_t$  in Wasserstein distance (of order 1)  $W_1$ :

$$W_1(X_t^{\mathcal{P}}, X_t) \leq C|\mathcal{P}| \rightarrow 0,$$

as  $|\mathcal{P}| \rightarrow 0$ , with

$$|\mathcal{P}| := \max_{k \in \{0, \dots, n-1\}} (r_{k+1} - r_k).$$

In our paper, under supplementary hypotheses, we prove a stronger result. We prove (see **Theorem 2.1** of Part II) 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

$$d_{TV}(X_t^{\mathcal{P}}, X_t) \leq C|\mathcal{P}|^{1-\varepsilon} \rightarrow 0, \quad (22)$$

as  $|\mathcal{P}| \rightarrow 0$ . 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 (21).

Since we have infinite numbers of jumps (due to **Hypothesis 2.4** in Section 2.2 of Part II), 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  $B_M = \{z \in \mathbb{R}^d : |z| \leq M\}$  and we assume that  $\mu(B_M) < \infty$ ,  $\forall M \geq 1$ . In some concrete example, the mass of  $\mu$  may be infinite around  $z = 0$ :  $\mu(|z| < \varepsilon) = \infty$ ,  $\forall \varepsilon > 0$ . In this case, one has to make a change of variable of type  $|z| \rightarrow 1/|z|$  to come back in our framework (as it is done in Part I). For  $M \in \mathbb{N}$ , we denote

$$c_M(r, v, z, x, \rho) := c(r, v, z, x, \rho) \mathbb{1}_{B_M}(z)$$

and

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

Now we discard the jumps of size  $|z| > M$  and we replace them by a Gaussian random variable:

$$\begin{aligned} X_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M}) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) N_{\rho_{\tau(r)-}^{\mathcal{P},M}}(dv, dz, dr), \end{aligned} \quad (23)$$

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  $\rho_t^{\mathcal{P},M}(dv) \mu(dz) dr$ ,  $\Delta$  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** of Part II) 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

$$d_{TV}(X_t^{\mathcal{P},M}, X_t) \leq C(\sqrt{\varepsilon_M} + |\mathcal{P}|)^{1-\varepsilon} \rightarrow 0, \quad (24)$$

as  $|\mathcal{P}| \rightarrow 0$  and  $M \rightarrow \infty$ , with

$$\varepsilon_M := \int_{\{|z|>M\}} |\bar{c}(z)|^2 \mu(dz) + \left| \int_{\{|z|>M\}} \bar{c}(z) \mu(dz) \right|^2.$$

Moreover, the law of  $X_t^{\mathcal{P},M}$  has a smooth density. We mention that the result (24) is non asymptotic: it holds for every partition  $\mathcal{P}$  and every  $M$ . We stress that there is an analogy between the construction of (15) in Part I and the construction of the equation (23) here: in both cases, we cancel the "big jumps" (of infinite mass) and we replace them by a Gaussian random variable. And in both cases, this is important in order to obtain the non degeneracy of the Malliavin covariance matrix. But in Part I, this also gives an improvement on the speed of convergence up to  $\int_{\{|z|>M\}} \sup_{s,x} |\bar{c}(s, z, x)|^3 \nu(dz)$  while here we keep  $(\int_{\{|z|>M\}} |\bar{c}(z)|^2 \mu(dz) + |\int_{\{|z|>M\}} \bar{c}(z) \mu(dz)|^2)^{\frac{1}{2}}$ . This is because here, the framework is more complex and we are not able to do the detailed analysis based on the Taylor expansion as in (3) and (4).

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 fix  $N$  (the number of particles) and we take an initial vector  $(X_0^1, \dots, X_0^N)$  with components which are independent and identically distributed with common law  $\rho_0$  (which is the law of  $X_0$ ), and  $(\Delta^1, \dots, \Delta^N)$  which is a  $N \times d$ -dimensional standard Gaussian random variable independent of  $(X_0^1, \dots, X_0^N)$ . Then we construct the particle system  $\vec{X}_t^{\mathcal{P},M} = (X_t^{\mathcal{P},M,1}, \dots, X_t^{\mathcal{P},M,N})$ :

$$\begin{aligned} X_t^{\mathcal{P},M,i} &= X_0^i + a_T^M \Delta^i + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(r)}^{\mathcal{P},M})) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(r)-}^{\mathcal{P},M})) N_{\hat{\rho}(\vec{X}_{\tau(r)-}^{\mathcal{P},M})}^i(dv, dz, dr), \quad i = 1, \dots, N, \end{aligned} \quad (25)$$

where

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

is the empirical measure of  $\vec{X}_t^{\mathcal{P},M}$  (with  $\delta_x(dv)$  denoting the Dirac measure),  $N_{\hat{\rho}(\vec{X}_t^{\mathcal{P},M})}^i(dv, dz, dr)$ ,  $i = 1, \dots, N$  are Poisson point measures that are independent each other conditionally to  $\vec{X}_t^{\mathcal{P},M}$  and independent of  $(X_0^1, \dots, X_0^N, \Delta^1, \dots, \Delta^N)$  with intensity  $\hat{\rho}(\vec{X}_t^{\mathcal{P},M})(dv) \mu(dz) dr$ .



Then we represent the jump's parts of the equation (25) by compound Poisson processes in order to give an explicit scheme of simulation. We denote  $I_1 = B_1$ ,  $I_k = B_k \setminus B_{k-1}$  for  $k \geq 2$  and we will produce the representation on each  $I_k$ . For  $i = 1, \dots, N$ ,  $k \in \mathbb{N}$ , we take  $(J_t^{k,i})_{t \in [0,T]}$  a Poisson process of intensity  $\mu(I_k)$ . We denote by  $(T_l^{k,i})_{l \in \mathbb{N}}$  the jump times of  $(J_t^{k,i})_{t \in [0,T]}$  and we consider some sequences of independent random variables  $Z_l^{k,i} \sim \mathbb{1}_{I_k}(z) \frac{\mu(dz)}{\mu(I_k)}$  and  $U_l^{k,i}$  uniformly distributed on  $\{1, \dots, N\}$ , for all  $i = 1, \dots, N$ ,  $k, l \in \mathbb{N}$ . Moreover,  $(J_t^{k,i})_{t \in [0,T]}$ ,  $Z_l^{k,i}$ ,  $U_l^{k,i}$ ,  $\Delta^i$ ,  $X_0^i$ ,  $i = 1, \dots, N$ ,  $k, l \in \mathbb{N}$  are taken to be independent. Then we write

$$\begin{aligned} X_t^{\mathcal{P},M,i} &= X_0^i + a_T^M \Delta^i + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M,i}, \tilde{\rho}(\vec{X}_{\tau(r)}^{\mathcal{P},M})) dr \\ &+ \sum_{k=1}^M \sum_{l=1}^{J_t^{k,i}} c(\tau(T_l^{k,i}), X_{\tau(T_l^{k,i})-}^{\mathcal{P},M,U_l^{k,i}}, Z_l^{k,i}, X_{\tau(T_l^{k,i})-}^{\mathcal{P},M,i}, \tilde{\rho}(\vec{X}_{\tau(T_l^{k,i})-}^{\mathcal{P},M})). \end{aligned}$$

It is clear that now  $\vec{X}_t^{\mathcal{P},M}$  is simulated in an explicit way. We also remark that we construct the representation by compound Poisson process on each set  $I_k$  instead of on the entire ball  $B_M$  in order to give an accurate description of the noise produced by  $Z_l^{k,i}$ . This is necessary in order to have sufficient noise which gives the non-degeneracy condition of the Malliavin covariance matrix (see Section 4.2 of Part II for details).

Given the dimension  $d$ , 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

$$\varphi(x) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{|x|^2}{2}}, \quad \varphi_\delta(x) = \frac{1}{\delta^d} \varphi\left(\frac{x}{\delta}\right), \quad 0 < \delta \leq 1. \quad (26)$$

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

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

Then we have

$$p_t(x) = \frac{1}{N} \sum_{i=1}^N \mathbb{E} \varphi_\delta(X_t^{\mathcal{P},M,i} - x) + \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{2}{d+3}}), \quad (27)$$

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

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

then we get moreover by Romberg method that

$$p_t(x) = \frac{2}{N} \sum_{i=1}^N \mathbb{E} \varphi_{\delta/\sqrt{2}}(X_t^{\mathcal{P},M,i} - x) - \frac{1}{N} \sum_{i=1}^N \mathbb{E} \varphi_\delta(X_t^{\mathcal{P},M,i} - x) + \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{\frac{4}{d+5}}). \quad (28)$$

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** of Part II, 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  $\tilde{\Delta}$  a  $d$ -dimensional standard Gaussian random variable independent of  $\vec{X}_t^{\mathcal{P},M}$ . For any  $\varepsilon > 0$ , we take

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

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

$$\int_{\mathbb{R}^d} f(x) p_t(x) dx = \frac{1}{N} \sum_{i=1}^N \mathbb{E} f(X_t^{\mathcal{P},M,i} + \delta \tilde{\Delta}) + \|f\|_{\infty} \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}). \quad (29)$$

We stress that the speed of convergence in (29) (which is  $1 - \varepsilon$ ) no longer depends on the dimension  $d$ , so it still behaves well for large dimension. However, the number of particles  $N$  depends on  $d$  because  $V_N \leq (|\mathcal{P}| + \varepsilon_M)^{\frac{d+3}{2}(1-\varepsilon'')}$  so that  $1/N \sim (|\mathcal{P}| + \varepsilon_M)^{d^2}$  when  $d$  is large. We also notice that the speed of convergence in (29) is the same as in (24) for the truncated Euler scheme. Moreover, for any  $\varepsilon > 0$ , we take

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

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

$$\int_{\mathbb{R}^d} f(x) p_t(x) dx = \frac{2}{N} \sum_{i=1}^N \mathbb{E} f(X_t^{\mathcal{P},M,i} + \frac{\delta}{\sqrt{2}} \tilde{\Delta}) - \frac{1}{N} \sum_{i=1}^N \mathbb{E} f(X_t^{\mathcal{P},M,i} + \delta \tilde{\Delta}) + \|f\|_{\infty} \times \mathcal{O}((|\mathcal{P}| + \sqrt{\varepsilon_M})^{1-\varepsilon}). \quad (30)$$

We remark that (30) is even a better simulation scheme than (29) in the sense that the numbers of particles  $N$  is smaller than the one in (29) and  $\delta$  is larger than the one in (29).

We give now a general view on the strategy used in Part II. Notice that the Poisson process which appears in the equation (18) has intensity  $\mu(dz)$  which is an infinite measure. It is convenient, from the point of view of Malliavin calculus, to introduce an intermediary equation driven by a Poisson point measure with intensity  $\mathbb{1}_{\{|z| \leq M\}} \mu(dz)$  which is a finite measure. We will use an intermediary equation. We denote by  $X_t^M$  the solution of this equation:

$$\begin{aligned} X_t^M &= X_0 + a_T^M \Delta + \int_0^t b(r, X_r^M, \rho_r) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathbb{1}_{B_M}(z) c(r, v, z, X_{r-}^M, \rho_{r-}) N_{\rho_{r-}}(dv, dz, dr), \end{aligned} \quad (31)$$

where  $\rho_r$  is the law of  $X_r$  (not of  $X_r^M$ ) and  $N_{\rho_t}(dv, dz, dr)$  is a Poisson point measure with intensity measure  $\rho_t(dv) \mu(dz) dr$  as in (18). Since  $X_t^M$  depends only on a finite number of jumps in any compact time interval, we can use Malliavin calculus for the amplitudes of jumps. We also replace the jumps larger than  $M$  (which have been canceled) by a Gaussian noise as in (23) - 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**) of Part II) that the measure  $\mu$  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. We first check that the Malliavin-Sobolev norms of  $X_t^M$  are bounded uniformly with respect to  $M$ . Using the convergence  $X_t^M \rightarrow X_t$  in  $L^1$ , we are able to prove that  $X_t$  is smooth in the sense of Malliavin calculus for jump processes and has bounded Malliavin-Sobolev norms. 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 the law of  $X_t$  and the density  $p_t$ . To do it, we consider the truncated Euler scheme  $X_t^{\mathcal{P},M}$  (see

(23)) and we focus on three equations (18), (31) and (23) with solutions  $X_t, X_t^M$  and  $X_t^{\mathcal{P},M}$  respectively. There is a supplementary difficulty which appears here: the Poisson point measures which govern the equations (31) and (23) have an intensity which depends on the law of the solution of each of these equations - so they are on different probability spaces. It is convenient to use similar equations driven by the same Poisson point measure (common to the three equations) such that we can use an  $L^2$  calculus. 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. The coupling procedure is as follows. We notice that our basic distance is the Wasserstein distance (of order 1)  $W_1$  (see (19)), however we need the distance  $W_{2+\varepsilon_*}$  (defined immediately below) for some small  $\varepsilon_* > 0$  because we need  $L^2$  estimate later in (37) 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

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

where  $\Pi(\rho_1, \rho_2)$  is the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\rho_1$  and  $\rho_2$ . Some basic properties of  $W_p, p \geq 1$  can be found in [52] and [63] for example, and we mention that  $W_1(\rho_1, \rho_2) \leq W_{2+\varepsilon_*}(\rho_1, \rho_2)$ .

Now we construct 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  $\rho_{t-}$  is the law of  $X_{t-}$ . 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 1.1.** *There exists a measurable map  $\Phi : [0, 1] \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}^d$  such that for any  $\rho \in \mathcal{P}_1(\mathbb{R}^d)$ , any bounded and measurable function  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ , we have*

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

This result is stated in [32] 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 1.1**, 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  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \int_0^1 f(\eta_t^1(w)) dw &= \int_{\mathbb{R}^d} f(v_1) \rho_{\tau(t)-}^{\mathcal{P},M}(dv_1), \quad \int_0^1 f(\eta_t^2(w)) dw = \int_{\mathbb{R}^d} f(v_2) \rho_{t-}(dv_2), \\ \int_0^1 |\eta_t^1(w) - \eta_t^2(w)|^{2+\varepsilon_*} dw &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |v_1 - v_2|^{2+\varepsilon_*} \Pi_t^{\mathcal{P},M}(dv_1, dv_2) = (W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t-}))^{2+\varepsilon_*}. \end{aligned} \quad (32)$$

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

$$x_t = X_0 + \int_0^t b(r, x_r, \rho_r) dr + \int_0^t \int_{[0,1] \times \mathbb{R}^d} c(r, \eta_r^2(w), z, x_{r-}, \rho_{r-}) \mathcal{N}(dw, dz, dr), \quad (33)$$

$$x_t^M = X_0 + a_T^M \Delta + \int_0^t b(r, x_r^M, \rho_r) dr + \int_0^t \int_{[0,1] \times \mathbb{R}^d} c_M(r, \eta_r^2(w), z, x_{r-}^M, \rho_{r-}) \mathcal{N}(dw, dz, dr), \quad (34)$$

$$\begin{aligned}
x_t^{\mathcal{P},M} &= X_0 + a_T^M \Delta + \int_0^t b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M}) dr \\
&+ \int_0^t \int_{[0,1] \times \mathbb{R}^d} c_M(\tau(r), \eta_r^1(w), z, x_{\tau(r)-}^{\mathcal{P},M}, \rho_{\tau(r)-}^{\mathcal{P},M}) \mathcal{N}(dw, dz, dr).
\end{aligned} \tag{35}$$

One can check by Itô formula that  $x_t^{\mathcal{P},M}$  has the same law as  $X_t^{\mathcal{P},M}$  (solution of (23)),  $x_t^M$  has the same law as  $X_t^M$  (solution of (31)) and  $x_t$  has the same law as  $X_t$  (solution of (18)). Then

$$\begin{aligned}
(W_{2+\varepsilon_*}(\rho_{\tau(t)-}^{\mathcal{P},M}, \rho_{t-}))^{2+\varepsilon_*} &= (W_{2+\varepsilon_*}(\mathcal{L}(X_{\tau(t)-}^{\mathcal{P},M}), \mathcal{L}(X_{t-})))^{2+\varepsilon_*} = (W_{2+\varepsilon_*}(\mathcal{L}(x_{\tau(t)-}^{\mathcal{P},M}), \mathcal{L}(x_{t-})))^{2+\varepsilon_*} \\
&\leq \mathbb{E}|x_{\tau(t)-}^{\mathcal{P},M} - x_{t-}|^{2+\varepsilon_*}.
\end{aligned} \tag{36}$$

We notice that now  $x_t, x_t^M$  and  $x_t^{\mathcal{P},M}$  are defined on the same probability space and verify equations driven by the same Poisson point measure  $\mathcal{N}(dw, dz, dr)$ . Meanwhile, we are able to compare them by using a  $L^{2+\varepsilon_*}$  calculus. So all our computations below (in Part II) concern the equations (33), (34) and (35) (and not the equations (18), (31), (23)).

In order to obtain estimates of the total variation distance between these processes, we will use Malliavin integration by parts techniques (which are presented in Section 3 of Part II) together with some results from [6] which allows us to pass from estimates in Wassestein distance to estimates in total variation distance. In the following, we define the distances between random variables  $F, G : \Omega \rightarrow \mathbb{R}^d$ :

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

We recall in Part I that if we suppose that the Malliavin-Sobolev norms (see the definition in (11)) of  $F$  and  $G$  are bounded (i.e.  $\|F\|_{L,l,p} + \|G\|_{L,l,p} < \infty$ ) and the Malliavin covariance matrix (see the definition in (12)) of  $F$  and  $G$  are non-degenerated (i.e.  $\Sigma_p(F) + \Sigma_p(G) < \infty$ ), then we are able to control the total variation distance between  $F$  and  $G$  by  $d_r$  distance-so (13) holds. However sometimes the condition that the Malliavin covariance matrices of both  $F$  and  $G$  are non-degenerated is too strong and we are not able to prove it. Actually this is our case in Part II. We are not able to prove that the Malliavin covariance matrix of  $x_t^{\mathcal{P},M}$  is non-degenerated since the tangent flow of the Euler scheme is not invertible (the inverse tangent flow plays a crucial role in our proof). So we need the following variant of (13). We still assume that the Malliavin-Sobolev norms of  $F$  and  $G$  are bounded. If we only assume the non-degeneracy condition on  $F$  but no non-degeneracy condition for  $G$ , then we have the following (see Proposition 3.6.1 of Part II). We take some  $r \in \mathbb{N}$ . For any  $\varepsilon > 0$ , there exists a constant  $C > 0$  such that

$$d_{TV}(F, G) \leq C(d_r(F, G) + \|DF - DG\|_{L^2(\Omega; \mathcal{H})}^2)^{1-\varepsilon}. \tag{37}$$

We remark that in this case, if we want to control the total variation distance between  $F$  and  $G$ , then we not only need to control the  $d_r$  distance between them but also control the distance between the Malliavin derivatives of  $F$  and  $G$ . We then apply (37) for  $r = 1$  to  $F = x_t$  and  $G = x_t^{\mathcal{P},M}$  (solutions of the equations (33) and (35)). Fortunately, we are able to estimate  $\|Dx_t - Dx_t^{\mathcal{P},M}\|_{L^2(\Omega; \mathcal{H})}$  since  $x_t$  and  $x_t^{\mathcal{P},M}$  are defined by the same Poisson point measure. Meanwhile, we notice that  $d_1(x_t, x_t^{\mathcal{P},M}) \leq W_1(x_t, x_t^{\mathcal{P},M}) \leq W_{2+\varepsilon_*}(x_t, x_t^{\mathcal{P},M})$  and that we have (36), so we can estimate  $d_1(x_t, x_t^{\mathcal{P},M})$  by a  $L^{2+\varepsilon_*}$  calculus. We recall that  $x_t$  has the same law of  $X_t$  (solution of (18)) and  $x_t^{\mathcal{P},M}$  has the same law as  $X_t^{\mathcal{P},M}$  (solution of (23)). Consequently, we are able to prove that  $X_t^{\mathcal{P},M} \rightarrow X_t$  in total variation distance and we obtain (24). In a similar way, we can also prove that the Euler scheme  $X_t^{\mathcal{P}} \rightarrow X_t$  in total variation distance and obtain (22). We stress that these are non asymptotic bounds.

Finally, we explain the main strategy in order to obtain (27), (28), (29) and (30), which provide some algorithms based on the particle system  $X_t^{\mathcal{P},M,i}$ ,  $i = 1, \dots, N$  to compute the density function  $p_t(x)$  of the law of  $X_t$ . We will apply the estimate (4.6) in Theorem 4.1 of [25] which gives a basic estimation

concerning the particle system  $X_t^{\mathcal{P},M,i}$ ,  $i = 1, \dots, N$ . Besides, we need the regularization lemma given below. We recall the regularization kernel  $\varphi_\delta$  given in (26) 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 (see **Lemma 3.5** and **Corollary 3.5.1** of Part II for precise statement).

**Lemma 1.2. (A)** *Suppose that the Malliavin-Sobolev norms of  $F$  are bounded and the Malliavin covariance matrix of  $F$  is non-degenerated. Then the law of  $F$  has a smooth density  $p_F(x)$  and there exists a constant  $C$  such that for every  $x$ , we have*

$$|p_F(x) - \mathbb{E}(\varphi_\delta(F - x))| \leq C \times \delta^2. \quad (38)$$

Using Romberg method, we have

$$\left| p_F(x) + \mathbb{E}(\varphi_\delta(F - x)) - 2\mathbb{E}(\varphi_{\delta/\sqrt{2}}(F - x)) \right| \leq C \times \delta^4. \quad (39)$$

**(B)** *We suppose that the Malliavin-Sobolev norms of  $F$  are bounded. We take some  $\rho \geq 1$  and suppose that there is another random variable  $G$  such that  $\mathbb{E}|DG|_{\mathcal{H}}^p$ ,  $\forall p \geq 1$  are bounded and  $\Sigma_\rho(G) < \infty$ . For any  $\varepsilon_0 > 0$ , we denote  $q = 2/(1 + \varepsilon_0)$ . Then there exists a constant  $C$  such that for any  $\eta > 0$  and  $\delta > 0$ , for any function  $f \in C_b^\infty(\mathbb{R}^d)$ , we have*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f_\delta(F))| \leq C \|f\|_\infty \times \left( \frac{\delta^2}{\eta^4} + (\eta^{-1} \|DF - DG\|_{L^2(\Omega; \mathcal{H})})^q + \eta^\rho \right). \quad (40)$$

Using moreover Romberg method, we have

$$\left| \mathbb{E}(f(F)) + \mathbb{E}(f_\delta(F)) - 2\mathbb{E}(f_{\delta/\sqrt{2}}(F)) \right| \leq C \|f\|_\infty \times \left( \frac{\delta^4}{\eta^8} + (\eta^{-1} \|DF - DG\|_{L^2(\Omega; \mathcal{H})})^q + \eta^\rho \right). \quad (41)$$

We remark that in **(A)**, we assume the non-degeneracy condition for  $F$ , then we can approximate the density of  $F$  by  $\mathbb{E}(\varphi_\delta(F - x))$  with error  $\delta^2$  (see (38)). Moreover, applying Romberg method, we have (39) which gives a faster speed of convergence than (38). While in **(B)**, we do not assume the non-degeneracy condition for  $F$ , but we need to assume that we have another random variable  $G$  such that  $\Sigma_\rho(G) < \infty$ . Then we obtain (40) and (41). We have to mention that (40) and (41) are slightly different from the regularization lemma in [6]. The kernel considered in [6] is the super kernel, but we are not able to simulate the super kernel. So here we consider the Gaussian kernel  $\varphi_\delta$  which allows us to do the simulation. In conclusion, we apply (38) and (39) for  $F = x_t$  so as to obtain (27) and (28) respectively, and apply (40) and (41) for  $F = x_t^{\mathcal{P},M}$  and  $G = x_t$  so as to obtain (29) and (30) respectively. We refer to Section 3.3 of Part II for details.

### 1.3 Part III

There has been a long history of research on invariant probability measures. One can see for example [73], [75], [77] for the existence of an invariant probability measure for a general Markov process. These are some classical results. We refer to [80], [81] for some basic computation of the invariant probability measure for a Lévy process.

In this paper, we concern the approximation of the invariant probability measure for a Markov process. Following the idea from Pagès and Panloup [79], we construct an Euler scheme with decreasing time steps which is called the unadjusted Langevin algorithm in the Monte Carlo literature. This has been studied in depth in [82], [72], [71]. One can see also [76] for more discussions on the unadjusted Langevin algorithm and [51], [78] for the Monte Carlo method.

As far as we know, here are the newest results concerning the approximation of the invariant probability measure recently. In [79], the authors use unadjusted Langevin algorithm to approximate the invariant

probability measure of a diffusion process driven by a Brownian motion and study the Wasserstein and total variation distance between them. In [70], the authors approximate the invariant probability measure of a Lévy process but only study the Wasserstein distance. Some other related results can also be found in [64] for example. In our paper, we aim at giving an abstract framework of the approximation for the invariant probability measure, which includes the case of diffusion process in [79] and Lévy process in [70]. We also apply this framework to jump processes to illustrate our results and study the total variation distance, which gives stronger results than [70]. We describe our ideas more precisely in the following.

We give in Section 2 of Part III an abstract framework of the approximation for the invariant probability measure. 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  $l$ . We consider a semigroup  $P_t, t \geq 0$  on the space  $\mathcal{M}_b(\mathbb{R}^d)$  of the bounded measurable functions on  $\mathbb{R}^d$  and assume that there exists at least one invariant probability measure  $\nu$  for the semigroup  $P_t, t \geq 0$ . We assume moreover the "exponential Lipschitz property": there exists two constants  $C_0 \geq 1$  and  $\rho > 0$  such that for every  $t > 0$  and every  $\varphi \in C_b^1(\mathbb{R}^d)$

$$(L_0) \quad \|\nabla P_t \varphi\|_\infty \leq C_0 \|\nabla \varphi\|_\infty e^{-\rho t}.$$

We notice that (see **Proposition 2.0.1** of Part III) the existence of an invariant probability measure  $\nu$  together with the "exponential Lipschitz property" implies that the invariant probability measure  $\nu$  is unique.

In order to approximate the invariant measure  $\nu$ , we introduce an Euler scheme with decreasing time steps (unadjusted Langevin algorithm). For every  $\gamma > 0$  we give an operator  $\bar{P}_\gamma : C_b^\infty \rightarrow C_b^\infty$  such that  $\|\bar{P}_\gamma \varphi\|_\infty \leq \|\varphi\|_\infty$  and which approximates our semigroup in the following sense: for every  $\gamma > 0$

$$A(k_0, \alpha) \quad \|(P_\gamma - \bar{P}_\gamma)\varphi\|_\infty \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma^{1+\alpha}. \quad (42)$$

Here  $\alpha > 0$  is a given number,  $k_0 \in \mathbb{N}$  and  $\|\psi\|_{k_0, \infty} = \sum_{|\alpha| \leq k_0} \|\partial^\alpha \psi\|_\infty$ . We consider a decreasing sequence of time steps  $\gamma_n \downarrow 0$  and define the time grid  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . We assume that

$$(\Gamma) \quad \sum_{i=1}^{\infty} \gamma_i = \lim_{n \rightarrow \infty} \Gamma_n = \infty.$$

We also introduce

$$\bar{\omega} = \bar{\omega}((\gamma_n)_{n \in \mathbb{N}}) = \lim_{n \rightarrow \infty} \frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2} < \infty.$$

The typical example is  $\gamma_n = \frac{1}{n}$  and then  $\bar{\omega} = 1$ . In the following we denote  $\{\Gamma\} = \{\Gamma_n, n \in \mathbb{N}\}$ . And, for  $\Gamma_i \leq t < \Gamma_{i+1}$  we denote  $N(t) = i$  and  $\tau(t) = \Gamma_i$ . Then, for  $s \in \{\Gamma\}$  and  $t \in \{\Gamma\}$  we define the Euler scheme

$$\bar{P}_{s,t} = \prod_{i=N(s)}^{N(t)-1} \bar{P}_{\gamma_i}, \quad (43)$$

the product being understood in the sense of composition. This means that we travel from  $\tau(s)$  to  $\tau(t)$  by using the Euler scheme associated to the one step Euler scheme  $\bar{P}_\gamma$ .

So now we use the Euler scheme with decreasing time steps  $\bar{P}_{0, \Gamma_n}$  (given in (43)) to approximate the invariant probability measure  $\nu$ . Our aim is to estimate the total variation distance between them. To do so, we need some regularization properties. First we give the regularization hypothesis concerning the semigroup  $P_t$ :

$$R_P(k) \quad \sup_{1 \leq t \leq 2} \|P_t \varphi\|_{k, \infty} \leq C_k \|\varphi\|_\infty, \quad \text{and} \\ R'_P(k) \quad \sup_{1 \leq t \leq 2} \|\nabla P_t \varphi\|_{k-1, \infty} \leq C'_k \|\nabla \varphi\|_\infty.$$

We also introduce the following variant of the Lipschitz property:

$$\bar{L}_k \quad \|\nabla P_t \varphi\|_{k, \infty} \leq C_k \|\nabla \varphi\|_{k, \infty}, \quad 1 \geq t > 0.$$

We give now the regularization properties for the Euler scheme  $\bar{P}_{s,t}$ . To begin, we introduce some notations. We recall that a super kernel  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a function which belongs to the Schwartz space and such that for every multi-indexes  $\beta_1$  and  $\beta_2$ , one has

$$\int_{\mathbb{R}^d} \phi(x) dx = 1, \quad \int_{\mathbb{R}^d} y^{\beta_1} \phi(y) dy = 0 \quad \text{for } |\beta_1| \geq 1, \quad \int_{\mathbb{R}^d} |y|^m |\partial^{\beta_2} \phi(y)| dy < \infty \quad \text{for } m \in \mathbb{N}. \quad (44)$$

We fix a super kernel  $\phi$ . For  $\delta \in (0, 1]$ , we denote

$$\phi_\delta(y) = \frac{1}{\delta^d} \phi\left(\frac{y}{\delta}\right) \quad (45)$$

and  $f_\delta$  the regularization by convolution with a super kernel:

$$f_\delta(x) = f * \phi_\delta(x) = \int_{\mathbb{R}^d} f(y) \phi_\delta(x - y) dy, \quad (46)$$

with  $*$  denoting convolution. For  $\delta > 0, \eta > 0$ , and  $q, \kappa, p \in \mathbb{N}$  we denote

$$A_{q,\kappa,p}^{\delta,\eta}(h) = \frac{\delta^q}{\eta^{2q}} + \eta^{-p} h^p + \eta^\kappa, \quad h > 0.$$

Let  $\beta > 0$  and  $p \geq 1$  be fixed and we assume the following regularization property for the Euler scheme  $\bar{P}_{s,t}$ : we assume that for every  $q, \kappa \in \mathbb{N}$  there exists a constant  $C = C_{q,\kappa,p}$  such that for every  $\delta > 0, \eta > 0$ , every  $1 < t < r < t + 2$  and every bounded measurable function  $f$

$$\begin{aligned} R_{\bar{P}}(p, \beta) & \quad \left\| \bar{P}_{t-1,t} P_{t,r} f - \bar{P}_{t-1,t} P_{t,r} f_\delta \right\|_\infty + \left\| \bar{P}_{t-1,t} \bar{P}_{t,r} f - \bar{P}_{t-1,t} \bar{P}_{t,r} f_\delta \right\|_\infty \\ & \leq C_{q,\kappa,p} \times A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) \|f\|_\infty. \end{aligned}$$

So now we can give our main result (see **Proposition 2.1.1** of Part III). We assume that an invariant probability measure  $\nu$  exists for the semigroup  $P_t, t \geq 0$ . We construct an Euler scheme with decreasing time steps  $\bar{P}_{s,t}$  by (43). Suppose that  $(L_0)$  holds for some  $\rho$ ,  $A(k_0, \alpha)$  holds for some  $k_0, \alpha$  with  $\rho > \alpha\bar{\omega}$ ,  $R_P(k)$ ,  $R'_P(k)$  and  $\bar{L}_k$  hold for every  $k$ , and  $R_{\bar{P}}(p, \beta)$  holds true for some  $p, \beta$ . Then the invariant probability measure  $\nu$  is unique and for any  $\varepsilon > 0$ , for every  $x \in \mathbb{R}^d$  and  $n$  large enough,

$$d_{TV}(\bar{P}_{0,\Gamma_n}(x, \cdot), \nu) \leq C_\varepsilon (\gamma_n^{((p\beta) \wedge \alpha) - \varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\rho\Gamma_n}). \quad (47)$$

We remark that we get the same speed of convergence as in [79] and [70], but in a more general framework.

We notice that we need some regularization properties (see  $R_P(k)$ ,  $R'_P(k)$  and  $R_{\bar{P}}(p, \beta)$ ). In order to obtain these properties, we use some integration by parts techniques as in Part I and Part II. We give now a regularization lemma which plays a crucial role in our paper. We recall the super kernel given in (44), (45) and (46). Then we have the following regularization lemma (see **Lemma 3.5** of Part III for precise statement) which is originally from the paper [6].

**Lemma 1.3.** *We fix some  $\kappa \geq 1$ . We consider a random variable  $F$  and suppose that the Malliavin-Sobolev norms of  $F$  (see (11)) are bounded. We also consider an auxiliary random variable  $Q$  such that  $\Sigma_\kappa(Q) < \infty$ . Then there exists a constant  $C$  such that for any  $q \in \mathbb{N}$ ,  $p \geq 1$ ,  $\eta > 0$  and  $\delta > 0$ , for any function  $f \in C_b^\infty(\mathbb{R}^d)$ , we have*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f_\delta(F))| \leq C \|f\|_\infty \times \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p) + \eta^\kappa \right). \quad (48)$$

*Remark.* We remark that we do not assume the non-degeneracy condition for  $F$ , but we need to assume that we have another random variable  $Q$  which is non-degenerated such that  $\det \sigma_Q$  is close to  $\det \sigma_F$ . The regularization lemma here is a variant of **Lemma 1.2 (B)**. In **Lemma 1.2 (B)**, we consider the Gaussian kernel (26) since we need to simulate the kernel in (29) and (30). However in (47) we do not need to simulate the kernel itself, so we consider the super kernel (44) in **Lemma 1.3** which is not able to be simulated but can give faster speed of convergence by optimization on  $q$ .

In Section 4 of Part III, we apply the result (47) in the abstract framework for jump processes. So we consider the  $d$ -dimensional stochastic differential equation with jumps as follows:

$$X_t = x + \int_0^t b(X_r)dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{r-})N(dz, dr), \quad (49)$$

where  $N(dz, dr)$  is a Poisson point measure on the state space  $\mathbb{R}^d$  with intensity measure  $\widehat{N}(dz, dr) = \mu(dz)dr$ ,  $x$  is the initial value,  $\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ , and  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ .

We need to give sufficient conditions to ensure the existence of an invariant probability measure for the jump equation (49). We recall by [73] the classical results of the existence of an invariant probability measure for a general Markov process. Recently, [74] gives some specific criterias for the existence of an invariant probability measure of a jump process and also discuss some ergodicity properties. Here we suppose that (see **Hypothesis 2.5** of Part III)

$$\begin{aligned} i) \quad & \langle x - y, b(x) - b(y) \rangle \leq -\bar{b} |x - y|^2 \\ ii) \quad & |c(z, x) - c(z, y)| \leq \bar{c}(z) |x - y| \end{aligned}$$

and

$$iii) \quad 2\bar{b} - \int_{\mathbb{R}^d} (2\bar{c}(z) + \bar{c}^2(z))\mu(dz) := \theta > 0.$$

Our conditions are based on [73] and are essentially the same as the conditions in [74]. Indeed, the conditions above implies that for some  $\bar{\beta}, \bar{\alpha} > 0$  and a Lyapunov function  $V(x) = |x|^2$ , we have

$$(\text{Lyapunov mean reverting condition}) \quad LV \leq \bar{\beta} - \bar{\alpha}V,$$

with  $L$  denoting the infinitesimal operator of (49). This guarantees the existence of an invariant probability measure  $\nu$ .

Moreover, in order to apply the Malliavin integration by parts framework and obtain regularization properties, we assume (see **Hypothesis 2.4 b**) of Part III) that the measure  $\mu$  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. We also need some regularity and ellipticity conditions on the coefficients (see **Hypothesis 2.1~2.3** of Part III for details). We mention that for every multi-indices  $\beta_1, \beta_2$ , we assume that there exists a non-negative function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that

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

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  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that for every  $\zeta \in \mathbb{R}^d$ ,

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

Now we construct the Euler scheme. We take a partition with decreasing time steps

$$\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{n-1} < \Gamma_n < \dots\}$$

with the time steps  $\gamma_n = \Gamma_n - \Gamma_{n-1}$ ,  $n \in \mathbb{N}$  verifying some suitable conditions (see Section 4.3 of Part III for details). For  $\Gamma_n \leq t < \Gamma_{n+1}$  we denote  $\tau(t) = \Gamma_n$ . We consider the Euler scheme:

$$X_t^{\mathcal{P}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}})dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{\tau(r)-}^{\mathcal{P}})N(dz, dr).$$

Since  $\mu(\mathbb{R}^d) = \infty$  (which is a consequence of **Hypothesis 2.4 a**) of Part III), we have infinitely many jumps. So we construct the truncated Euler scheme in order to have finite numbers of jumps for the sake of simulation and Malliavin calculus. For  $m \in \mathbb{N}$ , we denote  $B_m = \{z \in \mathbb{R}^d : |z| \leq m\}$  and denote

$$\varepsilon_m := \int_{\{|z| > m\}} |\bar{c}(z)|^2 \mu(dz) + \left| \int_{\{|z| > m\}} \bar{c}(z) \mu(dz) \right|^2.$$



For every  $\gamma > 0$ , we define the truncation function  $M(\gamma) \in \mathbb{N}$  such that

$$\varepsilon_{M(\gamma)} \leq \gamma^2.$$

For  $\Gamma_n < t \leq \Gamma_{n+1}$ , we denote  $M_{\mathcal{P}}(t) = M(\gamma_{n+1})$ . Now we cancel the "big jumps" (the jumps of size  $|z| > M_{\mathcal{P}}(t)$ ):

$$X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \int_0^t \int_{B_{M_{\mathcal{P}}}(r)} c(z, X_{\tau(r)-}^{\mathcal{P}, M_{\mathcal{P}}}) N(dz, dr). \quad (50)$$

The advantage of considering  $X_t^{\mathcal{P}, M_{\mathcal{P}}}$  is that we may represent it by means of compound Poisson processes. For  $k \in \mathbb{N}$ , we denote  $I_1 = B_1$ ,  $I_k = B_k \setminus B_{k-1}$  for  $k \geq 2$  and take  $(J_t^k)_{t \geq 0}$  a Poisson process of intensity  $\mu(I_k)$ . We denote by  $(T_i^k)_{i \in \mathbb{N}}$  the jump times of  $(J_t^k)_{t \geq 0}$  and we consider a sequences of independent random variables  $Z_i^k \sim \mathbb{1}_{I_k}(z) \frac{\mu(dz)}{\mu(I_k)}$ ,  $k, i \in \mathbb{N}$ . Moreover,  $((J_t^k)_{t \geq 0}, (Z_i^k)_{k, i \in \mathbb{N}})_{k \in \mathbb{N}}$  are taken to be independent. Then we represent (50) by compound Poisson process on each set  $I_k$ . We write

$$\begin{aligned} X_t^{\mathcal{P}, M_{\mathcal{P}}} &= x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{k=1}^{\infty} \sum_{i=1}^{J_t^k} \mathbb{1}_{B_{M_{\mathcal{P}}}(T_i^k)}(Z_i^k) c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}) \\ &= x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{k=1}^{\infty} \sum_{i=1}^{J_t^k} \sum_{n=0}^{N(t)} \mathbb{1}_{B_{M(\gamma_{n+1})}}(Z_i^k) c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}). \end{aligned}$$

Notice that  $Z_i^k \in B_{M(\gamma_{n+1})}$  is equivalent to  $k \leq M(\gamma_{n+1})$ . It follows that

$$X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}). \quad (51)$$

Now (51) can be constructed in an explicit way. We remark that rather than being a constant, the truncation rule  $M_{\mathcal{P}}(t)$  may change on different time intervals. We have to take the truncation like this in order to verify the condition (42).

Then we apply (47) in the abstract framework for  $X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}$  and obtain the following error estimate (see **Theorem 4.1** of Part III): An invariant probability measure  $\nu$  of the jump equation (49) exists and is unique, and for any  $\varepsilon > 0$ , there exists a constant  $C_\varepsilon$  such that for every  $x \in \mathbb{R}^d$  and  $n$  large enough, we have

$$d_{TV}(\mathcal{L}(X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}), \nu) \leq C_\varepsilon (\gamma_n^{1-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\frac{\varepsilon}{2} \Gamma_n}),$$

with  $\mathcal{L}(X)$  denoting the law of a random variable  $X$ . We notice that we obtain the same speed of convergence as in [79] but [79] concerns the diffusion process while here we consider the jump process. Comparing with the results in [70], we also obtain the same speed of convergence but [70] only deals with the Wasserstein distance while in our paper, we deal with the total variation distance, which gives stronger results.

## References

- [1] A. Alfonsi, E. Cancès, G. Turnici, B. Di Ventura, and W. Huisinga : Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. In ESAIM Proceedings, volume 14, pages 1–13(2005).
- [2] A. Alfonsi, E. Cancès, G. Turnici, B. Di Ventura, and W. Huisinga: Exact simulation of hybrid stochastic and deterministic models for biochemical systems. Research Report RR-5435, INRIA (2004).

- [3] S. Asmussen and J. Rosinski: Approximations of small jumps of Lévy processes with a view towards simulation, *Journal of Applied Probability* **38** , 482-493 (2001).
- [4] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala: Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, **16**:1925–1961 (2006).
- [5] V. Bally, L. Caramellino, G. Poly: Convergence in distribution norms in the CLT for non identical distributed random variables. *Electronic Journal of Probability*, Institute of Mathematical Statistics(IMS), **23**, paper 45,51 p. (2018).
- [6] V. Bally, L. Caramellino, G. Poly: Regularization lemmas and convergence in total variation. *Electron. J. Probab.* **25** 1-20. (2020).
- [7] V. Bally, E. Clément: Integration by parts formula and applications to equations with jumps. *Probab. Th. Rel.Fields*, **151**, 613-657. (2011).
- [8] K. Bichteler, J. B. Gravereaux, J. Jacod: *Malliavin calculus for processes with jumps*. Gordon and Breach, (1987).
- [9] A. Carpentier, C. Duval, E. Mariucci: Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps. arXiv:1810.02998 [math.ST] (2019).
- [10] A. Crudu, A. Debussche, A. Muller, O. Radulescu: Convergence of stochastic gene networks to hybrid piecewise deterministic processes, *Ann. Appl. Probab.* **10**, Nr 5, p 1822-1859.
- [11] C. Duval and E. Mariucci: Spectral free estimates of Lévy densities in high level frequency, to appear in Bernoulli. arXiv 1702. 08787 [Math PR].
- [12] N. Fournier: Simulation and approximation of Lévy driven stochastic differential equations. *ESAIM Proba and Stat*, Vol **15**, 249-269, (2011).
- [13] Y. Ishikawa: *Stochastic Calculus of Variations for Jump Processes*, Berlin, Boston: De Gruyter. (2013).
- [14] A. M. Kulik: Malliavin calculus for Lévy processes with arbitrary Lévy measures. *Theor. Probability ad Math. Statist.* No.**72**, 75-92. (2006).
- [15] 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).
- [16] T. Kurtz: Limits theorems for sequences of jump Markov processes approximating ordinary differential processes. *J.Appl.Prob.*, **8**:344–356 (1971).
- [17] T. Kurtz: Strong approximation theorems for density dependent Markov chains. *Stoch. Proc. Appl.*, **6**:223–240 (1978).
- [18] D. Nualart: *The Malliavin calculus and related topics*. Springer-Verlag, (2006).
- [19] P. Protter and D. Talay: The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.* Vol **25**, No.1, pg 393-423 (1997).
- [20] Y. Song and X. Zhang: Regularity of density for SDEs driven by degenerate Lévy noises. arXiv:1401.4624 (2014).
- [21] J. B. Walsh: *An introduction to stochastic partial differential equations*. Ecole d'été de probabilités de Saint Flour, 1984, p 265-439, ed. L. Hennequin, Springer Verlag.
- [22] X. Zhang: Densities for SDEs driven by degenerate  $\alpha$ -stable processes. *Ann. Probab.* Vol **42**, No.5, 1885-1910 (2014).
- [23] S. Albeverio, B. Rüdiger, and P. Sundar: The Enskog process. *J. Stat. Phys.* **167**(1):90–122 (2017).

- [24] R. Alexandre: A review of Boltzmann equation with singular kernels. *Kinet. Relat. Models*, **2**(4), 551-646 (2009).
- [25] A. Alfonsi, V. Bally : Construction of Boltzmann and McKean Vlasov type flows (the sewing lemma approach). arXiv:2105.12677 [math.PR](2021).
- [26] 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).
- [27] 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).
- [28] D. Applebaum: *Lévy Processes and Stochastic Calculus* (2nd ed., Cambridge Studies in Advanced Mathematics). Cambridge: Cambridge University Press. (2009) doi:10.1017/CBO9780511809781.
- [29] V. Bally, N. Fournier: Regularization properties of the 2D homogeneous Boltzmann equation without cutoff. *Probab. Theory Related Fields*, **151**(3-4), 659-704 (2011).
- [30] 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)
- [31] 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).
- [32] R. Carmona, F. Delarue: *Probability theory of mean field games with applications*. Springer Probability Theory and Stochastic Modelling Vol.**83** (2018).
- [33] C. Cercignani: *The Boltzmann equation and its applications*. Springer-Verlag Applied Mathematical Sciences, Vol.**67** (1988).
- [34] R. Cont, P. Tankov: *Finacial modelling with jump processes*. Chapman & Hall/CRC (2004).
- [35] 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).
- [36] 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).
- [37] 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).
- [38] 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).
- [39] M. Friesen, B. Rüdiger, and P. Sundar: On uniqueness and stability for the Enskog equation (2020).
- [40] E. Gobet, G. Pagès, H. Pham and J. Printems: Discretization and simulation for a class of SPDEs with applications to Zakai and McKean-Vlasov equations. (2005). (hal-00003917v1)
- [41] C. Graham: McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications* **40** 69-82 (1992).
- [42] 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).
- [43] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes*. 2nd ed. Amsterdam, Netherlands, North Holland, (1989).

- [44] 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).
- [45] 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).
- [46] 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).
- [47] A. Kohatsu-Higa, P. Protter: The Euler scheme for SDE'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).
- [48] A. Kohatsu-Higa and P. Tankov: Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*, Vol. **120**, 2258-2285 (2010).
- [49] 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).
- [50] H. Kunita: *Stochastic flows and jump-diffusions*. Springer,(2019).
- [51] B. Lapeyre, É. Pardoux and R. Sentis: *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*. Springer-Verlag, Berlin, (1998).
- [52] E. Mariucci, M. Reiß: Wasserstein and total variation distance between marginals of Lévy processes. *Electronic Journal of Statistics* **12**, 2482-2514 (2018).
- [53] 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).
- [54] M. A. Mezerdi: On the convergence of carathéodory numerical scheme for Mckean-Vlasov equations. *Stochastic Analysis and Applications*, **39**(5): 804-818 (2021).
- [55] R. Naito, T. Yamada: A higher order weak approximation of McKean–Vlasov type SDEs. *BIT Computer Science and Numerical Mathematics United States* (2022).
- [56] K. Sato. Lévy processes and infinitely divisible distributions. *Cambridge University press*, Cambridge (1999).
- [57] 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).
- [58] A.-S. Sznitman: Équations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebiete*, **66**(4):559–592 (1984).
- [59] 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).
- [60] H. Tanaka: Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete*. **46**(1), 67-105 (1978).
- [61] 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).
- [62] 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).

- [63] C. Villani: *Optimal Transport* Springer-Verlag, (2009).
- [64] A. Alfonsi, J. Corbetta, B. Jourdain: Evolution of the Wasserstein distance between the marginals of two Markov processes. *Bernoulli*, **24** (4A), pp.2461-2498. (hal-01390887) (2018).
- [65] D. Bakry, I. Gentil, M. Ledoux: Analysis and geometry of Markov diffusion operators. *Grundlehren der Mathematischen Wissenschaften* **348**, Springer, MR-3155209 (2014).
- [66] V. Bally and L. Caramellino: On the distances between probability density functions. *Electron. J. Probab.* **19**, no. 110, 1-33. MR-3296526 (2014).
- [67] V. Bally and L. Caramellino: Asymptotic development for the CLT in total variation distance. *Bernoulli* **22**, 2442–2485. MR-3498034 (2016).
- [68] V. Bally, L. Caramellino, G. Poly: Non universality for the variance of the number of real roots of random trigonometric polynomials. *Probab. Th. Rel. Fields.* **174**, 887-927. MR-3980307 (2019).
- [69] N. Bouleau, L. Denis: *Dirichlet forms methods for Poisson point measures and Lévy processes. With emphasis on the creation-annihilation techniques.* Probability Theory and Stochastic Modelling **76**, Springer. MR-3444890 (2015).
- [70] P. Chen, C. Deng, René L. Schilling, et. al.: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. arXiv:2205.01342v2 [math.PR] (2023).
- [71] A. Durmus, É. Moulines: High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm. arXiv:1605.01559v4 (2018).
- [72] A. Durmus, É. Moulines: Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *The Annals of Applied Probability* Vol. **27**, No. 3, 1551–1587 (2017).
- [73] Stewart N. Ethier, Thomas G. Kurtz: *Markov Processes: characterization and convergence.* (1984).
- [74] E. Löcherbach, V. Rabinet: Ergodicity for multidimensional jump diffusions with position dependent jump rate. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* <hal-01144260> (2015).
- [75] Sean P. Meyn and R. L. Tweedie: Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Advances in Applied Probability* Vol. **25**, No. 3, pp. 518-548 (1993).
- [76] W. Mou, N. Flammarion, M. J. Wainwright and P. L. Bartlett: Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity. arXiv:1907.11331, (2019).
- [77] G. Pagès: Sur quelques algorithmes récurrents pour les probabilités numériques. *Probability and Statistics* Vol. **5**, 141-170 (2001).
- [78] G. Pagès: *Numerical Probability: An introduction with applications to finance.* Universitext. Springer, Cham (2018).
- [79] G. Pagès, F. Panloup: Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. (hal-03082311v3) (2022).
- [80] F. Panloup: Computation of the invariant measure for a Lévy driven SDE: Rate of convergence. *Stochastic Processes and their Applications*, Elsevier, **118** (8), pp.1351-1384. (2008).
- [81] F. Panloup: Recursive computation of the invariant measure of a stochastic differential equation driven by a Lévy process. *Annals of Applied Probability*, Institute of Mathematical Statistics (IMS), **18** (2), pp.379-426. (2008).
- [82] Gareth O. Roberts and Richard L. Tweedie: Exponential convergence of Langevin distributions and their discrete approximations *Bernoulli* **2**(4), 341-363. (1996).

## Part I

.

Total variation distance between a jump-equation and its  
Gaussian approximation

# Total variation distance between a jump-equation and its Gaussian approximation

Vlad Bally\* and Yifeng Qin\*

\*Université Gustave Eiffel, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. Email address: bally@univ-mlv.fr

## Abstract

We deal with stochastic differential equations with jumps. In order to obtain an accurate approximation scheme, it is usual to replace the "small jumps" by a Brownian motion. In this paper, we prove that for every fixed time  $t$ , the approximate random variable  $X_t^\varepsilon$  converges to the original random variable  $X_t$  in total variation distance and we estimate the error. We also give an estimate of the distance between the densities of the laws of the two random variables. These are done by using some integration by parts techniques in Malliavin calculus.

**Key words:** Stochastic differential equations with jumps, Malliavin calculus, Total variation distance, Gaussian approximation

## Contents

<b>1</b>	<b>Introduction</b>	<b>24</b>
<b>2</b>	<b>Main results</b>	<b>25</b>
2.1	The basic equation and the hypotheses . . . . .	25
2.2	Approximation . . . . .	27
2.3	The main theorem . . . . .	29
2.4	A typical example . . . . .	30
<b>3</b>	<b>Abstract integration by parts framework</b>	<b>31</b>
3.1	Main consequences: Convergence in total variation distance . . . . .	34
<b>4</b>	<b>Malliavin calculus and stochastic differential equations with jumps</b>	<b>35</b>
4.1	The splitting method . . . . .	36
4.2	Malliavin calculus for Poisson point measures and space-time Brownian motions . . . . .	38
4.3	Malliavin calculus applied to stochastic differential equations with jumps . . . . .	39
4.4	Proof of the main result (Theorem 2.2) . . . . .	39
<b>5</b>	<b>Appendix</b>	<b>40</b>
5.1	Proof of Lemma 4.1 . . . . .	40
5.1.1	Construction of the Euler scheme . . . . .	40
5.1.2	Preliminary estimates . . . . .	42
5.1.3	Estimations of $\ X_t^{n,M}\ _{L,l,p}$ . . . . .	46
5.2	Proof of Lemma 4.2 . . . . .	51
5.3	Some proofs concerning Section 4.2 . . . . .	55

# 1 Introduction

In this paper we consider the stochastic differential equation with jumps

$$X_t = x + \int_0^t \int_{(0,1]} c(s, z, X_{s-}) N_\mu(ds, dz)$$

where  $N_\mu$  is a Poisson random measure on  $(0, 1]$  with compensator  $\mu(dz)ds$  and  $c$  is a coefficient which verifies strong regularity hypotheses (see **Hypotheses 2.1-2.4** in Section 2.1). The typical example that we have in mind is  $\mu(dz) = \frac{dz}{z^{1+\rho}} 1_{\{z \in (0,1]\}}$ , with  $\rho \in [0, 1)$ , so this is a truncated stable process - however, throughout the paper, we keep the general framework in which  $\mu$  is a measure which has infinite mass around zero. Our aim is to replace the "small jumps" by a space-time Brownian motion:

$$\begin{aligned} X_t^\varepsilon &= x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, X_{s-}^\varepsilon) N_\mu(ds, dz) \\ &\quad + \int_0^t b_\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \int_{(0, \varepsilon]} c(s, z, X_s^\varepsilon) W_\mu(ds, dz), \end{aligned} \tag{1}$$

where  $W_\mu(ds, dz)$  is a space-time Brownian motion (in the sense of Walsh [36]) with covariance  $\mu(dz)ds$ ,  $x \in \mathbb{R}$ , and the coefficient  $b_\varepsilon$  is defined by

$$b_\varepsilon(s, x) = \int_{(0, \varepsilon]} c(s, z, x) \mu(dz).$$

The interest of such approximations appears in various frameworks.

Our main motivation comes from numerical computations. If  $\mu(E) < \infty$  then there are a finite number of jumps in any compact interval of time, so  $X_t$  may be represented by means of a compound Poisson process which may be explicitly simulated. But if  $\mu(E) = \infty$  this is not possible anymore (except in very particular situations - see Talay and Protter [34] for example), and the "small jumps" should be truncated to revert to the case of a finite measure. This procedure is rather rough and gives large errors. In order to improve the approximation scheme, one may replace the "small jumps", namely those smaller than  $\varepsilon$ , by a stochastic integral with respect to  $W_\mu(ds, dz)$ . Note that the Poisson measure  $dN_\mu$  is not compensated, which is why the drift corresponding to  $b_\varepsilon$  appears. This idea goes back to Asmussen and Rosinski [3]. In the case of *SDE's* driven by a Lévy process, Fournier [16] gives a precise estimate of the error and compares the approximation obtained just by truncating the small jumps to the one obtained by adding a Gaussian noise as in (1). An enlightening discussion on the complexity of the two methods is also given. However, in that paper, the strong error is considered, while in our paper we discuss the weak error.

A second motivation comes from modelization problems in chemistry and biology: we are concerned by reactions which are naturally modelled by means of jump processes containing two regimes: one is very rapid but the jumps are small, and another is much slower and the jumps are larger – see for example [1], [2], [4], [13], [28], [29]. In this case the regime corresponding to the rapid scale may be modelled by a stochastic integral with respect to a Gaussian process and the slow regime by a compound Poisson process. It may also be reasonable to consider an intermediary regime and this would be modelled by a drift term.

A third motivation is given by a class of statistical problems (see [11], [15] and references therein), where a stochastic process is observed at various times and it should be decided whether its increments are due to small jumps or to a Gaussian component. In this framework it is important to estimate the error in total variation sense. The authors explain that, if the error in total variation between the laws of  $X_t$  and of  $X_t^\varepsilon$  goes to zero, then there is no way to construct a test which decides if the noise comes from small jumps or from the Brownian motion. So, asymptotically, the two models contain the same information.

Let us now discuss briefly our results and the relation to previously available estimates. If  $L_t$  (respectively  $L_t^\varepsilon$ ) represents the infinitesimal operator of  $X_t$  (respectively of  $X_t^\varepsilon$ ) then a development in Taylor series of order two gives

$$\|(L_t - L_t^\varepsilon)f\|_\infty \leq C \|f\|_{3,\infty} \int_{(0, \varepsilon]} |\hat{c}(z)|^3 \mu(dz),$$



where  $\hat{c}(z) := \sup_{s \leq T} \sup_{x \in \mathbb{R}} |c(s, z, x)|$  and  $\|f\|_{3,\infty} := \sum_{0 \leq i \leq 3} \|f^{(i)}\|_\infty$ . Then a Trotter-Kato type argument yields

$$\sup_{s \leq T} \|(P_t - P_t^\varepsilon)f\|_\infty \leq C \|f\|_{3,\infty} \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz)$$

where  $P_t$  (respectively of  $P_t^\varepsilon$ ) represents the semigroup of  $X_t$  (respectively of  $X_t^\varepsilon$ ).

The drawback of the above estimate is that the bound on the error involves  $\|f\|_{3,\infty}$ , so it only applies to smooth test functions. The main contribution of our paper is to replace  $\|f\|_{3,\infty}$  by  $\|f\|_\infty$ , so as to prove convergence in total variation distance. This is done under non-degeneracy and regularity assumptions on the coefficient  $c$ . Moreover, under these hypotheses, we prove that  $\mathbb{P}(X_t(x) \in dy) = p_t(x, y)dy$  and  $\mathbb{P}(X_t^\varepsilon(x) \in dy) = p_t^\varepsilon(x, y)dy$  with smooth densities  $y \mapsto p_t(x, y)$  and  $y \mapsto p_t^\varepsilon(x, y)$ . And, for every  $k$  and every  $\delta > 0$ , we obtain

$$\|\partial_y^k p_t - \partial_y^k p_t^\varepsilon\|_\infty \leq C_{k,\delta} \left( \int_{(0,\varepsilon]} |\hat{c}(z)|^3 \mu(dz) \right)^{1-\delta}.$$

This proves that  $p_t^\varepsilon$  converges to  $p_t$  in distribution norms as  $\varepsilon \rightarrow 0$ .

Our approach uses a strategy based on integration by parts (an abstract Malliavin calculus) developed in [6].

The paper is organized as follows. In Section 2, we give the main results and in Section 3, we recall the integration by parts technique introduced in [6] and used here. In Section 4, we use these results in the framework of stochastic equations with jumps and we prove the main result (**Theorem 2.2**). The Appendix contains technical estimates concerning Sobolev norms in Malliavin sense.

## 2 Main results

### 2.1 The basic equation and the hypotheses

A time horizon  $T > 0$  will be fixed throughout the paper. As already mentioned, we deal with the one-dimensional jump equation

$$X_t = x + \int_0^t \int_{(0,1]} c(s, z, X_{s-}) N_\mu(ds, dz), \quad (2)$$

where  $N_\mu$  is a Poisson point measure with intensity  $\widehat{N}_\mu(ds, dz) = \mu(dz)ds$ , and  $\mu$  is a positive  $\sigma$ -finite measure on  $(0, 1]$ ,  $t \in [0, T]$ .

For technical reasons which will be discussed in Section 4, we introduce the following change of variables. Let  $\theta : (0, 1] \rightarrow [1, \infty)$  be the function defined by  $\theta(z) = \frac{1}{z}$ , and let  $\nu(dz) := \mu \circ \theta^{-1}(dz)$ . Then  $\nu$  is a positive  $\sigma$ -finite measure on  $[1, \infty)$ . Consider a Poisson point measure  $N_\nu(ds, dz)$  with intensity  $\widehat{N}_\nu(ds, dz) = \nu(dz)ds$ . One may then check that for every  $t \in [0, T]$ ,  $X_t$  has the same law as  $\widehat{X}_t$ , with  $(\widehat{X}_t)_{t \in [0, T]}$  the solution of

$$\widehat{X}_t = x + \int_0^t \int_{[1,\infty)} \tilde{c}(s, z, \widehat{X}_{s-}) N_\nu(ds, dz), \quad (3)$$

where  $\tilde{c}(s, z, x) := c(s, \frac{1}{z}, x)$ .

Since this paper deals with the laws of the solution to (2), it is equivalent to consider the equation (3). We formulate our hypotheses in terms of  $\tilde{c}$  and  $\nu$  (instead of  $c$  and  $\mu$ ).

**Hypothesis 2.1 (Regularity with parameter  $q^*$ )** The map  $(s, z) \mapsto \tilde{c}(s, z, x)$  is continuous, and there exists a non-negative and decreasing function  $\bar{c} : [1, \infty) \rightarrow \mathbb{R}_+$  and a constant  $q^* \in \mathbb{N}$  such that for every indices  $\beta_1, \beta_2$ , with  $\beta_1 \leq q^*$  and  $\beta_2 \leq q^*$ , we have

$$\sup_{s \in [0, T]} \sup_{x \in \mathbb{R}} (|\tilde{c}(s, z, x)| + |\partial_z^{\beta_2} \partial_x^{\beta_1} \tilde{c}(s, z, x)|) \leq \bar{c}(z), \quad \forall z \in [1, \infty),$$

with

$$\int_{[1,\infty)} |\bar{c}(z)|^p \nu(dz) =: \bar{c}_p < \infty, \quad \forall p \geq 1. \quad (4)$$

*Remark.* We will use several times the following consequence of (4) and of Burkholder's inequality (see for example the Theorem 2.11 in [26]): We assume that  $\Phi(s, z, \omega)$  and  $\varphi(s, \omega)$  are two non-negative functions such that

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

Then for any  $p \geq 2$ ,

$$\mathbb{E} \left| \int_0^T \int_{[1,\infty)} \Phi(s, z, \omega) N_\nu(ds, dz) \right|^p \leq C \times C_p \mathbb{E} \int_0^T |\varphi(s, \omega)|^p ds, \quad (5)$$

where  $C_p = \max\{(\bar{c}_2)^{\frac{p}{2}}, \bar{c}_p, (\bar{c}_1)^p\}$  and  $C$  is a constant depending on  $p$  and  $T$ .

*Proof.* By compensating  $N_\nu$ , using Burkholder's inequality and (4), we have

$$\begin{aligned} \mathbb{E} \left| \int_0^T \int_{[1,\infty)} \Phi(s, z, \omega) N_\nu(ds, dz) \right|^p &\leq C (\mathbb{E} \left| \int_0^T \int_{[1,\infty)} \Phi(s, z, \omega) \widetilde{N}_\nu(ds, dz) \right|^p + \mathbb{E} \left| \int_0^T \int_{[1,\infty)} \Phi(s, z, \omega) \nu(dz) ds \right|^p) \\ &\leq C (\mathbb{E} \int_0^T \left( \int_{[1,\infty)} |\Phi(s, z, \omega)|^2 \nu(dz) \right)^{\frac{p}{2}} ds + \mathbb{E} \int_0^T \int_{[1,\infty)} |\Phi(s, z, \omega)|^p \nu(dz) ds + \mathbb{E} \int_0^T \left| \int_{[1,\infty)} \Phi(s, z, \omega) \nu(dz) \right|^p ds) \\ &\leq C \times C_p \mathbb{E} \int_0^T |\varphi(s, \omega)|^p ds. \end{aligned}$$

□

**Hypothesis 2.2** There exists a non-negative function  $\check{c} : [1, \infty) \rightarrow \mathbb{R}_+$  such that  $\int_{[1,\infty)} |\check{c}(z)|^p \nu(dz) =: \check{c}_p < \infty$ ,  $\forall p \geq 1$ , and

$$\left| \frac{\partial_x \tilde{c}(s, z, x)}{1 + \partial_x \tilde{c}(s, z, x)} \right| \leq \check{c}(z), \quad \forall s \in [0, T], x \in \mathbb{R}, z \in [1, \infty).$$

To avoid overburdening notation, since both hypotheses 2.1 and 2.2 apply, we will take  $\check{c}(z) = \bar{c}(z)$ .

**Hypothesis 2.3 (Ellipticity)** There exists a non-negative function  $\underline{c} : [1, \infty) \rightarrow \mathbb{R}_+$  such that for every  $s \in [0, T], x \in \mathbb{R}, z \in [1, \infty)$ ,

$$|\partial_z \tilde{c}(s, z, x)|^2 \geq \underline{c}(z) \quad \text{and} \quad |\tilde{c}(s, z, x)|^2 \geq \underline{c}(z).$$

**Hypothesis 2.4 (Sector condition)** This is a supplementary hypothesis concerning the measure  $\nu$ . Two version of this hypothesis will be used; we state them separately below. Let  $I_k = [k, k+1), k \in \mathbb{N}$  and  $m_k = \nu(I_k)$ .

(a) **Strong sector condition:** We say that the strong sector condition is satisfied if there exist constants  $\varepsilon_* > 0$  and  $\alpha_1 \geq \alpha_0 > \alpha_2 > 0$ , such that

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \frac{\varepsilon_*}{z^{1-\alpha_1}} dz \quad \text{for all } k \in \mathbb{N}, \quad (6)$$

$$\begin{aligned} \underline{c}(z) &\geq e^{-z^{\alpha_2}} && \text{for all } z \geq 1 \text{ and,} \\ \int_1^\infty \frac{|\bar{c}(z)|^p}{z^{1-\alpha_0}} dz &< \infty && \text{for all } p \geq 1. \end{aligned} \quad (7)$$

Notice that if (6) is true for some  $\alpha_1$ , then it is also true for any  $\alpha \leq \alpha_1$ . So (6) also implies

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k dz, \quad \text{with} \quad \varepsilon_k = \frac{\varepsilon_*}{(k+1)^{1-\alpha}},$$

for any  $\alpha \leq \alpha_1$ .

(b) **Weak sector condition:** We say that the weak sector condition holds if there exist constants  $\varepsilon_* > 0$  and  $\alpha > 0$ , such that for every  $k \in \mathbb{N}$ , we have

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \frac{\varepsilon_*}{z} dz \quad \text{for all } k \in \mathbb{N}, \quad (8)$$

$$\underline{c}(z) \geq \frac{1}{z^\alpha} \quad \text{for all } z \geq 1 \text{ and,}$$

$$\int_1^\infty \frac{|\tilde{c}(z)|^p}{z} dz < \infty \quad \text{for all } p \geq 1. \quad (9)$$

We notice that (8) also implies  $\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k dz$  with  $\varepsilon_k = \frac{\varepsilon_*}{k+1}$ .

*Remark.* Notice that hypotheses 2.1, 2.2 and 2.3 are analogous to those in [10] (2 – 7, 2 – 26, 2 – 24).

Henceforth, we will suppose that hypotheses 2.1-2.3 hold, as well as either 2.4(a) or 2.4(b).

## 2.2 Approximation

We come back now to equation (2). The goal of this paper is to replace the small jumps in (2) by a drift and a Brownian motion. In equation (2), the Poisson point measure  $N_\mu$  is not compensated, so the first step is to introduce a drift (see  $b_\varepsilon$  below) which represents the compensator. Afterwards, we introduce a space-time Brownian motion  $W_\mu$  in order to replace the "compensated small jumps":

$$\begin{aligned} X_t^\varepsilon &= x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, X_{s-}^\varepsilon) N_\mu(ds, dz) \\ &\quad + \int_0^t b_\varepsilon(s, X_s^\varepsilon) ds + \int_0^t \int_{(0, \varepsilon]} c(s, z, X_s^\varepsilon) W_\mu(dz, ds), \end{aligned} \quad (10)$$

where

$$b_\varepsilon(s, x) = \int_{(0, \varepsilon]} c(s, z, x) \mu(dz)$$

and  $W_\mu$  is a space-time Brownian motion with covariance measure  $\mu(dz)ds$ , which is independent of  $N_\mu$ .

Let us discuss this equation. We notice that we keep the "big jumps" with  $z > \varepsilon$  but we eliminate the "small jumps" with  $z \leq \varepsilon$ . We replace the "small jumps" by the drift with coefficient  $b_\varepsilon$  and by the stochastic integral with coefficient  $c$ . This stochastic integral is driven by the so called space-time Brownian motion  $W_\mu$ , as introduced by Walsh in [36]. The existence and uniqueness of the solution to this equation (10) are also given by Kunita (see [26], [27]).

We recall that we work on a fixed interval of time  $[0, T]$ . We now precise the filtration that we consider. Let

$$\begin{aligned} \mathcal{F}_t^W &= \sigma(W_\mu(\varphi \mathbb{1}_{[0, t]})) : \varphi \in L^2((0, 1] \times [0, T], \mu \times Leb), \\ \mathcal{F}_t^N &= \sigma(N_\mu(\varphi \mathbb{1}_{[0, t]})) : \varphi \in L^1((0, 1] \times [0, T], \mu \times Leb), \\ \mathcal{F}_t &= \mathcal{F}_t^W \vee \mathcal{F}_t^N, \end{aligned} \quad (11)$$

where  $Leb$  denotes the Lebesgue measure and

$$W_\mu(\varphi) = \int_0^T \int_{(0, 1]} \varphi(s, z) W_\mu(ds, dz), \quad N_\mu(\varphi) = \int_0^T \int_{(0, 1]} \varphi(s, z) N_\mu(ds, dz).$$

So,  $X_t^\varepsilon$  is  $\mathcal{F}_t$ -measurable and  $X_t$  is  $\mathcal{F}_t^N$ -measurable.

We denote

$$L^2(W) = \{F \in \mathcal{F}_T^W : \mathbb{E}|F|^2 < \infty\}, \quad L^2(N) = \{G \in \mathcal{F}_T^N : \mathbb{E}|G|^2 < \infty\}. \quad (12)$$

*Remark.* Let  $\Phi$  be an adapted and piecewise constant process, that is

$$\Phi(s, z, \omega) = \sum_{i=1}^n \sum_{j=1}^m \Phi_{i,j}(\omega) \mathbb{1}_{[s_i, s_{i+1})}(s) \mathbb{1}_{B_j}(z),$$

where  $0 \leq s_1 < \dots < s_n, i = 1, \dots, n, B_j \in \mathcal{B}((0, 1]), j = 1, \dots, m$ , are disjoint sets. Suppose that  $\Phi_{i,j}$  are  $\mathcal{F}_{s_i}^W$ -measurable for all  $j = 1, \dots, m$ , and  $\sup_{i,j} \mathbb{E}|\Phi_{i,j}|^2 < \infty$ . Then for every  $G \in L^2(N)$ , we have

$$\mathbb{E}\left[G \times \int_0^T \int_{(0,1]} \Phi(s, z, \omega) W_\mu(ds, dz)\right] = 0. \quad (13)$$

*Proof.* Since  $W_\mu([s_i, s_{i+1}) \times B_j)$  is centered and independent of  $\Phi_{i,j}$  and of  $G$ , for all  $i = 1, \dots, n, j = 1, \dots, m$ , it follows that  $\mathbb{E}[G \Phi_{i,j} W_\mu([s_i, s_{i+1}) \times B_j)] = 0$ . Then it extends by linearity, and so (13) is true.  $\square$

Now we write the infinitesimal operator of  $X_s$  and  $X_s^\varepsilon$ , respectively: For  $\phi \in C_b^3(\mathbb{R})$  (the space of functions with continuous and bounded derivatives up to order 3),

$$\begin{aligned} L_s \phi(x) &= \int_{(0,1]} (\phi(x + c(s, z, x)) - \phi(x)) \mu(dz) \quad \text{and} \\ L_s^\varepsilon \phi(x) &= \int_{\{z > \varepsilon\}} (\phi(x + c(s, z, x)) - \phi(x)) \mu(dz) + \phi'(x) b_\varepsilon(s, x) + \frac{1}{2} \phi''(x) a_\varepsilon(s, x), \end{aligned} \quad (14)$$

where

$$a_\varepsilon(s, x) = \int_{(0, \varepsilon]} |c(s, z, x)|^2 \mu(dz).$$

Using Taylor's formula of order 2, we find

$$L_s \phi(x) = \int_{\{z > \varepsilon\}} (\phi(x + c(s, z, x)) - \phi(x)) \mu(dz) + \phi'(x) b_\varepsilon(s, x) + \frac{1}{2} \phi''(x) a_\varepsilon(s, x) + R_s(x),$$

where

$$|R_s(x)| \leq \frac{1}{6} \|\phi\|_{3, \infty} \int_{(0, \varepsilon]} |c(s, z, x)|^3 \mu(dz),$$

with  $\|\phi\|_{l, \infty} := \sum_{0 \leq i \leq l} \|\phi^{(i)}\|_\infty$ , the sum of all the uniform norms of the derivatives of function  $\phi$  up to order  $l$ . In conclusion, we find

$$\|(L_s - L_s^\varepsilon) \phi\|_\infty = \|R_s\|_\infty \leq \frac{1}{6} \|\phi\|_{3, \infty} \eta_3(\varepsilon), \quad (15)$$

with

$$\eta_p(\varepsilon) = \int_{(0, \varepsilon]} |\bar{c}(1/z)|^p \mu(dz) = \int_{[\varepsilon^{-1}, \infty)} |\bar{c}(z)|^p \nu(dz), \quad p \geq 1. \quad (16)$$

Then, we can give an estimate of the distance between the semigroups. We use the standard semigroup notation, which we remind below. Let  $[X_t(s, x)]_{t \geq s}$  and  $[X_t^\varepsilon(s, x)]_{t \geq s}$  be the solutions to (2) and (10), respectively, starting at time  $s$  from point  $x$ . Denote by  $P_{s,t} \phi(x) = \mathbb{E} \phi(X_t(s, x))$  and  $P_{s,t}^\varepsilon \phi(x) = \mathbb{E} \phi(X_t^\varepsilon(s, x))$ . Also, set  $P_t := P_{0,t}$  and  $P_t^\varepsilon := P_{0,t}^\varepsilon$ .

**Lemma 2.1.** *There exists a constant  $C$  depending on  $T$  such that for  $\phi \in C_b^3(\mathbb{R})$  and  $0 \leq t \leq T$ , we have*

$$\|P_t \phi - P_t^\varepsilon \phi\|_\infty \leq C \|\phi\|_{3, \infty} \eta_3(\varepsilon). \quad (17)$$

*Proof. Step 1 Trotter-Kato method:* We know from Kunita [27] (Theorem 4.5.1) that we have the Kolmogorov forward and backward equations:

$$\partial_t P_{s,t} \phi(x) = P_{s,t} L_t \phi(x), \quad \partial_t P_{s,t}^\varepsilon \phi(x) = P_{s,t}^\varepsilon L_t^\varepsilon \phi(x); \quad (18)$$

$$\partial_s P_{s,t} \phi(x) = -L_s P_{s,t} \phi(x), \quad \partial_s P_{s,t}^\varepsilon \phi(x) = -L_s^\varepsilon P_{s,t}^\varepsilon \phi(x). \quad (19)$$

Then using Newton-Leibniz's formula and (18), (19),

$$P_t^\varepsilon \phi(x) - P_t \phi(x) = \int_0^t \partial_s (P_{0,s}^\varepsilon P_{s,t}) \phi(x) ds = \int_0^t (P_{0,s}^\varepsilon (L_s^\varepsilon - L_s) P_{s,t}) \phi(x) ds.$$

It follows that

$$\begin{aligned} \|P_t \phi - P_t^\varepsilon \phi\|_\infty &\leq \int_0^t \|P_{0,s}^\varepsilon (L_s^\varepsilon - L_s) P_{s,t} \phi\|_\infty ds \\ &\leq \int_0^t \|(L_s^\varepsilon - L_s) P_{s,t} \phi\|_\infty ds \\ &\leq \frac{1}{6} \eta_3(\varepsilon) \int_0^t \|P_{s,t} \phi\|_{3,\infty} ds. \end{aligned} \quad (20)$$

**Step 2 (propagation of regularity)** In [27], Kunita has shown in Theorem 3.4.1 and Theorem 3.4.2 the regularity of the flow associated with the jump-diffusion. So in our case, we have

$$\begin{aligned} \|P_{s,t} \phi\|_{3,\infty} &\leq \sup_{x \in \mathbb{R}} (\mathbb{E}|\phi(X_t(s, x))| + \mathbb{E}|\partial_x \phi(X_t(s, x))| + \mathbb{E}|\partial_x^2 \phi(X_t(s, x))| + \mathbb{E}|\partial_x^3 \phi(X_t(s, x))|) \\ &\leq \|\phi\|_{3,\infty} \sup_{x \in \mathbb{R}} \mathbb{E}[1 + 3|\partial_x X_t(s, x)| + 3|\partial_x^2 X_t(s, x)| + |\partial_x^3 X_t(s, x)|] \leq C \|\phi\|_{3,\infty}. \end{aligned} \quad (21)$$

Substituting (21) into (20), we obtain (17).  $\square$

*Remark.* A similar result has been obtained in [21] (Theorem 4.7). Besides, one may also consider an approximate equation obtained just by discarding the small jumps:

$$\widetilde{X}_t^\varepsilon = x + \int_0^t \int_{\{z > \varepsilon\}} c(s, z, \widetilde{X}_{s-}^\varepsilon) N_\mu(ds, dz).$$

Then, if  $\widetilde{L}_s^\varepsilon$  is the infinitesimal operator of  $\widetilde{X}_s^\varepsilon$ , we have  $\|(L_s - \widetilde{L}_s^\varepsilon) \phi\|_\infty \leq \|\phi\|_{1,\infty} \eta_1(\varepsilon)$ . So the same reasoning as above gives

$$\|P_t \phi - \widetilde{P}_t^\varepsilon \phi\|_\infty \leq C \|\phi\|_{1,\infty} \times \eta_1(\varepsilon) \rightarrow 0. \quad (22)$$

The gain in (17) is that we have  $\eta_3(\varepsilon)$  instead of  $\eta_1(\varepsilon)$  in (22), which means that we have a faster speed of convergence.

## 2.3 The main theorem

We are finally ready to state the main results of this paper. Denote by  $d_{TV}(F, G)$  the total variation distance between the laws of two random variables  $F$  and  $G$ .

**Theorem 2.2.** Assume that **Hypotheses 2.1, 2.2. and 2.3** hold with  $q^* \geq \frac{3}{\delta} + 1$  for some  $\delta > 0$ .

(a) If in addition we assume **Hypothesis 2.4 (a)**, then there exists a constant  $C$  depending on  $\delta$  and  $T$  such that

$$d_{TV}(X_t, X_t^\varepsilon) \leq C \eta_3(\varepsilon)^{1-\delta}. \quad (23)$$

Under the above hypotheses, the laws of  $X_t$  and  $X_t^\varepsilon$  are absolutely continuous with respect to the Lebesgue measure, with smooth densities  $p_{X_t}(x)$  and  $p_{X_t^\varepsilon}(x)$ . Moreover, if  $l$  is an index such that  $q^* \geq \frac{3+l}{\delta} + 1$ , then there exists a constant  $C$  depending on  $\delta$ ,  $T$  and  $l$  such that

$$\|p_{X_t} - p_{X_t^\varepsilon}\|_{l,\infty} \leq C \eta_3(\varepsilon)^{1-\delta}. \quad (24)$$

(b) If in addition we assume **Hypothesis 2.4** (b), then there exists a constant  $C$  depending on  $\delta$  and  $T$  such that for every  $t \in [0, T]$  with  $t > \frac{8\alpha(\frac{3}{\delta}-1)}{\varepsilon_*}$  (with  $\varepsilon_*$  and  $\alpha$  given in **Hypothesis 2.4** (b)), we have

$$d_{TV}(X_t, X_t^\varepsilon) \leq C\eta_3(\varepsilon)^{1-\delta}. \quad (25)$$

For any index  $l$  and for  $t > \frac{8\alpha(3l+2)}{\varepsilon_*}$ , both the laws of  $X_t$  and  $X_t^\varepsilon$  have  $l$ -times differentiable densities  $p_{X_t}(x)$  and  $p_{X_t^\varepsilon}(x)$ . Assume moreover that  $q^* \geq \frac{3+l}{\delta} + 1$ . Then there exists a constant  $C$  depending on  $\delta, T$  and  $l$  such that for  $t > \max\{\frac{8\alpha}{\varepsilon_*}(\frac{3+l}{\delta} - 1), \frac{8\alpha(3l+2)}{\varepsilon_*}\}$ , we have

$$\|p_{X_t} - p_{X_t^\varepsilon}\|_{l,\infty} \leq C\eta_3(\varepsilon)^{1-\delta}. \quad (26)$$

The proof of this theorem is left to Section 4.4.

*Remark.* Some recent results concerning the weak approximation of the SDE with jumps are also given in [14], [22], [23] for example. But they do not concern the convergence in total variation distance.

## 2.4 A typical example

For  $t \in [0, T]$ , we consider the following SDE driven by a Lévy process:

$$X_t = x + \int_0^t \sigma(X_{s-}) dZ_s, \quad (27)$$

where  $(Z_t)_{t \in [0, T]}$  is a Lévy process of Lévy triplet  $(0, 0, \mu)$ , with  $\mu(dz) = \mathbb{1}_{(0,1]}(z) \frac{dz}{z^{1+\rho}}$ ,  $0 \leq \rho < 1$ .

We approximate (27) by

$$X_t^\varepsilon = x + \int_0^t \sigma(X_{s-}^\varepsilon) dZ_s^\varepsilon + b(\varepsilon) \int_0^t \sigma(X_s^\varepsilon) ds + c(\varepsilon) \int_0^t \sigma(X_s^\varepsilon) dB_s, \quad (28)$$

where  $(Z_t^\varepsilon)_{t \in [0, T]}$  is a Lévy process of Lévy triplet  $(0, 0, \mathbb{1}_{\{z > \varepsilon\}} \mu(dz))$ ,  $(B_t)_{t \in [0, T]}$  is a standard Brownian motion independent of  $(Z_t^\varepsilon)_{t \in [0, T]}$ , and

$$b(\varepsilon) = \int_{(0,\varepsilon]} z \mu(dz), \quad c(\varepsilon) = \sqrt{\int_{(0,\varepsilon]} z^2 \mu(dz)}.$$

Then we have the following theorem.

**Theorem 2.3.** We assume that  $\sigma \in C_b^\infty(\mathbb{R})$ ,  $0 < \underline{\sigma} \leq \sigma(x) \leq \bar{\sigma}$  and  $-1 < a \leq \sigma'(x) \leq \bar{\sigma}$ ,  $\forall x \in \mathbb{R}$ , for some universal constants  $\bar{\sigma}, \underline{\sigma}, a$ , where  $\sigma'$  is the differential of  $\sigma$  in  $x$ . Then for any  $\delta > 0$ , there is a constant  $C > 0$  such that for any  $t \in [0, T]$ ,

$$d_{TV}(X_t, X_t^\varepsilon) \leq C\varepsilon^{3-\rho-\delta}.$$

Moreover, the laws of  $X_t$  and  $X_t^\varepsilon$  have smooth densities  $p_{X_t}(x)$  and  $p_{X_t^\varepsilon}(x)$  respectively. And for any index  $l$  and any  $\delta > 0$ , there exists a constant  $C > 0$  such that

$$\|p_{X_t} - p_{X_t^\varepsilon}\|_{l,\infty} \leq C\varepsilon^{3-\rho-\delta}.$$

*Proof.* We notice that

$$Z_t = \int_0^t \int_{(0,1]} z N_\mu(ds, dz),$$

where  $N_\mu$  is a Poisson point measure with intensity  $\mu(dz)ds$ . Then (27) coincides with (2) with  $c(s, z, x) = \sigma(x)z$ , and (28) coincides with (10) with  $c(s, z, x) = \sigma(x)z$ ,  $b_\varepsilon(s, x) = b(\varepsilon)\sigma(x)$ , and  $\int_{\{z \leq \varepsilon\}} z W_\mu(ds, dz) = c(\varepsilon)dB_s$ .

Let  $\theta : (0, 1] \rightarrow [1, \infty)$  be a function defined by  $\theta(z) = \frac{1}{z}$ . By a change of variables,

$$\tilde{c}(s, z, x) = c(s, \frac{1}{z}, x) = \sigma(x) \times \frac{1}{z}, \quad \nu(dz) = \mu \circ \theta^{-1}(dz) = \mathbb{1}_{[1, \infty)}(z) \frac{dz}{z^{1-\rho}}.$$

One can easily check that **Hypothesis 2.1** is verified (for every  $q^* \in \mathbb{N}$ ) with  $\bar{c}(z) = \bar{\sigma} \times \frac{1}{z}$  and

$$\int_1^\infty |\bar{c}(z)|^p \nu(dz) = \int_1^\infty \frac{\bar{\sigma}^p}{z^{p+1-\rho}} dz < \infty, \quad \forall p \geq 1.$$

We recall that  $I_k = [k, k+1)$ ,  $k \in \mathbb{N}$  and  $m_k = \nu(I_k)$ . Then for sufficiently large  $z$ , we have

$$\min\{|\partial_z \tilde{c}(s, z, x)|^2, |\tilde{c}(s, z, x)|^2\} \geq \underline{\sigma}^2 \times \frac{1}{z^4} \geq e^{-z^{\alpha_2}},$$

with some  $0 < \alpha_2 < 1$ . We also have

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \frac{1}{2} dz \geq \mathbb{1}_{I_k}(z) \frac{1}{2} \frac{dz}{z^{1-\alpha_0}},$$

with some  $\alpha_0 \in (\alpha_2, 1)$ . Moreover, since for any  $p \geq 1$ ,  $p+1-\alpha_0 > 1$ , we have

$$\int_1^\infty \frac{|\bar{c}(z)|^p}{z^{1-\alpha_0}} dz = \bar{\sigma} \int_1^\infty \frac{1}{z^{p+1-\alpha_0}} dz < \infty.$$

So **Hypothesis 2.3** and **Hypothesis 2.4** (a) are satisfied. Finally,

$$\left| \frac{\partial_x \tilde{c}(s, z, x)}{1 + \partial_x \tilde{c}(s, z, x)} \right| \leq \frac{\bar{\sigma} \times \frac{1}{z}}{1 + a \times \frac{1}{z}} \leq \max\left\{\frac{1}{1+a}, 1\right\} \times \bar{\sigma} \times \frac{1}{z},$$

so **Hypothesis 2.2** is satisfied as well. Then we can apply **Theorem 2.2(a)** for the equation (27) and (28). Since

$$\eta_3(\varepsilon) = \int_{(0, \varepsilon]} \bar{\sigma}^3 \times z^3 \mu(dz) = \frac{\bar{\sigma}^3}{3-\rho} \varepsilon^{3-\rho},$$

we obtain the estimates from **Theorem 2.3**. □

### 3 Abstract integration by parts framework

In order to obtain the main theorem (**Theorem 2.2**), we will apply some techniques of Malliavin calculus. So firstly, we give the abstract integration by parts framework introduced in [6]. This is a variant of the integration by parts framework given in [10].

We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , 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  $\phi(F) \in \mathcal{S}$  (with  $C_p^\infty$  the space of smooth functions which, together with all the derivatives, have polynomial growth). 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  $D : \mathcal{S} \rightarrow \bigcap_{p=1}^\infty L^p(\Omega; \mathcal{H})$  which is a linear application which satisfies

a)

$$D_h F := \langle DF, h \rangle_{\mathcal{H}} \in \mathcal{S}, \text{ for any } h \in \mathcal{H}, \tag{29}$$

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

$$D\phi(F) = \sum_{i=1}^d \partial_i \phi(F) DF_i, \quad (30)$$

Since  $D_h F \in \mathcal{S}$ , we may define by iteration the derivative operator of higher order  $D^q : \mathcal{S} \rightarrow \bigcap_{p=1}^{\infty} L^p(\Omega; \mathcal{H}^{\otimes q})$  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  $D_{h_1, \dots, h_q}^q F := D_{h_q} D_{h_{q-1}} \cdots D_{h_1} F$ . Then,  $D_{h_1, \dots, h_q}^q F = D_{h_q} D_{h_1, \dots, h_{q-1}}^{q-1} F$  ( $q \geq 2$ ).

For  $F = (F_1, \dots, F_{\tilde{d}}) \in \mathcal{S}^{\tilde{d}}$ , we define  $\sigma_F = (\sigma_F^{i,j})_{i,j=1, \dots, \tilde{d}}$  to be the Malliavin covariance matrix with  $\sigma_F^{i,j} = \langle DF_i, DF_j \rangle_{\mathcal{H}}$  and we denote

$$\Sigma_p(F) = \mathbb{E}(1/\det \sigma_F)^p. \quad (31)$$

For  $\tilde{d} = 1$ , which is the case that we discuss in this paper,  $\det \sigma_F = \sigma_F = \langle DF, DF \rangle_{\mathcal{H}}$ . 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 operator  $L : \mathcal{S} \rightarrow \mathcal{S}$  which is a linear operator satisfying the following duality formula:

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

$$\mathbb{E}\langle DF, DG \rangle_{\mathcal{H}} = \mathbb{E}(FLG) = \mathbb{E}(GLF). \quad (32)$$

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

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

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

$$|F|_{1,l} = \sum_{q=1}^l |D^q F|_{\mathcal{H}^{\otimes q}}, \quad |F|_l = |F| + |F|_{1,l}. \quad (33)$$

We remark that  $|F|_0 = |F|$  and  $|F|_{1,l} = 0$  for  $l = 0$ . For  $F = (F_1, \dots, F_d) \in \mathcal{S}^d$ , 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  $l, p \geq 1$ ,

$$\begin{aligned} \|F\|_{l,p} &= (\mathbb{E}|F|_l^p)^{1/p}, \quad \|F\|_p = (\mathbb{E}|F|^p)^{1/p}, \\ \|F\|_{L,l,p} &= \|F\|_{l,p} + \|LF\|_{l-2,p}. \end{aligned} \quad (34)$$

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

$$\mathcal{D}_{l,p} = \overline{\mathcal{S}}^{\|\cdot\|_{L,l,p}}, \quad (35)$$

and

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

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  $\|F - F_n\|_p \rightarrow 0, \|F_n - F_n\|_{q,p} \rightarrow 0$  and  $\|LF_n - LF_n\|_{q-2,p} \rightarrow 0$ . Since  $D^q$  and  $L$  are closable, we can define

$$D^q F = \lim_{n \rightarrow \infty} D^q F_n \quad \text{in } L^p(\Omega; \mathcal{H}^{\otimes q}), \quad LF = \lim_{n \rightarrow \infty} LF_n \quad \text{in } L^p(\Omega). \quad (36)$$

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



**Lemma 3.1.** *The triplet  $(\mathcal{D}_\infty, D, L)$  is an IbP framework.*

*Proof.* Here we just show that  $D$  verifies (29): For  $F \in \mathcal{D}_\infty$  and  $h \in \mathcal{H}$ , we have  $\langle DF, h \rangle_{\mathcal{H}} \in \mathcal{D}_\infty$ .

In fact, for any  $k \geq 1, p \geq 2$ , any  $F \in \mathcal{D}_{k+1,p}$ , there is a sequence  $F_n \in \mathcal{S}$  such that  $\|F_n - F\|_{k+1,p} \rightarrow 0$ . Then for any  $u_1, \dots, u_k \in L^p(\Omega; \mathcal{H}), h \in \mathcal{H}$ , any  $n, m \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E} \langle D^k(\langle DF_m, h \rangle_{\mathcal{H}} - \langle DF_n, h \rangle_{\mathcal{H}}), u_1 \otimes \dots \otimes u_k \rangle_{\mathcal{H}^{\otimes k}}^{\frac{p}{2}} &= \mathbb{E} |D_{u_k} D_{u_{k-1}} \dots D_{u_1} \langle D(F_m - F_n), h \rangle_{\mathcal{H}}|^{\frac{p}{2}} \\ &= \mathbb{E} |D_{u_k} D_{u_{k-1}} \dots D_{u_1} D_h(F_m - F_n)|^{\frac{p}{2}} = \mathbb{E} |D^{k+1}(F_m - F_n), h \otimes u_1 \otimes \dots \otimes u_k|_{\mathcal{H}^{\otimes(k+1)}}^{\frac{p}{2}} \\ &\leq \mathbb{E} |D^{k+1}(F_m - F_n)|_{\mathcal{H}^{\otimes(k+1)}}^{\frac{p}{2}} |h \otimes u_1 \otimes \dots \otimes u_k|_{\mathcal{H}^{\otimes(k+1)}}^{\frac{p}{2}} \rightarrow 0, \end{aligned}$$

which yields that  $\mathbb{E} |D^k(\langle DF_m, h \rangle_{\mathcal{H}} - \langle DF_n, h \rangle_{\mathcal{H}})|_{\mathcal{H}^{\otimes k}}^p \rightarrow 0$ . Therefore,  $\langle DF, h \rangle_{\mathcal{H}} \in \mathcal{D}_{k,p}$  and (29) is verified.  $\square$

The following lemma is useful in order to control the Sobolev norms  $\|F\|_{L,l,q}$ .

**Lemma 3.2.** *We fix  $p \geq 2, l \geq 2$ . Let  $F \in L^1(\Omega)$  and let  $F_n \in \mathcal{S}, n \in \mathbb{N}$  such that*

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

*Then for every  $1 \leq \bar{p} < p$ , we have  $F \in \mathcal{D}_{l,\bar{p}}$  and  $\|F\|_{L,l,\bar{p}} \leq K_{l,\bar{p}}$ .*

*Proof.* 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's theorem, there exists a functional  $G \in \mathcal{H}_l$  and a subsequence (we still denote it by  $n$ ), such that  $F_n \rightarrow G$  weakly in the Hilbert space  $\mathcal{H}_l$ . This means that for every  $Q \in \mathcal{H}_l$ ,  $\langle F_n, Q \rangle_{L,l,2} \rightarrow \langle G, Q \rangle_{L,l,2}$ . Therefore, by Mazur's theorem, we can construct some convex combination

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

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

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

In particular we have

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

Also, we notice that by *i)*,

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

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

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

By passing to a subsequence, we have  $|G_n - F|_l + |LG_n - LF|_{l-2} \rightarrow 0$  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,  $Y_n \rightarrow Y$  almost surely, and for any  $\tilde{q} \in [\bar{p}, p]$ ,

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

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

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

□

### 3.1 Main consequences: Convergence in total variation distance

We will use the abstract framework in [6] for the IbP framework  $(\mathcal{D}_\infty, D, L)$ , with  $D$  and  $L$  defined in (36). Using Malliavin type arguments, [6] proves the following results. The first result, concerning the density, is classical:

**Lemma 3.3.** *Let  $F \in \mathcal{D}_\infty$ . If  $\Sigma_{6p+4}(F) < \infty$ , then the law of random variable  $F$  has a density  $p_F(x)$  which is  $p$ -times differentiable.*

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

$$d_k(F, G) = \sup\{|\mathbb{E}(f(F)) - \mathbb{E}(f(G))| : \sum_{0 \leq i \leq k} \|f^{(i)}\|_\infty \leq 1\}$$

For  $k = 1$ , this is the Fortet Mourier distance (which is a variant of the Wasserstein distance), while for  $k = 0$ , this is the total variation distance and we denote it by  $d_{TV}$ . Now we present the second result concerning the total variation distance:

**Lemma 3.4.** *We fix some index  $l$ , some  $r \in \mathbb{N}$  and some  $\delta > 0$ . We define  $p_1 = 2(r(\frac{1}{\delta} - 1) + 2)$ ,  $p_2 = \max\{6l + 4, 2(\frac{r+l}{\delta} - r + 2)\}$ ,  $q_1 \geq r(\frac{1}{\delta} - 1) + 4$ ,  $q_2 \geq \frac{r+l}{\delta} - r + 4$ . Let  $F, G \in \mathcal{D}_\infty$ . Then one may find  $C \in \mathbb{R}_+$ ,  $p \in \mathbb{N}$  (depending on  $r$  and  $\delta$ ) such that*

$$i) \quad d_{TV}(F, G) \leq C(1 + \Sigma_{p_1}(F) + \Sigma_{p_1}(G) + \|F\|_{L,q_1,p} + \|G\|_{L,q_1,p}) \times d_r(F, G)^{1-\delta}, \quad (37)$$

and

$$ii) \quad \|p_F - p_G\|_{l,\infty} \leq C(1 + \Sigma_{p_2}(F) + \Sigma_{p_2}(G) + \|F\|_{L,q_2,p} + \|G\|_{L,q_2,p}) \times d_r(F, G)^{1-\delta}, \quad (38)$$

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

**Comment** The significance of this lemma is the following. Suppose that one has already obtained an estimate of a "smooth" distance  $d_r$  between two random vectors  $F$  and  $G$  (in our case  $r = 3$  in (17)). But we would like to control the total variation distance between them. In order to do this, one employs some integration by parts techniques which are developed in [6] and conclude the following. We need to assume that both  $F$  and  $G$  are "smooth" in the sense that  $\|F\|_{L,q,p} + \|G\|_{L,q,p} < \infty$  for sufficiently large  $q, p$ . Moreover, we need some non degeneracy condition: both  $F$  and  $G$  are non-degenerated, that is  $\Sigma_p(F) + \Sigma_p(G) < \infty$ , with  $p$  large enough. Then (37) asserts that one may control  $d_{TV}$  by  $d_r$ , and the control is quasi optimal: we loose just a power  $\delta > 0$  which we may take as small as we want. And (38) says that we may also control the distance between the derivatives of density functions by  $d_r$ .

Then we can get the following corollary.

**Corollary 3.4.1.** *We fix some index  $l$ , some  $r \in \mathbb{N}$  and some  $\delta > 0$ . We define  $p_1, p_2, q_1, q_2$  as in **Lemma 3.4**. Let  $F_M \in \mathcal{D}_\infty$ ,  $M \in \mathbb{N}$  such that for every  $p \geq 1$ ,*

$$\sup_M (\|F_M\|_{L, q_1, p} + \Sigma_{p_1}(F_M)) \leq Q_{q_1, p, p_1} < \infty,$$

*with  $Q_{q_1, p, p_1}$  a constant not dependent on  $M$ . Consider moreover some random variable  $F$  such that  $d_r(F, F_M) \rightarrow 0$ . Then there exists a constant  $C > 0$  such that*

$$i) \quad d_{TV}(F, F_M) \leq C d_r(F, F_M)^{1-\delta}.$$

*Moreover, if  $\sup_M (\|F_M\|_{L, q_2, p} + \Sigma_{p_2}(F_M)) \leq Q_{q_2, p, p_2} < \infty$ , the law of  $F$  is absolutely continuous with smooth density  $p_F$  and one has*

$$ii) \quad \|p_F - p_{F_M}\|_{l, \infty} \leq C d_r(F, F_M)^{1-\delta}.$$

*Proof.* We take  $C$  to be a constant depending on  $p, p_1, q_1, r$  and  $\delta$  which can change from one line to another. By **Lemma 3.4**, for every  $M < M'$ , one has

$$d_{TV}(F_M, F_{M'}) \leq C d_r(F_M, F_{M'})^{1-\delta} \leq C [d_r(F_M, F)^{1-\delta} + d_r(F, F_{M'})^{1-\delta}]. \quad (39)$$

So  $(F_M)_{M \in \mathbb{N}}$  is a Cauchy sequence in  $d_{TV}$ . It follows that it has a limit  $G$ . But since  $d_r(F_M, F) \rightarrow 0$ , it follows that  $F = G$ . Passing to the limit  $M' \rightarrow \infty$  in (39), we get

$$d_{TV}(F_M, F) \leq C d_r(F_M, F)^{1-\delta}.$$

The proof of *ii*) is analogous. □

## 4 Malliavin calculus and stochastic differential equations with jumps

In this section we present the integration by parts framework that will be used in the following. To begin we give a quick informal presentation of our strategy. We will work with the solution of the equation (10), but, for technical reasons, we make the change of variable  $z \mapsto \frac{1}{z}$  so the equation of interest is now the equation (49). We use the notation from that section. The intensity measure for our random measure is  $\mathbb{1}_{[1, M)}(z) \nu(dz) ds$  and this is a finite measure. Then the corresponding Poisson Point measure  $N_\nu$  may be represented by means of a compound Poisson process. For some technical reasons, we produce the representation on each set  $\{z \in I_k = [k, k+1)\}, k \in \mathbb{N}$ , so the equation (49) reads

$$\begin{aligned} \hat{X}_t^M &= x + \int_0^t \sum_{k=1}^{M-1} \int_{\{z \in I_k\}} \tilde{c}(s, z, \hat{X}_{s-}^M) N_\nu(ds, dz) \\ &\quad + \int_0^t b_M(s, \hat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \tilde{c}(s, z, \hat{X}_s^M) W_\nu(ds, dz) \\ &= x + \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \tilde{c}(T_i^k, Z_i^k, \hat{X}_{T_i^k-}^M) \\ &\quad + \int_0^t b_M(s, \hat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \tilde{c}(s, z, \hat{X}_s^M) W_\nu(ds, dz). \end{aligned}$$

Here  $T_i^k, k, i \in \mathbb{N}$  are the jump times of the Poisson process  $(J_t^k)_{t \in [0, T]}$  of parameter  $\nu(I_k)$ , and  $Z_i^k, k, i \in \mathbb{N}$  are independent random variables of law  $\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{\nu(I_k)}$ , which are independent of  $J^k$  as well. We will work conditionally to  $T_i^k, k, i \in \mathbb{N}$ , so the randomness in the system comes from  $W_\nu$  on one hand and from  $Z_i^k, k, i \in \mathbb{N}$  on the other hand. Concerning  $W_\nu$  we will use the standard Malliavin calculus (which fits in

the framework presented in Section 3). But we will also use this integration by parts calculus with respect to the amplitude of the jumps given by  $Z_i^k, k, i \in \mathbb{N}$ . We present this kind of calculus now.

Suppose for a moment (just for simplicity) the law of  $Z_i^k$  is absolutely continuous with respect to the Lebesgue measure and has a smooth density  $h_k(z)$  which has compact support. We also assume that the logarithm of the density  $\ln h_k$  is smooth. Then we look to  $\hat{X}_t^M$  as to a functional  $F(Z_1^1, \dots, Z_{J_t^k}^{M-1})$  and we define the derivative operators

$$D_{k,i}^Z F = \frac{\partial}{\partial z_i^k} F(Z_1^1, \dots, Z_{J_t^k}^{M-1}),$$

and

$$L^Z F = - \sum_{k,i} D_{k,i}^Z D_{k,i}^Z F + D_{k,i}^Z F \times \partial_z \ln h_k(Z_i^k).$$

And we check that these operators verify the conditions in Section 3. Since we want to use integration by parts with respect to both  $W_\nu$  and  $Z_i^k, k, i \in \mathbb{N}$ , we will consider the derivative operator  $D = (D^W, D^Z)$  and the operator  $L = (L^W, L^Z)$  where  $D^W$  and  $L^W$  are the derivative and Ornstein Uhlenbeck operators from the standard Malliavin calculus for Gaussian random variables. With these operators at hand we check the hypotheses of **Lemma 3.4** and of **Corollary 3.4.1**, and these are the results which allow as to prove our **Theorem 2.2**.

Roughly speaking this is our strategy. But there is one more point: the hypotheses we raise for the law of  $Z_i^k$  that it has a smooth density with compact support and has a smooth logarithm density, is rather strong and we want to weaken it. This is the aim of the "splitting method". This amounts to produce three independent random variables  $V_i^k, U_i^k$  and  $\xi_i^k$  such that  $Z_i^k$  has the same law as  $\xi_i^k V_i^k + (1 - \xi_i^k) U_i^k$  with  $\xi_i^k$  a Bernoulli random variable and  $V_i^k$  a random variable with good properties. So we split  $Z_i^k$  in two parts,  $V_i^k$  and  $U_i^k$ . We may do it in such a way that  $V_i^k$  has the law  $\psi_k(v)dv$  with  $\psi_k \in C_c^\infty(\mathbb{R})$  (see Section 4.1 for the precise procedure). And we perform the Malliavin calculus with respect to  $V_i^k$  instead of  $Z_i^k$  (we work conditionally to  $\xi_i^k$  and  $U_i^k$  which appear as constants).

## 4.1 The splitting method

We consider a Poisson point measure  $N_\nu(ds, dz)$  with compensator  $\hat{N}_\nu(ds, dz) = \nu(dz)ds$  on the state space  $[1, \infty)$ . We will make use of the noise  $z \in [1, \infty)$  in order to apply the results from the previous section. We recall that  $I_k = [k, k+1)$  and  $m_k = \nu(I_k)$ , and we suppose that for every  $k$ , there exists  $\varepsilon_k > 0$ , such that

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k \times dz. \quad (40)$$

*Remark.* Under **Hypothesis 2.4 (a)**, the splitting condition (40) is satisfied with  $\varepsilon_k = \frac{\varepsilon_*}{(k+1)^{1-\alpha}}$ , for any  $\alpha \leq \alpha_1$ . If instead we assume **Hypothesis 2.4 (b)**, (40) is also satisfied, with  $\varepsilon_k = \frac{\varepsilon_*}{k+1}$ .

When (40) is satisfied, we are able to use the "splitting method" as follows. To begin we define the functions

$$a(y) = 1 - \frac{1}{1 - (4y - 1)^2} \quad \text{for } y \in [\frac{1}{4}, \frac{1}{2}] \quad (41)$$

$$\psi(y) = \mathbb{1}_{\{|y| \leq \frac{1}{4}\}} + \mathbb{1}_{\{\frac{1}{4} < |y| \leq \frac{1}{2}\}} e^{a(|y|)}. \quad (42)$$

We notice that  $\psi \in C_c^\infty(\mathbb{R})$  and that its support is included in  $[-\frac{1}{2}, \frac{1}{2}]$ . We also notice that for every  $q, p \in \mathbb{N}$  the function  $y \mapsto |a^{(q)}(y)|^p \psi(y)$  is continuous and has support included in  $[-\frac{1}{2}, \frac{1}{2}]$ , so it is bounded: one may find  $C_{q,p}$  such that

$$|a^{(q)}(y)|^p \psi(y) \leq C_{q,p} \quad \forall y \in \mathbb{R}. \quad (43)$$

We denote

$$\psi_k(y) = \psi(y - (k + \frac{1}{2})), \quad \theta_k(y) := \partial_y \ln \psi_k(y). \quad (44)$$

By (43) (which is uniform with respect to  $y$ ), we have

$$\sup_k |( \ln \psi_k )^{(q)}(y)|^p \psi_k(y) \leq C_{q,p} \quad \forall y \in \mathbb{R}. \quad (45)$$

We denote

$$m(\psi) = \int_{-1/2}^{1/2} \psi(y) dy. \quad (46)$$

We consider a sequence of independent random variables  $Z^k$  such that

$$Z^k \sim \mathbb{1}_{I_k}(z) \frac{1}{m_k} \nu(dz).$$

This is the sequence of random variables which are involved in the representation of the measure  $N_\nu(ds, dz)$  as long as  $z \in [1, \infty)$  is concerned. We notice that, according to our hypothesis (40),

$$\mathbb{P}(Z^k \in dz) = \mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k \times dz.$$

Then we construct some independent random variables  $V^k, U^k, \xi^k$  with laws

$$\begin{aligned} \mathbb{P}(V^k \in dz) &= \frac{1}{m(\psi)} \psi(z - (k + \frac{1}{2})) dz \\ \mathbb{P}(U^k \in dz) &= \frac{1}{1 - \varepsilon_k m(\psi)} (\mathbb{P}(Z^k \in dz) - \varepsilon_k \psi(z - (k + \frac{1}{2})) dz) \\ \mathbb{P}(\xi^k = 1) &= \varepsilon_k m(\psi), \quad \mathbb{P}(\xi^k = 0) = 1 - \varepsilon_k m(\psi). \end{aligned} \quad (47)$$

We choose  $\varepsilon_k < 1/m(\psi)$  so that  $1 - \varepsilon_k m(\psi) > 0$ . Using (40), one may check that  $\mathbb{P}(U^k \in dz)$  is a positive measure and has mass one. So it is a probability measure. And finally one can easily check the identity of laws:

$$Z^k \sim \xi^k V^k + (1 - \xi^k) U^k. \quad (48)$$

In the following, we will work directly with  $Z^k = \xi^k V^k + (1 - \xi^k) U^k$ . This is possible because all the results that we discuss here concern the law of the random variables, and the law remains unchanged.

The Poisson point measure  $N_\nu$  can be written as the following sum:

$$N_\nu(ds, dz) = \sum_{k=1}^{\infty} \mathbb{1}_{I_k}(z) N_\nu(ds, dz) = \sum_{k=1}^{\infty} N_{\nu_k}(ds, dz),$$

where  $\nu_k(dz) = \mathbb{1}_{I_k}(z) \nu(dz)$  and  $N_{\nu_k}$  is a Poisson point measure with intensity  $\nu_k(dz) ds$ .

The Poisson point measure  $N_{\nu_k}$  can be represented by means of compound Poisson processes as follows. For each  $k \in \mathbb{N}$ , we denote by  $T_i^k, i \in \mathbb{N}$  the jump times of a Poisson process  $(J_t^k)_{t \in [0, T]}$  of parameter  $m_k$ , and we consider a sequence of independent random variables  $Z_i^k \sim \mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k}, i \in \mathbb{N}$ , which are

independent of  $J^k$  as well. Then, for any  $t > 0$  and  $A \in \mathcal{B}([k, k+1))$ ,  $N_{\nu_k}([0, t] \times A) = \sum_{i=1}^{J_t^k} \mathbb{1}_A(Z_i^k)$ . And for each  $k, i \in \mathbb{N}$ , we will split  $Z_i^k$  as  $Z_i^k = \xi_i^k V_i^k + (1 - \xi_i^k) U_i^k$ .

*Remark.* The law of  $Z_i^k$  could be very irregular and it is not possible to make integration by parts based on it. So we construct the  $V_i^k$ , which has all the good regularity properties in order to make Malliavin calculus. This is the idea of the splitting method. The splitting method presented here is analogous to the one in [10]. Therein, Bichteler, Gravereau and Jacod deal with 2 kinds of independent Poisson point measures. One is very regular, and smooth enough to make Malliavin calculus on it (in our paper,  $V_i^k$  play the same role). The other one can be arbitrary, and it may be very irregular (in our paper, it corresponds to  $U_i^k$ ). But the difference is that instead of splitting the Poisson point measure, we split the random variables, and so this method can also be applied in a large class of different problems. For example, Bally, Caramellino and Poly use the splitting method to show the convergence in total variation distance in the central limit theorem in [5]. Other possible approaches to the Malliavin calculus for jump processes are given in the papers [24], [25], [35], [37] and the book [19] for example.

## 4.2 Malliavin calculus for Poisson point measures and space-time Brownian motions

In this section we present the IbP framework on a space where we have the Poisson point measure  $N_\nu$  presented in the previous section and moreover we have a space-time Brownian motion  $W_\nu(ds, dz)$  with covariance measure  $\nu(dz)ds$ , which is independent of  $N_\nu$ . We recall that in Section 2.2 we have introduced the random variables  $W_\nu(\varphi)$ ,  $N_\nu(\varphi)$  and the filtrations  $(\mathcal{F}_t^W)_{t \in [0, T]}$ ,  $(\mathcal{F}_t^N)_{t \in [0, T]}$ , and we denote  $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^N$ . We present now the Malliavin calculus. We recall the random variables  $T_i^k$ , and  $Z_i^k = \xi_i^k V_i^k + (1 - \xi_i^k)U_i^k$  introduced in the previous section and we take  $\mathcal{G} = \sigma(U_i^k, \xi_i^k, T_i^k : k, i \in \mathbb{N})$  to be the  $\sigma$ -algebra associated to the noise from  $U_i^k, \xi_i^k, T_i^k$ ,  $k, i \in \mathbb{N}$ . These are the noises which will not be involved in the Malliavin calculus. We denote by  $C_{\mathcal{G}, p}$  the space of the functions  $f : \Omega \times \mathbb{R}^{m' \times m} \times \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $f$  is  $\mathcal{F}_T$ -measurable, and for each  $\omega$ , the function  $(v_1^1, \dots, v_{m'}^m, w_1, \dots, w_n) \mapsto f(\omega, v_1^1, \dots, v_{m'}^m, w_1, \dots, w_n)$  belongs to  $C_p^\infty(\mathbb{R}^{m' \times m} \times \mathbb{R}^n)$ , and for each  $(v_1^1, \dots, v_{m'}^m, w_1, \dots, w_n)$ , the function  $\omega \mapsto f(\omega, v_1^1, \dots, v_{m'}^m, w_1, \dots, w_n)$  is  $\mathcal{G}$ -measurable. Then we define the space of simple functionals

$$S = \{F = f(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) : f \in C_{\mathcal{G}, p}, \varphi_1, \dots, \varphi_n \in L^2([1, \infty) \times [0, T], \nu \times Leb), m', m, n \in \mathbb{N}\}.$$

On the space  $S$  we define the derivative operators

$$D_{(k_0, i_0)}^Z F = \mathbb{1}_{\{k_0 \leq m\}} \mathbb{1}_{\{i_0 \leq m'\}} \xi_{i_0}^{k_0} \frac{\partial f}{\partial v_{i_0}^{k_0}}(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n), \quad k_0, i_0 \in \mathbb{N}$$

$$D_{(s, z)}^W F = \sum_{r=1}^n \frac{\partial f}{\partial w_r}(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) \varphi_r(s, z), \quad (s, z) \in [0, T] \times [1, \infty).$$

We regard  $D^Z F$  as an element of the Hilbert space  $l_2$  (the space of the sequences  $h = (h_i^k)_{k, i \in \mathbb{N}}$  with  $|h|_{l_2}^2 = \sum_{k=1}^\infty \sum_{i=1}^\infty |h_i^k|^2 < \infty$ ) and  $D^W F$  as an element of the Hilbert space  $L^2([1, \infty) \times [0, T], \nu \times Leb)$ . Then

$$DF := (D^Z F, D^W F) \in l_2 \otimes L^2([1, \infty) \times [0, T], \nu \times Leb).$$

We also denote  $D^{Z, W} F = DF$  and  $\mathcal{H} = l_2 \otimes L^2([1, \infty) \times [0, T], \nu \times Leb)$ . And we have

$$\langle DF, DG \rangle_{\mathcal{H}} = \sum_{k=1}^\infty \sum_{i=1}^\infty D_{(k, i)}^Z F \times D_{(k, i)}^Z G + \int_{[0, T] \times [1, \infty)} D_{(s, z)}^W F \times D_{(s, z)}^W G \nu(dz)ds.$$

Moreover, we define the derivatives of order  $q \in \mathbb{N}$  recursively:

$$D_{(k_1, i_1) \dots (k_q, i_q), (s_1, z_1) \dots (s_q, z_q)}^{Z, W, q} F := D_{(k_q, i_q), (s_q, z_q)}^{Z, W} D_{(k_{q-1}, i_{q-1}), (s_{q-1}, z_{q-1})}^{Z, W} \dots D_{(k_1, i_1), (s_1, z_1)}^{Z, W} F,$$

and we denote  $D^q F = D^{Z, W, q} F$ . We also denote  $D^{Z, q}$  (respectively  $D^{W, q}$ ) as the derivative  $D^Z$  (respectively  $D^W$ ) of order  $q$ .

We recall the function  $\theta_k$  defined in (44) and we define the Ornstein-Uhlenbeck operators  $L^Z$ ,  $L^W$  and  $L = L^Z + L^W$  (which verify the duality relation), with

$$L^Z F = - \sum_{k=1}^m \sum_{i=1}^{m'} (D_{(k, i)}^Z D_{(k, i)}^Z F + \xi_i^k D_{(k, i)}^Z F \times \theta_k(V_i^k)),$$

$$L^W F = \sum_{r=1}^n \frac{\partial f}{\partial w_r}(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) W_\nu(\varphi_r) \\ - \sum_{l, r=1}^n \frac{\partial^2 f}{\partial w_l \partial w_r}(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) \langle \varphi_l, \varphi_r \rangle_{L^2([1, \infty) \times [0, T], \nu \times Leb)}.$$

One can check that the triplet  $(S, D, L)$  is consistent with the IbP framework given in Section 3. The proof is left to Appendix 5.3.

In the following, we will close the operator  $D^q$  and  $L$ , so we will use the IbP framework  $(\mathcal{D}_\infty, D, L)$  associated to  $(S, D, L)$  in **Lemma 3.1**.

### 4.3 Malliavin calculus applied to stochastic differential equations with jumps

Now we will use the IbP framework presented in Section 4.2 for the equation (10).

Let  $\theta : (0, 1] \rightarrow [1, \infty)$  be a function such that  $\theta(z) = \frac{1}{z}$ . By a change of variables, instead of dealing with equation (10), it is equivalent to consider the following equation.

$$\begin{aligned}\hat{X}_t^M &= x + \int_0^t \int_{[1, M)} \tilde{c}(s, z, \hat{X}_{s-}^M) N_\nu(ds, dz) \\ &\quad + \int_0^t b_M(s, \hat{X}_s^M) ds + \int_0^t \int_{\{z \geq M\}} \tilde{c}(s, z, \hat{X}_s^M) W_\nu(ds, dz),\end{aligned}\tag{49}$$

where  $M = \frac{1}{\varepsilon}$ ,  $\nu(dz) = \mu \circ \theta^{-1}(dz)$ ,  $\tilde{c}(s, z, x) = c(s, \frac{1}{z}, x)$ ,

$$b_M(s, x) = \int_{\{z \geq M\}} \tilde{c}(s, z, x) \nu(dz),\tag{50}$$

and  $W_\nu$  is the space-time Brownian motion with covariance measure  $\nu(dz)ds$ . One can check that  $\hat{X}_t^M$  has the same law as  $X_t^\varepsilon$ .

Here we give two lemmas, concerning the Malliavin-Sobolev norms and the Malliavin covariance. We recall that  $\varepsilon_*$  and  $\alpha$  are introduced in **Hypothesis 2.4** (b), and  $q^*$  is introduced in **Hypothesis 2.1**.

**Lemma 4.1.** *Assuming **Hypothesis 2.1** with  $q^* \geq 2$  and **Hypothesis 2.4** (either 2.4(a) or 2.4(b)), we have  $\hat{X}_t^M \in \mathcal{D}_\infty$ , and for all  $p \geq 1$ ,  $2 \leq l \leq q^*$ , there exists a constant  $C_{l,p}(T)$  depending on  $l, p, x$  and  $T$ , such that  $\sup_M \|\hat{X}_t^M\|_{L,l,p} \leq C_{l,p}(T)$ .*

**Lemma 4.2.** *Assume that **Hypothesis 2.1** with  $q^* \geq 1$  and **Hypothesis 2.2, 2.3** hold true.*

a) *If we also assume **Hypothesis 2.4** (a), then for every  $p \geq 1$ ,  $t \in [0, T]$ , we have*

$$\sup_M \mathbb{E}(1/\sigma_{\hat{X}_t^M})^p \leq C_p,\tag{51}$$

with  $C_p$  a constant only depending on  $p$  and  $T$ .

b) *If we assume **Hypothesis 2.4** (b), then for every  $p \geq 1$ ,  $t \in [0, T]$  such that  $t > \frac{4p\alpha}{\varepsilon_*}$ , we have  $\sup_M \mathbb{E}(1/\sigma_{\hat{X}_t^M})^p \leq C_p$ .*

The proofs of these lemmas are rather technical and are postponed for the Appendix (Section 5.1 and 5.2).

### 4.4 Proof of the main result (Theorem 2.2)

*Proof.* (a) By **Lemma 4.1** and **Lemma 4.2** a), we know that for any  $\delta > 0$ , for any  $p, p_1 \geq 1$ ,  $2 \leq q \leq q^*$ , with  $q^* \geq \frac{3}{\delta} + 1$ , there exists a constant  $C_{q,p,p_1}(T)$  such that for any  $M \geq 1$ ,  $t \in [0, T]$ , we have

$$\Sigma_{p_1}(\hat{X}_t^M) + \|\hat{X}_t^M\|_{L,q,p} \leq C_{q,p,p_1}(T).$$

By **Lemma 2.1**, we know that

$$d_3(\hat{X}_t, \hat{X}_t^M) = d_3(X_t, X_t^\varepsilon) \leq C\eta_3(\varepsilon).$$

Then applying **Corollary 3.4.1** i) for  $r = 3$ , we have

$$d_{TV}(X_t, X_t^\varepsilon) = d_{TV}(\hat{X}_t, \hat{X}_t^M) \leq C d_3(\hat{X}_t, \hat{X}_t^M)^{1-\delta} \leq C_\delta \eta_3(\varepsilon)^{1-\delta}.$$

So we obtain (23). The proof of (24) is obtained by **Corollary 3.4.1 ii)**, since  $q^* \geq \frac{3+l}{\delta} + 1$ .

(b) The proof is almost the same. If  $t > \frac{8\alpha(\frac{3}{\delta}-1)}{\varepsilon_*}$ , then by **Lemma 4.2 b)**,  $\Sigma_{p_1}(\hat{X}_t^M) < \infty$  for  $p_1 = 2(\frac{3}{\delta} - 1)$ . So **Corollary 3.4.1 i)** still holds, and we can obtain (25). For

$$t > \max\left\{\frac{8\alpha}{\varepsilon_*}\left(\frac{3+l}{\delta} - 1\right), \frac{8\alpha(3l+2)}{\varepsilon_*}\right\},$$

by **Lemma 4.2 b)**,  $\Sigma_{p_2}(\hat{X}_t^M) < \infty$  for  $p_2 = \max\{2(\frac{3+l}{\delta} - 1), 6l+4\}$ . So **Corollary 3.4.1 ii)** still holds, and we obtain (26).  $\square$

## 5 Appendix

### 5.1 Proof of Lemma 4.1

In the following, we will only work with the measure  $\nu$  supported on  $[1, \infty)$  and with the processes  $(\hat{X}_t)_{t \in [0, T]}$  and  $(\hat{X}_t^M)_{t \in [0, T]}$ . So in order to simplify the notation, from now on we denote  $\hat{X}_t = X_t$  and  $\hat{X}_t^M = X_t^M$ . We remark that  $M = \frac{1}{\varepsilon}$  is generally not an integer, but for simplicity, we assume in the following that  $M$  is an integer.

Here is the idea of the proof. Since  $X_t^M$  is not a simple functional, we construct first the Euler scheme  $(X_t^{n, M})_{t \in [0, T]}$  in subsection 5.1.1 and check that  $X_t^{n, M} \rightarrow X_t^M$  in  $L^1$  when  $n \rightarrow \infty$ . We will prove that  $\mathbb{E}|X_t^{n, M}|_l^p$  and  $\mathbb{E}|LX_t^{n, M}|_l^p$  are bounded (uniformly in  $n, M$ ) in subsection 5.1.3. Then based on **Lemma 3.2**, we obtain that  $X_t^M \in \mathcal{D}_\infty$  and the norms  $\|X_t^M\|_{L, l, p}$  are bounded (uniformly in  $M$ ).

#### 5.1.1 Construction of the Euler scheme

We take a time-partition  $\mathcal{P}_t^n = \{r_j = \frac{jt}{n}, j = 0, \dots, n\}$  and a space-partition  $\tilde{\mathcal{P}}_M^n = \{z_j = M + \frac{j}{n}, j = 0, 1, \dots\}$ . We denote  $\tau_n(r) = r_j$  when  $r \in [r_j, r_{j+1})$ , and denote  $\gamma_n(z) = z_j$  when  $z \in [z_j, z_{j+1})$ . Let

$$\begin{aligned} X_t^{n, M} &= x + \int_0^t \int_{[1, M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n, M}) N_\nu(dr, dz) \\ &\quad + \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n, M}) dr + \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n, M}) W_\nu(dr, dz). \end{aligned} \quad (52)$$

Then we can obtain the following lemma.

**Lemma 5.1.** *Assume that the **Hypothesis 2.1** holds true with  $q^* \geq 1$ . Then for any  $p \geq 1, M \geq 1$ , we have  $\mathbb{E}|X_t^{n, M} - X_t^M|^p \rightarrow 0$  as  $n \rightarrow \infty$ .*

*Proof.* We first notice that since  $\bar{c}(z)$  (in **Hypothesis 2.1**) is decreasing,  $\sup_{n \in \mathbb{N}} \bar{c}(\gamma_n(z)) \leq \bar{c}(\gamma_1(z))$ . So

$$\int_1^\infty \sup_{n \in \mathbb{N}} |\bar{c}(\gamma_n(z))|^2 \nu(dz) \leq \int_1^\infty |\bar{c}(\gamma_1(z))|^2 \nu(dz) \leq |\bar{c}(1)|^2 \nu[1, 2] + \int_1^\infty |\bar{c}(z)|^2 \nu(dz) \leq C < \infty. \quad (53)$$

Then by the Lebesgue dominated convergence theorem, (53) implies that

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} \int_0^T \int_{[1, \infty)} |\tilde{c}(s, z, x) - \tilde{c}(\tau_n(s), \gamma_n(z), x)|^2 \nu(dz) ds = 0, \quad (54)$$

and

$$\sup_{n \in \mathbb{N}} \sup_{x \in \mathbb{R}} \int_0^T \int_{[1, \infty)} |\tilde{c}(\tau_n(s), \gamma_n(z), x)|^2 \nu(dz) ds \leq C. \quad (55)$$



In the following proof,  $C_p(T)$  will be a constant depending on  $p$  and  $T$  which may be changed from line to line. For  $p \geq 2$ , we write  $\mathbb{E}|X_t^{n,M} - X_t^M|^p \leq C_p(T)(E_1 + E_2 + E_3)$ , where

$$E_1 = \mathbb{E} \left| \int_0^t \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M) N_\nu(dr, dz) \right|^p,$$

$$E_2 = \mathbb{E} \left| \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) - b_M(r, X_r^M) dr \right|^p,$$

$$E_3 = \mathbb{E} \left| \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) - \tilde{c}(r, z, X_r^M) W_\nu(dr, dz) \right|^p.$$

Then, compensating  $N_\nu$  and using Burkholder's inequality (see for example the Theorem 2.11 in [26]),

$$\begin{aligned} E_1 &\leq C_p(T) \mathbb{E} \int_0^t \left( \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M)|^2 \nu(dz) \right)^{\frac{p}{2}} dr \\ &\quad + \mathbb{E} \int_0^t \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M)|^p \nu(dz) dr \\ &\quad + \mathbb{E} \int_0^t \left| \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M) \nu(dz) \right|^p dr \\ &\leq C_p(T) [R_n^1 + ((\bar{c}_2)^{\frac{p}{2}} + \bar{c}_p + (\bar{c}_1)^p) \int_0^t \mathbb{E}|X_{\tau_n(r)}^{n,M} - X_r^M|^p dr], \end{aligned}$$

with

$$\begin{aligned} R_n^1 &= \mathbb{E} \int_0^t \left( \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M)|^2 \nu(dz) \right)^{\frac{p}{2}} dr \\ &\quad + \mathbb{E} \int_0^t \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M)|^p \nu(dz) dr \\ &\quad + \mathbb{E} \int_0^t \left| \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M) \nu(dz) \right|^p dr. \end{aligned}$$

Since  $|\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) - \tilde{c}(r, z, X_{r-}^M)|^p \leq |2\bar{c}(z)|^p \in L^1(\Omega \times [1, \infty) \times [0, T], \mathbb{P} \times \nu \times Leb)$ , we apply the Lebesgue's dominated convergence theorem and we obtain that  $R_n^1 \rightarrow 0$ . Next,

$$\begin{aligned} E_2 &\leq C_p(T) \mathbb{E} \int_0^t \left| \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)}^{n,M}) - \tilde{c}(r, z, X_r^M) \nu(dz) \right|^p dr \\ &\leq C_p(T) [R_n^2 + (\bar{c}_1)^p \int_0^t \mathbb{E}|X_{\tau_n(r)}^{n,M} - X_r^M|^p dr], \end{aligned}$$

with

$$R_n^2 = \mathbb{E} \int_0^t \left| \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)}^{n,M}) - \tilde{c}(r, z, X_r^M) \nu(dz) \right|^p dr \rightarrow 0.$$

Finally, using Burkholder's inequality,

$$\begin{aligned} E_3 &\leq C_p(T) \mathbb{E} \int_0^t \int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) - \tilde{c}(r, z, X_r^M)|^2 \nu(dz) \right)^{\frac{p}{2}} dr \\ &\leq C_p(T) [R_n^3 + (\bar{c}_2)^{\frac{p}{2}} \int_0^t \mathbb{E}|X_{\tau_n(r)}^{n,M} - X_r^M|^p dr], \end{aligned}$$

where (by (54)),

$$R_n^3 = \mathbb{E} \left| \int_0^t \int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_r^M) - \tilde{c}(r, z, X_r^M)|^2 \nu(dz) \right|^{\frac{p}{2}} dr \rightarrow 0.$$

Therefore,  $\mathbb{E}|X_t^{n,M} - X_t^M|^p \leq C_p(T)[R_n + \int_0^t \mathbb{E}|X_{\tau_n(r)}^{n,M} - X_r^M|^p dr]$ , with  $R_n = R_n^1 + R_n^2 + R_n^3 \rightarrow 0$  as  $n \rightarrow \infty$ . One can easily check that  $\mathbb{E}|X_t^{n,M} - X_{\tau_n(t)}^{n,M}|^p \rightarrow 0$ . Also there exists a constant  $C_p(T)$  depending on  $p$  and  $T$  such that for any  $n, M \in \mathbb{N}$  and any  $t \in [0, T]$ ,  $\mathbb{E}|X_t^{n,M}|^p \leq C_p(T)$  (see (71) for details). Then, by the dominated convergence theorem, these yield  $\int_0^t \mathbb{E}|X_r^{n,M} - X_{\tau_n(r)}^{n,M}|^p dr \rightarrow 0$ . So we have  $\mathbb{E}|X_t^{n,M} - X_t^M|^p \leq C_p(T)[\widetilde{R}_n + \int_0^t \mathbb{E}|X_r^{n,M} - X_r^M|^p dr]$ , with  $\widetilde{R}_n \rightarrow 0$  as  $n \rightarrow \infty$ . We conclude by using Gronwall's lemma.  $\square$

*Remark.* Some results on the convergence of the Euler scheme of a jump-diffusion can be found for example in [32], [34]. The special thing in our paper is that we deal with the space-time Brownian motion instead of the classical Brownian motion, and this is why we need to assume (54).

Now we represent the jump's part of  $(X_t^{n,M})_{t \in [0, T]}$  by means of compound Poisson processes. We recall that for each  $k \in \mathbb{N}$ , we denote by  $T_i^k, i \in \mathbb{N}$  the jump times of a Poisson process  $(J_t^k)_{t \in [0, T]}$  of parameter  $m_k$ , and we consider a sequence of independent random variables  $Z_i^k \sim \mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k}, i \in \mathbb{N}$ , which are independent of  $J^k$  as well. Then we write

$$\begin{aligned} X_t^{n,M} &= x + \int_0^t \sum_{k=1}^{M-1} \int_{\{z \in I_k\}} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) N_\nu(dr, dz) \\ &\quad + \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) dr + \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) W_\nu(dr, dz) \\ &= x + \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M}) \\ &\quad + \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) dr + \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) W_\nu(dr, dz). \end{aligned} \quad (56)$$

So for every  $t \in [0, T]$ ,  $X_t^{n,M}$  is a simple functional.

### 5.1.2 Preliminary estimates

In order to estimate the Sobolev norms of the Euler scheme, we need the following preliminary lemmas.

**Lemma 5.2.** *We fix  $M \geq 1$ . Let  $y : \Omega \times [0, T] \times [M, \infty) \rightarrow \mathbb{R}$  be a function which is piecewise constant with respect to both  $t$  and  $z$ . We assume that  $y_t(z)$  is progressively measurable with respect to  $\mathcal{F}_t$  (defined in (11)),  $y_t(z) \in \mathcal{S}$ , and  $\mathbb{E}(\int_0^t \int_{\{z \geq M\}} |y_r(z)|^2 \nu(dz) dr) < \infty$ . We denote  $I_t(y) = \int_0^t \int_{\{z \geq M\}} y_r(z) W_\nu(dr, dz)$ . Then for any  $l \geq 1, p \geq 2$ , there exists a constant  $C_{l,p}(T)$  such that*

$$a) \quad \mathbb{E}|I_t(y)|_l^p \leq C_{l,p}(T) \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |y_r(z)|_l^2 \nu(dz) \right)^{\frac{p}{2}} dr,$$

$$b) \quad \mathbb{E}|LI_t(y)|_l^p \leq C_{l,p}(T) [\mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |Ly_r(z)|_l^2 \nu(dz) \right)^{\frac{p}{2}} dr + \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |y_r(z)|_l^2 \nu(dz) \right)^{\frac{p}{2}} dr].$$

*Proof.* Proof of a): Let  $C_{l,p}(T)$  be a constant depending on  $l, p$  and  $T$  which may change from one line to another. For any  $l \geq 1$ , we take  $l_W \geq 0$  and  $l_Z \geq 0$  such that  $0 < l_W + l_Z \leq l$ .

It is easy to check that

$$D_{(k_1, i_1) \dots (k_{l_Z}, i_{l_Z})}^{Z, l_Z} I_t(y) = \int_0^t \int_{\{z \geq M\}} D_{(k_1, i_1) \dots (k_{l_Z}, i_{l_Z})}^{Z, l_Z} y_r(z) W_\nu(dr, dz).$$

And by recurrence, one can show that

$$D_{(s_1, z_1) \dots (s_{l_W}, z_{l_W})}^{W, l_W} I_t(y) = \int_0^t \int_{\{z \geq M\}} D_{(s_1, z_1) \dots (s_{l_W}, z_{l_W})}^{W, l_W} y_r(z) W_\nu(dr, dz) + \sum_{j=1}^{l_W} D_{(s_j, z_j)}^{W, l_W-1} y_{s_j}(z_j) \mathbb{1}_{s_j \leq t},$$

with

$$\widehat{(s_j, z_j)}^{l_W-1} := (s_1, z_1) \cdots (s_{j-1}, z_{j-1}) (s_{j+1}, z_{j+1}) \cdots (s_{l_W}, z_{l_W}).$$

We denote

$$\bar{y}_r(z)(k_1, i_1, \dots, k_{l_Z}, i_{l_Z}) := D_{(k_1, i_1) \dots (k_{l_Z}, i_{l_Z})}^{Z, l_Z} y_r(z), \quad \bar{y}_r^{l_Z}(z) := D^{Z, l_Z} y_r(z) \in l_2^{\otimes l_Z}.$$

Then  $D^{Z, l_Z} I_t(y) = I_t(\bar{y}^{l_Z})$ , and

$$\begin{aligned} D_{(s_1, z_1) \dots (s_{l_W}, z_{l_W})}^{W, l_W} D_{(k_1, i_1) \dots (k_{l_Z}, i_{l_Z})}^{Z, l_Z} I_t(y) &= \int_0^t \int_{\{z \geq M\}} D_{(s_1, z_1) \dots (s_{l_W}, z_{l_W})}^{W, l_W} \bar{y}_r(z)(k_1, i_1, \dots, k_{l_Z}, i_{l_Z}) W_\nu(dr, dz) \\ &+ \sum_{j=1}^{l_W} D_{(s_j, z_j)}^{W, l_W-1} \bar{y}_{s_j}(z_j)(k_1, i_1, \dots, k_{l_Z}, i_{l_Z}) \mathbb{1}_{s_j \leq t}. \end{aligned}$$

Let  $H_{l_Z, l_W, T} = l_2^{\otimes l_Z} \otimes L^2([0, T] \times [M, \infty), Leb \times \nu)^{\otimes l_W}$ . We have

$$\begin{aligned} |D^{W, l_W} D^{Z, l_Z} I_t(y)|_{H_{l_Z, l_W, T}}^2 &= \int_{[0, T]^{l_W}} \int_{[M, \infty)^{l_W}} |D_{(s_1, z_1) \dots (s_{l_W}, z_{l_W})}^{W, l_W} I_t(\bar{y}^{l_Z})|_{l_2^{\otimes l_Z}}^2 \nu(dz_1) ds_1 \cdots \nu(dz_{l_W}) ds_{l_W} \\ &\leq 2 \left| \int_0^t \int_{\{z \geq M\}} D^{W, l_W} \bar{y}_r^{l_Z}(z) W_\nu(dr, dz) \right|_{H_{l_Z, l_W, T}}^2 + l_W 2^{l_W} \int_0^t \int_{\{z \geq M\}} |D^{W, l_W-1} \bar{y}_r^{l_Z}(z)|_{H_{l_Z, l_W-1, T}}^2 \nu(dz) dr. \end{aligned}$$

Using Burkholder's inequality for Hilbert-space-valued martingales (see [30] for example), we have

$$\begin{aligned} \mathbb{E} |D^{W, l_W} D^{Z, l_Z} I_t(y)|_{H_{l_Z, l_W, T}}^p &\leq C_{l, p}(T) \left[ \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |D^{W, l_W} D^{Z, l_Z} y_r(z)|_{H_{l_Z, l_W, T}}^2 \nu(dz) \right)^{\frac{p}{2}} dr \right. \\ &\quad \left. + \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |D^{W, l_W-1} D^{Z, l_Z} y_r(z)|_{H_{l_Z, l_W-1, T}}^2 \nu(dz) \right)^{\frac{p}{2}} dr \right]. \end{aligned} \quad (57)$$

We recall that for  $F \in \mathcal{D}_\infty$ , we have  $|D^{W, l_W} D^{Z, l_Z} F|_{H_{l_Z, l_W, T}} \leq |F|_{l_Z + l_W}$  (see the definition in (33)). Then (57) gives

$$\mathbb{E} |I_t(y)|_{1, l}^p \leq C_{l, p}(T) \sum_{l_Z + l_W \leq l} \mathbb{E} |D^{W, l_W} D^{Z, l_Z} I_t(y)|_{H_{l_Z, l_W, T}}^p \leq C_{l, p}(T) \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |y_r(z)|_l^2 \nu(dz) \right)^{\frac{p}{2}} dr. \quad (58)$$

Finally, using Burkholder's inequality, we have

$$\mathbb{E} |I_t(y)|^p \leq C_{l, p}(T) \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |y_r(z)|^2 \nu(dz) \right)^{\frac{p}{2}} dr. \quad (59)$$

So a) is proved.

Proof of b): We first show that

$$L I_t(y) = I_t(Ly) + I_t(y). \quad (60)$$

We denote

$$I_k^t(f_k) = k! \int_0^t \int_0^{s_1} \cdots \int_0^{s_{k-1}} \int_{[M, +\infty)^k} f_k(s_1, \dots, s_k, z_1, \dots, z_k) W_\nu(ds_k, dz_k) \cdots W_\nu(ds_1, dz_1)$$

the multiple stochastic integral for a deterministic function  $f_k$ , which is square integrable with respect to  $(\nu(dz)ds)^{\otimes k}$  and is symmetric with respect to the time variation  $(s_1, \dots, s_k)$  for each fixed  $(z_1, \dots, z_k)$ . Notice that  $L^Z I_k^t(f_k) = 0$  and  $L^W I_k^t(f_k) = k I_k^t(f_k)$ . So,  $L I_k^t(f_k) = k I_k^t(f_k)$ . Then by the duality relation (32),

$$\mathbb{E}(I_k^t(f_k) L(I_t(y))) = \mathbb{E}(I_t(y) \times L I_k^t(f_k)) = k \mathbb{E}(I_t(y) \times I_k^t(f_k)). \quad (61)$$

On the other hand, using the isometry property and the duality relation,

$$\begin{aligned} \mathbb{E}(I_k^t(f_k) \times I_t(Ly)) &= k \mathbb{E} \int_0^t \int_{\{z \geq M\}} I_{k-1}^r(f_k(r, z, \cdot)) L y_r(z) \nu(dz) dr \\ &= k \int_0^t \int_{\{z \geq M\}} \mathbb{E}[y_r(z) \times L I_{k-1}^r(f_k(r, z, \cdot))] \nu(dz) dr = k(k-1) \mathbb{E} \int_0^t \int_{\{z \geq M\}} y_r(z) I_{k-1}^r(f_k(r, z, \cdot)) \nu(dz) dr \\ &= k(k-1) \mathbb{E}(I_t(y) \times \int_0^t \int_{\{z \geq M\}} I_{k-1}^r(f_k(r, z, \cdot)) W_\nu(dr, dz)) \\ &= (k-1) \mathbb{E}(I_t(y) \times I_k^t(f_k)). \end{aligned}$$

Combining this with (61), we get

$$\mathbb{E}(I_k^t(f_k)(I_t(y) + I_t(Ly))) = k \mathbb{E}(I_k^t(f_k) I_t(y)) = \mathbb{E}[I_k^t(f_k) \times L I_t(y)]. \quad (62)$$

Since every element in  $L^2(W)$  (defined by (12)) can be represented as the direct sum of multiple stochastic integrals, we have for any  $F \in L^2(W)$ ,

$$\mathbb{E}[F L I_t(y)] = \mathbb{E}[F(I_t(Ly) + I_t(y))]. \quad (63)$$

For  $G \in L^2(N)$ , one has  $L^W G = 0$  and  $L^Z G \in L^2(N)$ . Then by using duality and (13),

$$\mathbb{E}[G L I_t(y)] = \mathbb{E}[I_t(y) L G] = \mathbb{E}[I_t(y) L^Z G] = 0,$$

and by (13),

$$\mathbb{E}[G(I_t(Ly) + I_t(y))] = 0.$$

So,

$$\mathbb{E}[G L I_t(y)] = \mathbb{E}[G(I_t(Ly) + I_t(y))]. \quad (64)$$

Combining (63) and (64), for any  $\tilde{G} \in L^2(W) \otimes L^2(N)$ , we have  $\mathbb{E}[\tilde{G} L I_t(y)] = \mathbb{E}[\tilde{G}(I_t(Ly) + I_t(y))]$ , which proves (60).

Then, by **Lemma 5.2 a)**,

$$\begin{aligned} \mathbb{E}|L I_t(y)|_l^p &\leq 2^{p-1} (\mathbb{E} \int_0^t \int_{\{z \geq M\}} L y_r(z) W_\nu(dr, dz)|_l^p + \mathbb{E} \int_0^t \int_{\{z \geq M\}} y_r(z) W_\nu(dr, dz)|_l^p) \\ &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |L y_r(z)|_l^2 \nu(dz))^{\frac{p}{2}} dr + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} |y_r(z)|_l^2 \nu(dz))^{\frac{p}{2}} dr]. \end{aligned}$$

□

We will also need the following lemma from [7] (Lemma 8 and Lemma 10), which is a consequence of the chain rule for  $D^q$  and  $L$ .

**Lemma 5.3.** *Let  $F \in S^d$ . For every  $l \in \mathbb{N}$ , if  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $C^l(\mathbb{R}^d)$  function ( $l$ -times differentiable function), then*

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

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

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

For  $l = 0$ , we have

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

We finish this section with a first estimate concerning the operator  $L$ .

**Lemma 5.4.** *Under the **Hypothesis 2.4** (either 2.4(a) or 2.4(b)), for every  $p \geq 2, \tilde{p} \geq 1, l \geq 0$ , there exists a constant  $C_{l,p,\tilde{p}}(T)$  such that*

$$\sup_{M \in \mathbb{N}} \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LZ_i^k|_l^{\tilde{p}} \right)^p \leq C_{l,p,\tilde{p}}(T). \quad (65)$$

*Proof.* We notice that (with  $\psi_k$  given in (44)),  $LZ_i^k = \xi_i^k (\ln \psi_k)'(V_i^k)$  and  $D^{W,l} LZ_i^k = 0$ . Moreover,

$$D_{(r_1, m_1) \dots (r_l, m_l)}^{Z,l} LZ_i^k = \prod_{j=1}^l (\delta_{r_j k} \delta_{m_j i}) \xi_i^k (\ln \psi_k)^{(l+1)}(V_i^k),$$

with  $\delta_{r,k}$  the Kronecker delta, so that

$$|LZ_i^k|_l = \xi_i^k \sum_{0 \leq \tilde{l} \leq l} |(\ln \psi_k)^{(\tilde{l}+1)}(V_i^k)|. \quad (66)$$

It follows that

$$\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LZ_i^k|_l^{\tilde{p}} \right)^p \leq C_{l,p,\tilde{p}} \sum_{0 \leq \tilde{l} \leq l} \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) \xi_i^k |(\ln \psi_k)^{(\tilde{l}+1)}(V_i^k)|^{\tilde{p}} \right)^p.$$

Since  $\bar{c}(Z_i^k) \xi_i^k = \bar{c}(V_i^k) \xi_i^k$ , we may replace  $Z_i^k$  by  $V_i^k$  in the right hand side of the above estimate. This gives

$$C_{l,p,\tilde{p}} \sum_{0 \leq \tilde{l} \leq l} \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(V_i^k) \xi_i^k |(\ln \psi_k)^{(\tilde{l}+1)}(V_i^k)|^{\tilde{p}} \right)^p = C_{l,p,\tilde{p}} \sum_{0 \leq \tilde{l} \leq l} \mathbb{E} \left| \int_0^t \int_{[1,M)} \int_{\{0,1\}} \bar{c}(v) \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{\tilde{p}} \Lambda(ds, d\xi, dv) \right|^p,$$

where  $\bar{\psi}(v) := \sum_{k=1}^{\infty} \mathbb{1}_{I_k}(v) \psi(v - (k + \frac{1}{2}))$  and  $\Lambda$  is a Poisson point measure on  $\{0, 1\} \times [1, \infty)$  with compensator

$$\widehat{\Lambda}(ds, d\xi, dv) = \sum_{k=1}^{\infty} \left[ \frac{\psi(v - (k + \frac{1}{2}))}{m(\psi)} \mathbb{1}_{I_k}(v) dv \times b(v, d\xi) \right] ds,$$

with  $b(v, d\xi)$  the Bernoulli probability measure on  $\{0, 1\}$  with parameter  $\varepsilon_k m(\psi)$ , if  $v \in I_k$ . Then by compensating  $\Lambda$  and using Burkholder's inequality (the same proof as for (5)),

$$\begin{aligned} & C_{l,p,\tilde{p}} \sum_{0 \leq \tilde{l} \leq l} \mathbb{E} \left| \int_0^t \int_{[1,M)} \int_{\{0,1\}} \bar{c}(v) \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{\tilde{p}} \Lambda(ds, d\xi, dv) \right|^p \\ & \leq C_{l,p,\tilde{p}}(T) \sum_{0 \leq \tilde{l} \leq l} \left[ \left( \int_0^t \int_{[1,M) \times \{0,1\}} |\bar{c}(v)|^2 \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{2\tilde{p}} \widehat{\Lambda}(ds, d\xi, dv) \right)^{\frac{p}{2}} \right. \\ & \quad \left. + \int_0^t \int_{[1,M) \times \{0,1\}} |\bar{c}(v)|^p \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{p\tilde{p}} \widehat{\Lambda}(ds, d\xi, dv) + \int_0^t \int_{[1,M) \times \{0,1\}} \bar{c}(v) \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{\tilde{p}} \widehat{\Lambda}(ds, d\xi, dv) \right]^p. \end{aligned} \quad (67)$$

We notice that by (45),

$$\begin{aligned} \int_0^t \int_{[1,M) \times \{0,1\}} |\bar{c}(v)|^p \xi |(\ln \bar{\psi})^{(\tilde{l}+1)}(v)|^{p\tilde{p}} \widehat{\Lambda}(ds, d\xi, dv) &= t \sum_{k=1}^{M-1} \int_{I_k} \varepsilon_k m(\psi) |\bar{c}(v)|^p |(\ln \psi_k)^{(\tilde{l}+1)}(v)|^{p\tilde{p}} \frac{\psi_k(v)}{m(\psi)} dv \\ &\leq C_{\tilde{l},p,\tilde{p}}(T) \sum_{k=1}^{M-1} \varepsilon_k \int_{I_k} |\bar{c}(v)|^p dv. \end{aligned}$$

Similar upper bounds hold for the two other terms in the right hand side of (67), so (67) is upper bounded by

$$C_{l,p,\tilde{p}}(T) \left[ \left( \sum_{k=1}^{M-1} \varepsilon_k \int_{I_k} |\bar{c}(v)|^2 dv \right)^{\frac{p}{2}} + \sum_{k=1}^{M-1} \varepsilon_k \int_{I_k} |\bar{c}(v)|^p dv + \left( \sum_{k=1}^{M-1} \varepsilon_k \int_{I_k} |\bar{c}(v)| dv \right)^p \right]. \quad (68)$$

If we assume the **Hypothesis 2.4** (a), then we have  $\varepsilon_k = \varepsilon_*/(k+1)^{1-\alpha_0}$ , with  $\alpha_0$  given in (7). So the above term is less than

$$C_{l,p,\tilde{p}}(T) \left[ \left( \int_1^\infty \frac{|\bar{c}(v)|^2}{v^{1-\alpha_0}} dv \right)^{\frac{p}{2}} + \int_1^\infty \frac{|\bar{c}(v)|^p}{v^{1-\alpha_0}} dv + \left( \int_1^\infty \frac{|\bar{c}(v)|}{v^{1-\alpha_0}} dv \right)^p \right],$$

which is upper bounded by a constant  $C_{l,p,\tilde{p}}(T)$  thanks to (7). On the other hand, if we assume the **Hypothesis 2.4** (b), then  $\varepsilon_k = \varepsilon_*/(k+1)$ . So (68) is upper bounded by a constant  $C_{l,p,\tilde{p}}(T)$  thanks to (9). □

### 5.1.3 Estimations of $\|X_t^{n,M}\|_{L,l,p}$

In this section, our aim is to prove the following lemma.

**Lemma 5.5.** *Under the **Hypothesis 2.1** with  $q^* \geq 2$  and **Hypothesis 2.4** (either 2.4(a) or 2.4(b)), for all  $p \geq 2, 0 \leq l \leq q^*$ , there exists a constant  $C_{l,p}(T)$  depending on  $l, p, x$  and  $T$ , such that*

$$a) \quad \sup_n \sup_M \mathbb{E} |X_t^{n,M}|_l^p \leq C_{l,p}(T), \quad (69)$$

and for  $0 \leq l \leq q^* - 2$ ,

$$b) \quad \sup_n \sup_M \mathbb{E} |LX_t^{n,M}|_l^p \leq C_{l,p}(T). \quad (70)$$

*Proof.* In the following proof,  $C_{l,p}(T)$  will be a constant which depends on  $l, p, x$  and  $T$ , and which may change from a line to another.  $q^* \geq 2$  is fixed throughout the proof.

a) We prove (69) for  $0 \leq l \leq q^*$  by recurrence on  $l$ .

**Step 1** For  $l = 0$ , using Burkholder's inequality, **Hypothesis 2.1** and (55),

$$\begin{aligned} \mathbb{E} |X_t^{n,M}|^p &\leq C_{0,p}(T) [x^p + \mathbb{E} \left| \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) dr \right|^p + \mathbb{E} \left| \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) W_\nu(dr, dz) \right|^p \\ &\quad + \mathbb{E} \left| \int_0^t \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) N_\nu(dr, dz) \right|^p] \\ &\leq C_{0,p}(T) [1 + \mathbb{E} \int_0^t \left| \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)}^{n,M}) \nu(dz) \right|^p dr \\ &\quad + \mathbb{E} \int_0^t \left( \int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 \nu(dz) \right)^{\frac{p}{2}} dr + \mathbb{E} \int_0^t \left( \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M})|^2 \nu(dz) \right)^{\frac{p}{2}} dr \\ &\quad + \mathbb{E} \int_0^t \int_{[1,M)} |\tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M})|^p \nu(dz) dr + \mathbb{E} \int_0^t \left| \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) \nu(dz) \right|^p dr] \\ &\leq C_{0,p}(T). \end{aligned} \quad (71)$$

**Step 2** Now we assume that (69) holds for  $l-1$ , with  $l \geq 1$  and for every  $p \geq 2$ , and we prove that it holds for  $l$  and for every  $p \geq 2$ . We write  $\mathbb{E}|X_t^{n,M}|_l^p \leq C_{l,p}(T)(A_1 + A_2 + A_3)$ , with

$$\begin{aligned} A_1 &= \mathbb{E} \left| \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) dr \right|_l^p, \\ A_2 &= \mathbb{E} \left| \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) W_\nu(dr, dz) \right|_l^p, \\ A_3 &= \mathbb{E} \left| \int_0^t \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) N_\nu(dr, dz) \right|_l^p. \end{aligned}$$

We notice that by **Hypothesis 2.1**,  $\|b_M\|_{l,\infty} \leq \bar{c}_1$ . Then using **Lemma 5.3 a)** and the recurrence hypothesis, we get

$$\begin{aligned} A_1 &\leq C_{l,p}(T) \mathbb{E} \int_0^t |b_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|_l^p dr \\ &\leq C_{l,p}(T) [(\bar{c}_1)^p + \mathbb{E} \int_0^t |\partial_x b_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|^p |X_{\tau_n(r)}^{n,M}|_{1,l}^p dr \\ &\quad + \mathbb{E} \int_0^t \sup_{2 \leq |\beta| \leq l} |\partial_x^\beta b_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|^p |X_{\tau_n(r)}^{n,M}|_{1,l-1}^{lp} dr] \\ &\leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr]. \end{aligned} \tag{72}$$

Next, we estimate  $A_2$ . By **Hypothesis 2.1**, for every  $n$ ,  $\|\tilde{c}(\tau_n(r), \gamma_n(z), \cdot)\|_{l,\infty} \leq |\bar{c}(\gamma_n(z))|$ . Then using **Lemma 5.2 a)**, **Lemma 5.3 a)**, (53) and the recurrence hypothesis, we get

$$\begin{aligned} A_2 &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|_l^2 \nu(dz))^{\frac{p}{2}} dr \\ &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\partial_x \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 |X_{\tau_n(r)}^{n,M}|_{1,l}^2 \nu(dz))^{\frac{p}{2}} dr \\ &\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} \sup_{2 \leq |\beta| \leq l} |\partial_x^\beta \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 |X_{\tau_n(r)}^{n,M}|_{1,l-1}^{2l} \nu(dz))^{\frac{p}{2}} dr \\ &\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 \nu(dz))^{\frac{p}{2}} dr \\ &\leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr]. \end{aligned} \tag{73}$$

Finally we estimate  $A_3$ . We notice that  $D_{(r,m)}^Z Z_i^k = \xi_i^k \delta_{rk} \delta_{mi}$ ,  $D_{(s,z)}^W Z_i^k = 0$ , and for  $l \geq 2$ ,  $D_{(r_1, m_1) \dots (r_l, m_l), (s_1, z_1) \dots (s_l, z_l)}^{Z,W,l} Z_i^k = 0$ . So we have  $|Z_i^k|_{1,l}^p = |\xi_i^k|^p \leq 1$ . By **Lemma 5.3 a)** for  $d = 2$ , **Hypothesis 2.1**, for any  $k, i \in \mathbb{N}$ ,

$$\begin{aligned} &|\tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|_l \leq |\bar{c}(Z_i^k)| \\ &+ (|\partial_z \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})| + |\partial_x \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|) (|Z_i^k|_{1,l} + |X_{\tau_n(T_i^k)-}^{n,M}|_{1,l}) \\ &+ C_{l,p}(T) \sup_{2 \leq |\beta_1 + \beta_2| \leq l} (|\partial_z^{\beta_2} \partial_x^{\beta_1} \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|) (|Z_i^k|_{1,l-1}^l + |X_{\tau_n(T_i^k)-}^{n,M}|_{1,l-1}^l) \\ &\leq C_{l,p}(T) \bar{c}(Z_i^k) (1 + |X_{\tau_n(T_i^k)-}^{n,M}|_l + |X_{\tau_n(T_i^k)-}^{n,M}|_{l-1}^l). \end{aligned}$$

It follows that

$$\begin{aligned}
A_3 &\leq \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} |\tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|_l \right)^p \leq C_{l,p}(T) \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \tilde{c}(Z_i^k) (1 + |X_{\tau_n(T_i^k)-}^{n,M}|_l + |X_{\tau_n(T_i^k)-}^{n,M}|_{l-1}^l) \right)^p \\
&= C_{l,p}(T) \mathbb{E} \left( \int_0^t \int_{[1,M)} \tilde{c}(z) (1 + |X_{\tau_n(r)-}^{n,M}|_l + |X_{\tau_n(r)-}^{n,M}|_{l-1}^l) N_\nu(dr, dz) \right)^p \\
&\leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr],
\end{aligned} \tag{74}$$

where the last inequality is obtained by using (5) and recurrence hypothesis. Then combining (72), (73) and (74),

$$\mathbb{E} |X_t^{n,M}|_l^p \leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr]. \tag{75}$$

So  $\mathbb{E} |X_{\tau_n(t)}^{n,M}|_l^p \leq C_{l,p}(T) [1 + \int_0^{\tau_n(t)} \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr] \leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |X_{\tau_n(r)}^{n,M}|_l^p dr]$ . We denote temporarily  $g(t) = \mathbb{E} |X_{\tau_n(t)}^{n,M}|_l^p$ , then we have  $g(t) \leq C_{l,p}(T) [1 + \int_0^t g(r) dr]$ . By Gronwall's lemma,  $g(t) \leq C_{l,p}(T) e^{TC_{l,p}(T)}$ , which means that

$$\mathbb{E} |X_{\tau_n(t)}^{n,M}|_l^p \leq C_{l,p}(T) e^{TC_{l,p}(T)}.$$

Substituting into (75), we conclude that

$$\sup_{n,M} \mathbb{E} |X_t^{n,M}|_l^p \leq C_{l,p}(T). \tag{76}$$

As a summary of the recurrence argument, we remark that the uniform bound in  $n, M$  of the operator  $D$  for  $l = 0$  is due to the **Hypothesis 2.1**, and it propagates to larger  $l$  thanks to **Lemma 5.3 a)**.

b) Now we prove (70) for  $0 \leq l \leq q^* - 2$ , by recurrence on  $l$ .

**Step 1** One has to check that (70) holds for  $l = 0$ . The proof is analogous to that in the following **Step 2**, but simpler. It is done by using **Lemma 5.3 c)**, (60), Burkholder's inequality, **Hypothesis 2.1, 2.4**, (53), (69) and Gronwall's lemma. So we skip it.

**Step 2** Now we assume that (70) holds for  $l - 1$ , with  $l \geq 1$  and for any  $p \geq 2$  and we prove that it holds for  $l$  and for any  $p \geq 2$ . We write  $\mathbb{E} |LX_t^{n,M}|_l^p \leq C_{l,p}(T) (B_1 + B_2 + B_3)$ , with

$$\begin{aligned}
B_1 &= \mathbb{E} |L \int_0^t b_M(\tau_n(r), X_{\tau_n(r)}^{n,M}) dr|_l^p, \\
B_2 &= \mathbb{E} |L \int_0^t \int_{\{z \geq M\}} \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M}) W_\nu(dr, dz)|_l^p, \\
B_3 &= \mathbb{E} |L \int_0^t \int_{[1,M)} \tilde{c}(\tau_n(r), z, X_{\tau_n(r)-}^{n,M}) N_\nu(dr, dz)|_l^p.
\end{aligned}$$

Using **Lemma 5.3 b)**, **Hypothesis 2.1**, the recurrence hypothesis and (69), we get

$$\begin{aligned}
B_1 &\leq C_{l,p}(T) \mathbb{E} \int_0^t |Lb_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|_l^p dr \\
&\leq C_{l,p}(T) \mathbb{E} \int_0^t |\partial_x b_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|^p |LX_{\tau_n(r)}^{n,M}|_l^p dr \\
&+ \mathbb{E} \int_0^t \sup_{2 \leq |\beta| \leq l+2} |\partial_x^\beta b_M(\tau_n(r), X_{\tau_n(r)}^{n,M})|^p (1 + |X_{\tau_n(r)}^{n,M}|_{l+1}^{(l+2)p}) (1 + |LX_{\tau_n(r)}^{n,M}|_{l-1}^p) dr \\
&\leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |LX_{\tau_n(r)}^{n,M}|_l^p dr].
\end{aligned} \tag{77}$$



Then by **Lemma 5.2 b)**, we get

$$\begin{aligned}
B_2 &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |L\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|_l^2 \nu(dz))^{\frac{p}{2}} dr \\
&\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|_l^2 \nu(dz))^{\frac{p}{2}} dr] \\
&:= C_{l,p}(T) [B_{2,1} + B_{2,2}].
\end{aligned}$$

As a consequence of **Lemma 5.3 b)**, we have

$$\begin{aligned}
B_{2,1} &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\partial_x \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 |LX_{\tau_n(r)}^{n,M}|_l^2 \nu(dz))^{\frac{p}{2}} dr \\
&\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} \sup_{2 \leq |\beta| \leq l+2} |\partial_x^\beta \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 (1 + |X_{\tau_n(r)}^{n,M}|_{l+1}^{2(l+2)}) (1 + |LX_{\tau_n(r)}^{n,M}|_{l-1}^2) \nu(dz))^{\frac{p}{2}} dr].
\end{aligned}$$

And using **Lemma 5.3 a)**,

$$\begin{aligned}
B_{2,2} &\leq C_{l,p}(T) [\mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 \nu(dz))^{\frac{p}{2}} dr \\
&\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} |\partial_x \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 |X_{\tau_n(r)}^{n,M}|_{1,l}^2 \nu(dz))^{\frac{p}{2}} dr \\
&\quad + \mathbb{E} \int_0^t (\int_{\{z \geq M\}} \sup_{2 \leq |\beta| \leq l} |\partial_x^\beta \tilde{c}(\tau_n(r), \gamma_n(z), X_{\tau_n(r)}^{n,M})|^2 |X_{\tau_n(r)}^{n,M}|_{1,l-1}^{2l} \nu(dz))^{\frac{p}{2}} dr].
\end{aligned}$$

Then by **Hypothesis 2.1**, (53), (69) and the recurrence hypothesis,

$$B_2 \leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |LX_{\tau_n(r)}^{n,M}|_l^p dr]. \quad (78)$$

Now we estimate  $B_3$ . By **Lemma 5.3 b)** for  $d = 2$ , **Hypothesis 2.1**, for any  $k, i \in \mathbb{N}$ ,

$$\begin{aligned}
&|L\tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|_l \leq (|\partial_z \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})| + |\partial_x \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|) (|LZ_i^k|_l + |LX_{\tau_n(T_i^k)-}^{n,M}|_l) \\
&+ C_{l,p}(T) \sup_{2 \leq |\beta_1 + \beta_2| \leq l+2} (|\partial_z^{\beta_2} \partial_x^{\beta_1} \tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|) \\
&\times (1 + |Z_i^k|_{l+1}^{l+2} + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2}) (1 + |LZ_i^k|_{l-1} + |LX_{\tau_n(T_i^k)-}^{n,M}|_{l-1}) \\
&\leq C_{l,p}(T) \bar{c}(Z_i^k) (1 + |LZ_i^k|_l + |LX_{\tau_n(T_i^k)-}^{n,M}|_l + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \times (|LZ_i^k|_{l-1} + |LX_{\tau_n(T_i^k)-}^{n,M}|_{l-1})).
\end{aligned}$$

Then

$$\begin{aligned}
B_3 &\leq \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} |L\tilde{c}(\tau_n(T_i^k), Z_i^k, X_{\tau_n(T_i^k)-}^{n,M})|_l \right)^p \\
&\leq C_{l,p}(T) \mathbb{E} \left[ \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) (1 + |LZ_i^k|_l + |LX_{\tau_n(T_i^k)-}^{n,M}|_l + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \right. \\
&\quad \left. + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \times (|LZ_i^k|_{l-1} + |LX_{\tau_n(T_i^k)-}^{n,M}|_{l-1})) \right]^p \\
&\leq C_{l,p}(T) (B_{3,1} + B_{3,2} + B_{3,3}),
\end{aligned}$$

where

$$B_{3,1} = \mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LX_{\tau_n(T_i^k)-}^{n,M}|_l \right)^p,$$

$$B_{3,2} = \mathbb{E} \left| \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) (|LZ_i^k|_l + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \times |LZ_i^k|_{l-1}) \right|^p,$$

$$B_{3,3} = \mathbb{E} \left| \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) (1 + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} + |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \times |LX_{\tau_n(T_i^k)-}^{n,M}|_{l-1}) \right|^p.$$

By (5),

$$\begin{aligned} B_{3,1} &= \mathbb{E} \left| \int_0^t \int_{[1,M)} \bar{c}(z) |LX_{\tau_n(r)-}^{n,M}|_l N_\nu(dr, dz) \right|^p \\ &\leq C_{l,p}(T) \int_0^t \mathbb{E} |LX_{\tau_n(r)-}^{n,M}|_l^p dr. \end{aligned} \quad (79)$$

Using Schwartz's inequality, (5) and (69), we have

$$\begin{aligned} &\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{l+2} \times |LZ_i^k|_{l-1} \right)^p \\ &\leq [\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |X_{\tau_n(T_i^k)-}^{n,M}|_{l+1}^{2(l+2)} \right)^p]^{\frac{1}{2}} \times [\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LZ_i^k|_{l-1}^2 \right)^p]^{\frac{1}{2}} \\ &= [\mathbb{E} \left| \int_0^t \int_{[1,M)} \bar{c}(z) |X_{\tau_n(r)-}^{n,M}|_{l+1}^{2(l+2)} N_\nu(dr, dz) \right|^p]^{\frac{1}{2}} \times [\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LZ_i^k|_{l-1}^2 \right)^p]^{\frac{1}{2}} \\ &\leq C_{l,p}(T) [\mathbb{E} \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \bar{c}(Z_i^k) |LZ_i^k|_{l-1}^2 \right)^p]^{\frac{1}{2}}. \end{aligned}$$

Then applying **Lemma 5.4**, we get

$$B_{3,2} \leq C_{l,p}(T). \quad (80)$$

By (5), (69) and recurrence hypothesis, we have

$$\begin{aligned} B_{3,3} &= \mathbb{E} \left| \int_0^t \int_{[1,M)} \bar{c}(z) (1 + |X_{\tau_n(r)-}^{n,M}|_{l+1}^{l+2} + |X_{\tau_n(r)-}^{n,M}|_{l+1}^{l+2} \times |LX_{\tau_n(r)-}^{n,M}|_{l-1}) N_\nu(dr, dz) \right|^p \\ &\leq C_{l,p}(T). \end{aligned} \quad (81)$$

So by (79), (80) and (81),

$$B_3 \leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |LX_{\tau_n(r)-}^{n,M}|_l^p dr]. \quad (82)$$

Then combining (77), (78) and (82),

$$\mathbb{E} |LX_t^{n,M}|_l^p \leq C_{l,p}(T) [1 + \int_0^t \mathbb{E} |LX_{\tau_n(r)-}^{n,M}|_l^p dr], \quad (83)$$

Using Gronwall's lemma for (83) as for (75), we conclude that

$$\sup_{n,M} \mathbb{E} |LX_t^{n,M}|_l^p \leq C_{l,p}(T). \quad (84)$$

As a summary of the recurrence argument, we remark that the uniform bound in  $n, M$  of the operator  $L$  for  $l = 0$  is due to the **Hypothesis 2.1, 2.4** and **Lemma 5.3 c)**, and it propagates to larger  $l$  thanks to **Lemma 5.3 b)**.  $\square$

*Proof of Lemma 4.1.*

By **Lemma 5.1** and **Lemma 5.5**, as a consequence of **Lemma 3.2**, we have  $X_t^M \in \mathcal{D}_{l,p}$  and  $\sup_M \|X_t^M\|_{L,l,p} \leq C_{l,p}(T)$ .  $\square$

## 5.2 Proof of Lemma 4.2

In the following, we turn to the non-degeneracy of  $X_t^M$ . We consider the approximate equation (49)

$$X_t^M = x + \int_0^t \int_{[1,M)} \tilde{c}(r, z, X_{r-}^M) N_\nu(dr, dz) + \int_0^t b_M(r, X_r^M) dr + \int_0^t \int_{\{z \geq M\}} \tilde{c}(r, z, X_r^M) W_\nu(dr, dz).$$

We can calculate the Malliavin derivatives of the Euler scheme and then by passing to the limit, we have

$$\begin{aligned} D_{(k,i)}^Z X_t^M &= \mathbb{1}_{\{k \leq M-1\}} \mathbb{1}_{\{i \leq J_t^k\}} \xi_i^k \partial_z \tilde{c}(T_i^k, Z_i^k, X_{T_i^k-}^M) + \int_{T_i^k}^t \int_{[1,M)} \partial_x \tilde{c}(r, z, X_{r-}^M) D_{(k,i)}^Z X_{r-}^M N_\nu(dr, dz) \\ &+ \int_{T_i^k}^t \partial_x b_M(r, X_r^M) D_{(k,i)}^Z X_s^M dr + \int_{T_i^k}^t \int_{\{z \geq M\}} \partial_x \tilde{c}(r, z, X_r^M) D_{(k,i)}^Z X_s^M W_\nu(dr, dz). \end{aligned} \quad (85)$$

$$\begin{aligned} D_{(s,z_0)}^W X_t^M &= \int_s^t \int_{[1,M)} \partial_x \tilde{c}(r, z, X_{r-}^M) D_{(s,z_0)}^W X_{r-}^M N_\nu(dr, dz) + \int_s^t \partial_x b_M(r, X_r^M) D_{(s,z_0)}^W X_r^M dr \\ &+ \mathbb{1}_{\{s \leq t\}} \mathbb{1}_{\{z_0 \geq M\}} \tilde{c}(s, z_0, X_s^M) + \int_s^t \int_{\{z \geq M\}} \partial_x \tilde{c}(r, z, X_r^M) D_{(s,z_0)}^W X_r^M W_\nu(dr, dz). \end{aligned} \quad (86)$$

We obtain now some explicit expressions for the Malliavin derivatives. We consider the tangent flow  $(Y_t^M)_{t \in [0, T]}$  which is the solution of the linear equation

$$Y_t^M = 1 + \int_0^t \int_{[1,M)} \partial_x \tilde{c}(r, z, X_{r-}^M) Y_{r-}^M N_\nu(dr, dz) + \int_0^t \partial_x b_M(r, X_r^M) Y_r^M dr + \int_0^t \int_{\{z \geq M\}} \partial_x \tilde{c}(r, z, X_r^M) Y_r^M W_\nu(dr, dz).$$

And using Itô's formula,  $\bar{Y}_t^M = 1/Y_t^M$  verifies the equation

$$\begin{aligned} \bar{Y}_t^M &= 1 - \int_0^t \int_{[1,M)} \partial_x \tilde{c}(r, z, X_{r-}^M) (1 + \partial_x \tilde{c}(r, z, X_{r-}^M))^{-1} \bar{Y}_{r-}^M N_\nu(dr, dz) - \int_0^t \partial_x b_M(r, X_r^M) \bar{Y}_r^M dr \\ &- \int_0^t \int_{\{z \geq M\}} \partial_x \tilde{c}(r, z, X_r^M) \bar{Y}_r^M W_\nu(dr, dz) + \frac{1}{2} \int_0^t \int_{\{z \geq M\}} |\partial_x \tilde{c}(r, z, X_r^M)|^2 \bar{Y}_r^M \nu(dz) dr. \end{aligned}$$

Applying **Hypothesis 2.1** with  $q^* \geq 1$  and **Hypothesis 2.2**, with  $K_p$  a constant only depending on  $p$ , one also has (the proof is standard)

$$\mathbb{E}(\sup_{s \leq t} (|Y_s^M|^p + |\bar{Y}_s^M|^p)) \leq K_p < \infty. \quad (87)$$

*Remark.* Due to (4), we have

$$\max \left\{ \int_{[1,M)} |\bar{c}(z)|^p \nu(dz), \int_{[M,\infty)} |\bar{c}(z)|^p \nu(dz) \right\} \leq \int_{[1,\infty)} |\bar{c}(z)|^p \nu(dz) = \bar{c}_p,$$

so the constant in (87) is uniform with respect to  $M$ .

Then using the uniqueness of solution to the equation (85) and (86), one obtains

$$\begin{aligned} D_{(k,i)}^Z X_t^M &= \mathbb{1}_{\{k \leq M-1\}} \mathbb{1}_{\{i \leq J_t^k\}} \xi_i^k Y_t^M \bar{Y}_{T_i^k-}^M \partial_z \tilde{c}(T_i^k, Z_i^k, X_{T_i^k-}^M), \\ D_{(s,z_0)}^W X_t^M &= \mathbb{1}_{\{s \leq t\}} \mathbb{1}_{\{z_0 \geq M\}} Y_t^M \bar{Y}_s^M \tilde{c}(s, z_0, X_s^M). \end{aligned} \quad (88)$$

And the Malliavin covariance of  $X_t^M$  is

$$\sigma_{X_t^M} = \langle DX_t^M, DX_t^M \rangle_{\mathcal{H}} = \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} |D_{(k,i)}^Z X_t^M|^2 + \int_0^T \int_{\{z \geq M\}} |D_{(s,z)}^W X_t^M|^2 \nu(dz) ds. \quad (89)$$

In the following, we denote  $\lambda_t^M = \sigma_{X_t^M}$ . So the aim is to prove that for every  $p \geq 1$ ,

$$\mathbb{E}(|\lambda_t^M|^{-p}) \leq C_p. \quad (90)$$

We proceed in 5 steps.

**Step 1** We notice that by (88) and (89)

$$\lambda_t^M = \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \xi_i^k |Y_t^M|^2 |\bar{Y}_{T_i^k-}^M|^2 |\partial_z \tilde{c}(T_i^k, Z_i^k, X_{T_i^k-}^M)|^2 + |Y_t^M|^2 \int_0^t |\bar{Y}_s^M|^2 \int_{\{z \geq M\}} |\tilde{c}(s, z, X_s^M)|^2 \nu(dz) ds.$$

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

$$|\partial_z \tilde{c}(s, z, x)|^2 \geq \underline{c}(z) \quad \text{and} \quad |\tilde{c}(s, z, x)|^2 \geq \underline{c}(z).$$

In particular

$$\int_{\{z \geq M\}} |\tilde{c}(s, z, x)|^2 \nu(dz) \geq \int_{\{z \geq M\}} \underline{c}(z) \nu(dz),$$

so that

$$\lambda_t^M \geq Q_t^{-2} \times \left( \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \xi_i^k \underline{c}(Z_i^k) + t \int_{\{z \geq M\}} \underline{c}(z) \nu(dz) \right) \quad \text{with} \quad Q_t = \inf_{s \leq t} |Y_s^M \bar{Y}_t^M|.$$

We denote

$$\rho_t^M = \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \xi_i^k \underline{c}(Z_i^k), \quad \bar{\rho}_t^M = \sum_{k=M}^{\infty} \sum_{i=1}^{J_t^k} \xi_i^k \underline{c}(Z_i^k), \quad \alpha^M = \int_{\{z \geq M\}} \underline{c}(z) \nu(dz).$$

By (87),  $(\mathbb{E} \sup_{s \leq t} |Y_s^M \bar{Y}_t^M|^{4p})^{1/2} \leq C < \infty$ , so that

$$\mathbb{E}(|\lambda_t^M|^{-p}) \leq C(\mathbb{E}(|\rho_t^M + t\alpha^M|^{-2p}))^{\frac{1}{2}}. \quad (91)$$

**Step 2** Let  $\Gamma(p) = \int_0^\infty s^{p-1} e^{-s} ds$ . By a change of variables, we have the numerical equality

$$\frac{1}{(\rho_t^M + t\alpha^M)^p} = \frac{1}{\Gamma(p)} \int_0^\infty s^{p-1} e^{-s(\rho_t^M + t\alpha^M)} ds$$

which, by taking expectation, gives

$$\mathbb{E}\left(\frac{1}{(\rho_t^M + t\alpha^M)^p}\right) = \frac{1}{\Gamma(p)} \int_0^\infty s^{p-1} \mathbb{E}(e^{-s(\rho_t^M + t\alpha^M)}) ds. \quad (92)$$

**Step 3 (splitting).** In order to compute  $\mathbb{E}(e^{-s(\rho_t^M + t\alpha^M)})$  we have to interpret  $\rho_t^M$  in terms of Poisson measures. We recall that we suppose the "splitting hypothesis" (40):

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k dz,$$

with  $I_k = [k, k+1)$ ,  $m_k = \nu(I_k)$ . We also have the function  $\psi$  and  $m(\psi) = \int_{\mathbb{R}} \psi(t) dt$ . And we use the basic decomposition

$$Z_i^k = \xi_i^k V_i^k + (1 - \xi_i^k) U_i^k$$

where  $V_i^k, U_i^k, \xi_i^k, k, i \in \mathbb{N}$  are some independent random variables with laws given in (47).

For every  $k$  we consider a Poisson point measure  $N_k(ds, d\xi, dv, du)$  with  $\xi \in \{0, 1\}, v, u \in [1, \infty), s \in [0, T]$  with compensator

$$\begin{aligned}\widehat{N}_k(ds, d\xi, dv, du) &= \widehat{M}_k(d\xi, dv, du) \times ds \\ \text{with } \widehat{M}_k(d\xi, dv, du) &= b_k(d\xi) \times \mathbb{1}_{I_k}(v) \frac{1}{m(\psi)} \psi(v - (k + \frac{1}{2})) dv \\ &\quad \times \frac{1}{1 - \varepsilon_k m(\psi)} \mathbb{1}_{I_k}(u) (\mathbb{P}(Z_1^k \in du) - \varepsilon_k \psi(u - (k + \frac{1}{2}))) du.\end{aligned}$$

Here  $b_k(d\xi)$  is the Bernoulli law of parameter  $\varepsilon_k m(\psi)$ . The intervals  $I_k, k \in \mathbb{N}$  are disjoint so the Poisson point measures  $N_k, k = 1, \dots, M-1$  are independent. Then

$$\sum_{i=1}^{J_t^k} \xi_i^k \mathcal{C}(Z_i^k) = \sum_{i=1}^{J_t^k} \xi_i^k \mathcal{C}(\xi_i^k V_i^k + (1 - \xi_i^k) U_i^k) = \int_0^t \int_{\{0,1\}} \int_{[1,\infty)^2} \xi \mathcal{C}(\xi v + (1 - \xi)u) N_k(ds, d\xi, dv, du).$$

In order to get compact notation, we put together all the measures  $N_k, k \leq M-1$ . Since they are independent we get a new Poisson point measure that we denote by  $\Theta$ . And we have

$$\rho_t^M = \sum_{k=1}^{M-1} \sum_{i=1}^{J_t^k} \xi_i^k \mathcal{C}(Z_i^k) = \int_0^t \int_{\{0,1\}} \int_{[1,\infty)^2} \xi \mathcal{C}(\xi v + (1 - \xi)v) \Theta(ds, d\xi, dv, du).$$

**Step 4** Using Itô's formula,

$$\begin{aligned}\mathbb{E}(e^{-s\rho_t^M}) &= 1 + \mathbb{E} \int_0^t \int_{\{0,1\}} \int_{[1,\infty)^2} (e^{-s(\rho_{r-}^M + \xi \mathcal{C}(\xi v + (1-\xi)v))} - e^{-s\rho_{r-}^M}) \widehat{\Theta}(dr, d\xi, dv, du) \\ &= 1 - \int_0^t \mathbb{E}(e^{-s\rho_{r-}^M}) dr \int_{\{0,1\}} \int_{[1,\infty)^2} (1 - e^{-s\xi \mathcal{C}(\xi v + (1-\xi)v)}) \sum_{k=1}^{M-1} \widehat{M}_k(d\xi, dv, du).\end{aligned}$$

Solving the above equation we obtain

$$\mathbb{E}(e^{-s\rho_t^M}) = \exp(-t \sum_{k=1}^{M-1} \int_{\{0,1\}} \int_{[1,\infty)^2} (1 - e^{-s\xi \mathcal{C}(\xi v + (1-\xi)v)}) \widehat{M}_k(d\xi, dv, du)).$$

We compute

$$\int_{\{0,1\} \times [1,\infty)^2} (1 - e^{-s\xi \mathcal{C}(\xi v + (1-\xi)v)}) \widehat{M}_k(d\xi, dv, du) = \varepsilon_k m(\psi) \int_k^{k+1} (1 - e^{-s\mathcal{C}(v)}) \frac{1}{m(\psi)} \psi(v - (k + \frac{1}{2})) dv.$$

Since  $\psi \geq 0$  and  $\psi(z) = 1$  if  $|z| \leq \frac{1}{4}$  it follows that the above term is larger than

$$\varepsilon_k \int_{k+\frac{1}{4}}^{k+\frac{3}{4}} (1 - e^{-s\mathcal{C}(v)}) dv.$$

Finally this gives

$$\begin{aligned}\mathbb{E}(e^{-s\rho_t^M}) &\leq \exp(-t \sum_{k=1}^{M-1} \varepsilon_k \int_{k+\frac{1}{4}}^{k+\frac{3}{4}} (1 - e^{-s\mathcal{C}(v)}) dv) \\ &= \exp(-t \int_1^M (1 - e^{-s\mathcal{C}(v)}) m(dv)),\end{aligned}$$

with

$$m(dv) = \sum_{k=1}^{\infty} \varepsilon_k \mathbb{1}_{(k+\frac{1}{4}, k+\frac{3}{4})}(v) dv. \tag{93}$$

In the same way, we get

$$\mathbb{E}(e^{-s\bar{\rho}_t^M}) \leq \exp(-t \int_M^\infty (1 - e^{-s\underline{c}(v)})m(dv)).$$

Notice that  $t\alpha^M \geq \mathbb{E}(\bar{\rho}_t^M)$ . Then using Jensen's inequality for the convex function  $f(x) = e^{-sx}$ ,  $s, x > 0$ , we have

$$e^{-st\alpha^M} \leq e^{-s\mathbb{E}\bar{\rho}_t^M} \leq \mathbb{E}(e^{-s\bar{\rho}_t^M}) \leq \exp(-t \int_M^\infty (1 - e^{-s\underline{c}(v)})m(dv)).$$

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

$$\begin{aligned} \mathbb{E}(e^{-s(\rho_t^M + t\alpha^M)}) &= e^{-st\alpha^M} \times \mathbb{E}(e^{-s\rho_t^M}) \\ &\leq \exp(-t \int_M^\infty (1 - e^{-s\underline{c}(v)})m(dv)) \times \exp(-t \int_1^M (1 - e^{-s\underline{c}(v)})m(dv)) \\ &= \exp(-t \int_1^\infty (1 - e^{-s\underline{c}(v)})m(dv)), \end{aligned} \tag{94}$$

and the last term does not depend on  $M$ .

Now we will use the Lemma 14 from [7], which states the following.

**Lemma 5.6.** *We consider an abstract measurable space  $E$ , a  $\sigma$ -finite measure  $\eta$  on this space and a non-negative measurable function  $f : E \rightarrow \mathbb{R}_+$  such that  $\int_E f d\eta < \infty$ . For  $t > 0$  and  $p \geq 1$ , we note*

$$\alpha_f(t) = \int_E (1 - e^{-tf(a)})\eta(da) \quad \text{and} \quad I_t^p(f) = \int_0^\infty s^{p-1} e^{-t\alpha_f(s)} ds.$$

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

$$\liminf_{u \rightarrow \infty} \frac{1}{\ln u} \eta(f \geq \frac{1}{u}) > p/t, \tag{95}$$

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

We will use the above lemma for  $\eta = m$  and  $f = \underline{c}$ . So if we have

$$\liminf_{u \rightarrow \infty} \frac{1}{\ln u} m(\underline{c} \geq \frac{1}{u}) = \infty, \tag{96}$$

then for every  $p \geq 1, t > 0, M \geq 1$ , (92), (94) and Lemma 5.6 give

$$\begin{aligned} \mathbb{E}(\frac{1}{\rho_t^M + t\alpha^M})^{2p} &= \frac{1}{\Gamma(2p)} \int_0^\infty s^{2p-1} \mathbb{E}(e^{-s(\rho_t^M + t\alpha^M)}) ds \\ &\leq \frac{1}{\Gamma(2p)} \int_0^\infty s^{2p-1} \exp(-t \int_1^\infty (1 - e^{-s\underline{c}(v)})m(dv)) ds < \infty. \end{aligned} \tag{97}$$

Finally using (91), we conclude that if (96) holds, then

$$\sup_M \mathbb{E}(\lambda_t^M)^{-p} < \infty. \tag{98}$$

**Step 5** Now the only problem left is to compute  $m(\underline{c} \geq \frac{1}{u})$ . It seems difficult to discuss this in a completely abstract framework. So we suppose **Hypothesis 2.4** (a): There exists a constant  $\varepsilon_* > 0$  and there are some  $\alpha_1 > \alpha_2 > 0$  such that for every  $k \in \mathbb{N}$ ,

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k dz \quad \text{with} \quad \varepsilon_k = \frac{\varepsilon_*}{(k+1)^{1-\alpha}}, \text{ for any } \alpha \in (\alpha_2, \alpha_1], \quad \text{and} \quad \underline{c}(z) \geq e^{-z^{\alpha_2}},$$

Then  $\{z : \underline{c}(z) \geq \frac{1}{u}\} \supseteq \{z : (\ln u)^{1/\alpha_2} \geq z\}$ . In particular, for  $k \leq \lfloor (\ln u)^{1/\alpha_2} \rfloor - 1 := k(u)$ , one has  $I_k \subseteq \{z : \underline{c}(z) \geq \frac{1}{u}\}$ . Then for  $u$  large enough, we compute

$$\begin{aligned} m(\underline{c} \geq \frac{1}{u}) &\geq \sum_{k=1}^{k(u)} m(I_k) \geq \frac{1}{2} \sum_{k=1}^{k(u)} \varepsilon_k \geq \frac{1}{2} \varepsilon_* \sum_{k=1}^{k(u)} \frac{1}{(k+1)^{1-\alpha}} \geq \frac{1}{2} \varepsilon_* \int_2^{(\ln u)^{1/\alpha_2}} \frac{1}{z^{1-\alpha}} dz \\ &= \frac{\varepsilon_*}{2\alpha} ((\ln u)^{\alpha/\alpha_2} - 2^\alpha). \end{aligned}$$

Since  $\alpha > \alpha_2$ , (96) is verified and we obtain (98).

Now we consider **Hypothesis 2.4** (b): We suppose that there exists a constant  $\varepsilon_* > 0$  and there are some  $\alpha > 0$  such that for every  $k \in \mathbb{N}$ ,

$$\mathbb{1}_{I_k}(z) \frac{\nu(dz)}{m_k} \geq \mathbb{1}_{I_k}(z) \varepsilon_k dz \quad \text{with} \quad \varepsilon_k = \frac{\varepsilon_*}{k+1}, \quad \text{and} \quad \underline{c}(z) \geq \frac{1}{z^\alpha}.$$

Now  $\{z : \underline{c}(z) \geq \frac{1}{u}\} \supseteq \{z : z \leq u^{1/\alpha}\}$ . Then for  $u$  large enough,

$$m(\underline{c} \geq \frac{1}{u}) \geq \frac{1}{2} \varepsilon_* \sum_{k=1}^{\lfloor u^{1/\alpha} \rfloor - 1} \frac{1}{k+1} \geq \frac{1}{2} \varepsilon_* \int_2^{u^{1/\alpha}} \frac{dz}{z} = \frac{1}{2} \varepsilon_* (\frac{1}{\alpha} \ln u - \ln 2).$$

And consequently

$$\lim_{u \rightarrow \infty} \frac{1}{\ln u} m(\underline{c} \geq \frac{1}{u}) \geq \frac{\varepsilon_*}{2\alpha}.$$

Using **Lemma 5.6**, this gives: if

$$\frac{2p}{t} < \frac{\varepsilon_*}{2\alpha} \quad \Leftrightarrow \quad t > \frac{4p\alpha}{\varepsilon_*}$$

then

$$\sup_M \mathbb{E} \left( \frac{1}{\rho_t^M + t\alpha^M} \right)^{2p} < \infty,$$

and we have  $\sup_M \mathbb{E}(\lambda_t^M)^{-p} < \infty$ . □

### 5.3 Some proofs concerning Section 4.2

We will prove that the triplet  $(\mathcal{S}, D, L)$  defined in Section 4.2 is an IbP framework. Here, we only show that  $D^q$  is closable and  $L$  verifies the duality formula (32). To do so, we introduce the divergence operator  $\delta$ . We denote the space of simple processes by

$$\mathcal{P} = \{u = ((\bar{u}_i^k)_{1 \leq i \leq m', \sum_{r=1}^n u_r \varphi_r}) : \bar{u}_i^k, u_r \in \mathcal{S}, \varphi_r \in L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb), m', m, n \in \mathbb{N}\}.$$

For  $u = ((\bar{u}_i^k)_{1 \leq i \leq m', \sum_{r=1}^n u_r \varphi_r}) \in \mathcal{P}$ , we denote  $u^Z = (\bar{u}_i^k)_{1 \leq i \leq m', \sum_{r=1}^n u_r \varphi_r}$  and  $u^W = \sum_{r=1}^n u_r \varphi_r$ , so that  $u = (u^Z, u^W)$ .

We notice that  $\mathcal{P}$  is dense in  $L^2(\Omega; \mathcal{H})$ , with  $\mathcal{H} = l_2 \otimes L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)$ .

Then we define the divergence operator  $\delta : \mathcal{P} \rightarrow \mathcal{S}$  by

$$\begin{aligned} \delta(u) &= \delta^Z(u^Z) + \delta^W(u^W) \\ \text{with} \quad \delta^Z(u^Z) &= - \sum_{k=1}^m \sum_{i=1}^{m'} (D_{(k,i)}^Z \bar{u}_i^k + \xi_i^k \bar{u}_i^k \times \theta_k(V_i^k)) \\ \delta^W(u^W) &= \sum_{r=1}^n u_r W_\nu(\varphi_r) - \sum_{r=1}^n \langle D^W u_r, \varphi_r \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)}. \end{aligned}$$

We will show that  $\delta$  satisfies the following duality formula: For every  $F \in \mathcal{S}, u \in \mathcal{P}$ ,

$$\mathbb{E}\langle DF, u \rangle_{\mathcal{H}} = \mathbb{E}F\delta(u). \quad (99)$$

In fact, if we denote  $\hat{V}_i^k(x)$  the sequence  $(V_{i_0}^{k_0})_{\substack{1 \leq i_0 \leq m' \\ 1 \leq k_0 \leq m}}$  after replacing  $V_i^k$  by  $x$ , then for any  $m', m \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E}\langle D^Z F, u^Z \rangle_{l_2} &= \mathbb{E} \sum_{k=1}^m \sum_{i=1}^{m'} D_{(k,i)}^Z F \times \bar{u}_i^k \\ &= \sum_{k=1}^m \sum_{i=1}^{m'} \mathbb{E} \xi_i^k \partial_{v_i^k} f(\omega, (V_{i_0}^{k_0})_{\substack{1 \leq i_0 \leq m' \\ 1 \leq k_0 \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) \bar{u}_i^k(\omega, (V_{i_0}^{k_0})_{\substack{1 \leq i_0 \leq m' \\ 1 \leq k_0 \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) \\ &= \sum_{k=1}^m \sum_{i=1}^{m'} \mathbb{E} \int_{\mathbb{R}} \xi_i^k \partial_{v_i^k} f(\omega, \hat{V}_i^k(x), (W_\nu(\varphi_j))_{j=1}^n) \times \bar{u}_i^k(\omega, \hat{V}_i^k(x), (W_\nu(\varphi_j))_{j=1}^n) \frac{\psi_k(x)}{m(\psi)} dx \\ &= - \sum_{k=1}^m \sum_{i=1}^{m'} \mathbb{E} \int_{\mathbb{R}} \xi_i^k f(\omega, \hat{V}_i^k(x), (W_\nu(\varphi_j))_{j=1}^n) \times [\partial_{v_i^k} \bar{u}_i^k(\omega, \hat{V}_i^k(x), (W_\nu(\varphi_j))_{j=1}^n) \\ &\quad + \bar{u}_i^k(\omega, \hat{V}_i^k(x), (W_\nu(\varphi_j))_{j=1}^n) \frac{\partial_x \psi_k(x)}{\psi_k(x)}] \frac{\psi_k(x)}{m(\psi)} dx \\ &= - \sum_{k=1}^m \sum_{i=1}^{m'} \mathbb{E} F[D_{(k,i)}^Z \bar{u}_i^k + \xi_i^k \bar{u}_i^k \partial_x (\ln \psi_k(V_i^k))] = \mathbb{E}(F \delta^Z(u^Z)). \end{aligned}$$

On the other hand, since  $L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)$  is a separable Hilbert space, we can assume without loss of generality that, in the definition of simple functionals,  $(\varphi_1, \dots, \varphi_m, \dots)$  is the orthogonal basis of the space  $L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)$ .

Then with  $p_r = \int_{\mathbb{R}_+ \times \mathbb{R}_+} \varphi_r^2(s, z) \nu(dz) ds$ , for any  $n \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E}\langle D^W F, u^W \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)} &= \mathbb{E} \int_{\mathbb{R}_+ \times \mathbb{R}_+} D_{(s,z)}^W F \times \sum_{r=1}^n u_r \varphi_r(s, z) \nu(dz) ds \\ &= \mathbb{E} \sum_{r=1}^n \partial_{w_r} f(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) u_r(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, (W_\nu(\varphi_j))_{j=1}^n) p_r \\ &= \sum_{r=1}^n \mathbb{E} \int_{\mathbb{R}} \partial_{w_r} f(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, W_\nu(\varphi_1), \dots, W_\nu(\varphi_{r-1}), y, W_\nu(\varphi_{r+1}), \dots, W_\nu(\varphi_n)) \\ &\quad \times u_r(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, W_\nu(\varphi_1), \dots, W_\nu(\varphi_{r-1}), y, W_\nu(\varphi_{r+1}), \dots, W_\nu(\varphi_n)) \frac{1}{\sqrt{2\pi p_r}} e^{-\frac{y^2}{2p_r}} dy \times p_r \\ &= - \sum_{r=1}^n \mathbb{E} \int_{\mathbb{R}} f(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, W_\nu(\varphi_1), \dots, W_\nu(\varphi_{r-1}), y, W_\nu(\varphi_{r+1}), \dots, W_\nu(\varphi_n)) \\ &\quad \times [\partial_{w_r} u_r(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, W_\nu(\varphi_1), \dots, W_\nu(\varphi_{r-1}), y, W_\nu(\varphi_{r+1}), \dots, W_\nu(\varphi_n)) \\ &\quad - \frac{y}{p_r} u_r(\omega, (V_i^k)_{\substack{1 \leq i \leq m' \\ 1 \leq k \leq m}}, W_\nu(\varphi_1), \dots, W_\nu(\varphi_{r-1}), y, W_\nu(\varphi_{r+1}), \dots, W_\nu(\varphi_n))] \frac{1}{\sqrt{2\pi p_r}} e^{-\frac{y^2}{2p_r}} dy \times p_r \\ &= \mathbb{E} F(\sum_{r=1}^n u_r W_\nu(\varphi_r) - \sum_{r=1}^n \langle D^W u_r, \varphi_r \rangle_{L^2(\mathbb{R}_+ \times \mathbb{R}_+, \nu \times Leb)}) = \mathbb{E}(F \delta^W(u^W)). \end{aligned}$$

Then (99) is proved. Using this duality formula recursively, we can show the closability of  $D^q$ . If there exists  $u \in L^2(\Omega; \mathcal{H}^{\otimes q})$  such that  $F_n \rightarrow 0$  in  $L^2(\Omega)$  and  $D^q F_n \rightarrow u$  in  $L^2(\Omega; \mathcal{H}^{\otimes q})$ , then for any  $h_1, \dots, h_q \in \mathcal{P}$ ,  $\mathbb{E}\langle u, h_1 \otimes \dots \otimes h_q \rangle_{\mathcal{H}^{\otimes q}} = \lim_{n \rightarrow \infty} \mathbb{E}\langle D^q F_n, h_1 \otimes \dots \otimes h_q \rangle_{\mathcal{H}^{\otimes q}} = \lim_{n \rightarrow \infty} \mathbb{E} F_n \delta(h_1 \delta(h_2 \dots \delta(h_q))) = 0$ . Since  $\mathcal{P}^{\otimes q}$  is dense in  $L^2(\Omega; \mathcal{H}^{\otimes q})$ , we conclude that  $u = 0$ . This implies that  $D^q$  is closable.



We notice that from the definition of  $\delta$  and  $L$ , we get immediately that  $LF = \delta(DF)$ ,  $\forall F \in \mathcal{S}$ . And if we replace  $u$  by  $DG$  in (99) for  $G \in \mathcal{S}$ , we get the duality formula of  $L$  (32).

**Data availability statement.** Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

## References

- [1] A. Alfonsi, E. Cancès, G. Turnici, B. Di Ventura, and W. Huisinga : Adaptive simulation of hybrid stochastic and deterministic models for biochemical systems. In ESAIM Proceedings, volume **14**, pages 1–13(2005).
- [2] A. Alfonsi, E. Cancès, G. Turnici, B. Di Ventura, and W. Huisinga: Exact simulation of hybrid stochastic and deterministic models for biochemical systems. Research Report RR-5435, INRIA (2004).
- [3] S. Asmussen and J. Rosinski: Approximations of small jumps of Lévy processes with a view towards simulation, *Journal of Applied Probability* **38** , 482-493 (2001).
- [4] K. Ball, T. G. Kurtz, L. Popovic, and G. Rempala: Asymptotic analysis of multiscale approximations to reaction networks. *Ann. Appl. Probab.*, **16**:1925–1961 (2006).
- [5] V. Bally, L. Caramellino, G. Poly: Convergence in distribution norms in the CLT for non identical distributed random variables. *Electronic Journal of Probability*, Institute of Mathematical Statistics(IMS), **23**, paper 45,51 p. (2018).
- [6] V. Bally, L. Caramellino, G. Poly: Regularization lemmas and convergence in total variation. (2020). <hal-02429512>
- [7] V. Bally, E. Clément: Integration by parts formula and applications to equations with jumps. *Probab. Th. Rel.Fields*, **151**, 613-657. (2011).
- [8] V. Bally, D. Goreac, V. Rabié: Regularity and stability for the semigroup of jump diffusions with state-dependent intensity. *The Annals of Applied Probability*, **28**(5), pp.3028-3074 (2018).
- [9] V. Bally, E. Pardoux: Malliavin Calculus for white noise driven parabolic SPDEs. *Potential Analysis* **9**, 27–64 (1998).
- [10] K. Bichteler, J. B. Gravereaux, J. Jacod: *Malliavin calculus for processes with jumps*. Gordon and Breach, (1987).
- [11] A. Carpentier, C. Duval, E. Mariucci: Total variation distance for discretely observed Lévy processes: a Gaussian approximation of the small jumps. arXiv:1810.02998 [math.ST] (2019).
- [12] R. Cont, P. Tankov: *Financial modelling with jump processes*. Chapman & Hall/CRC (2004).
- [13] A. Crudu, A. Debussche, A. Muller, O. Radulescu: Convergence of stochastic gene networks to hybrid piecewise deterministic processes, *Ann. Appl. Probab.* **10**, Nr 5, p 1822-1859.
- [14] S. Dereich: Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *Ann. Appl. Probab.* **21**(1): 283-311 (2011).
- [15] C. Duval and E. Mariucci: Spectral free estimates of Lévy densities in high level frequency, to appear in Bernoulli. arXiv 1702. 08787 [Math PR].
- [16] N. Fournier: Simulation and approximation of Lévy driven stochastic differential equations. *ESAIM Proba and Stat*, Vol **15**, 249-269, (2011).

- [17] C. Graham: McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications* **40** 69-82, (1992).
- [18] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes*. 2nd ed. Amsterdam, Netherlands, North Holland, (1989).
- [19] Y. Ishikawa: *Stochastic Calculus of Variations for Jump Processes*, Berlin, Boston: De Gruyter. (2013).
- [20] J. Jacod. and A. Shiryaev: *Limit theorems for stochastic processes*. Springer Verlag. (1987).
- [21] E. Jum: Numerical approximation of stochastic differential equations driven by Lévy motion with infinitely many jumps. PhD diss., University of Tennessee. (2015).
- [22] A. Kohatsu-Higa, P. Tankov: Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*. (2010).
- [23] A. Kohatsu-Higa, S. Ortiz-Latorre, P. Tankov: Optimal simulation schemes for Lévy driven stochastic differential equations. *Mathematics of Computation* **83**, 2293-2324, 201 (2010).
- [24] A. M. Kulik: Malliavin calculus for Lévy processes with arbitrary Lévy measures. *Theor. Probability and Math. Statist.* No. **72**, 75-92. (2006).
- [25] 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).
- [26] H. Kunita: Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In: Rao, MM, ed. *Real and stochastic analysis*. Boston, USA, Birkhäuser, 305-373. (2004).
- [27] H. Kunita: *Stochastic flows and jump-diffusions*. Springer, (2019).
- [28] T. Kurtz: Limits theorems for sequences of jump Markov processes approximating ordinary differential processes. *J. Appl. Prob.*, **8**:344–356 (1971).
- [29] T. Kurtz: Strong approximation theorems for density dependent Markov chains. *Stoch. Proc. Appl.*, **6**:223–240 (1978).
- [30] C. Marinelli, M. Röckner: On the maximal inequalities of Burkholder, Davis and Gundy. *ScienceDirect, Expo. Math.* **34** (2016).
- [31] E. Mariucci, M. Reiß: Wasserstein and total variation distance between marginals of Lévy processes. *Electronic Journal of Statistics* **12**, 2482-2514. (2018).
- [32] E. Mordecki, A. Szepessy, R. Tempone and G. E. Zouraris: Adaptive weak approximation of diffusions with jumps. arxiv:math/0609186v1 (2006).
- [33] D. Nualart: *The Malliavin calculus and related topics*. Springer-Verlag, (2006).
- [34] P. Protter and D. Talay: The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.* Vol **25**, No.1, pg 393-423 (1997).
- [35] Y. Song and X. Zhang: Regularity of density for SDEs driven by degenerate Lévy noises. arXiv:1401.4624 (2014).
- [36] J. B. Walsh: *An introduction to stochastic partial differential equations*. Ecole d'été de probabilités de Saint Flour, 1984, p 265-439, ed. L. Hennequin, Springer Verlag.
- [37] X. Zhang: Densities for SDEs driven by degenerate  $\alpha$ -stable processes. *Ann. Probab.* Vol **42**, No.5, 1885-1910 (2014).

## Part II

.

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

# 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

**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  $X_t^P$  of this equation converges to  $X_t$  in Wasserstein distance. In this paper, under more restricted assumptions, we show that the Euler scheme  $X_t^P$  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  $X_t^{P,M}$  which has a finite numbers of jumps in any compact interval. We prove that  $X_t^{P,M}$  also converges to  $X_t$  in total variation distance. Finally, we give an algorithm based on a particle system associated to  $X_t^{P,M}$  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

<b>1</b>	<b>Introduction</b>	<b>60</b>
<b>2</b>	<b>Main results</b>	<b>64</b>
2.1	Basic notations and the main equation . . . . .	64
2.2	Hypotheses . . . . .	64
2.3	The Euler scheme . . . . .	66
2.4	The truncated Euler scheme . . . . .	67
2.5	The particle system . . . . .	68
2.6	Some examples . . . . .	70
2.7	Preliminaries: coupling . . . . .	71
<b>3</b>	<b>Malliavin calculus</b>	<b>75</b>
3.1	Abstract integration by parts framework . . . . .	75
3.1.1	Main consequences . . . . .	80
3.2	Malliavin calculus for the jump equations . . . . .	85
3.3	Proofs of <b>Theorem 2.1~2.4</b> . . . . .	87
<b>4</b>	<b>Proofs</b>	<b>93</b>
4.1	Sobolev norms . . . . .	93
4.2	Covariance matrices . . . . .	98
<b>5</b>	<b>Appendix</b>	<b>101</b>

# 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  $X_t$ ,  $t \in [0, T]$ ,  $N_{\rho_t}$  is a Poisson point measure on the state space  $\mathbb{R}^d \times \mathbb{R}^d$  with intensity measure  $\rho_t(dv)\mu(dz)dr$ ,  $\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ ,  $X_0$  is the initial random variable independent of the Poisson point measure  $N_{\rho_t}$ , 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  $\beta_1, \beta_2$ , there exists a non-negative function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  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  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  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  $N_{\rho_t}$  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  $\mathcal{P} = \{0 = r_0 < r_1 < \dots < r_{n-1} < r_n = T\}$  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.

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

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

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

as  $|\mathcal{P}| \rightarrow 0$ , with  $|\mathcal{P}| := \max_{k \in \{0, \dots, n-1\}} (r_{k+1} - r_k)$ . 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  $a_T^M := \sqrt{T \int_{\{|z| > M\}} c(z) \mu(dz)}$ . Now we cancel the jumps of size  $|z| > M$  and we replace them by a Gaussian random variable.

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

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  $\rho_t^{\mathcal{P}, M}(dv) \mu(dz) dr$ ,  $\Delta$  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

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

as  $|\mathcal{P}| \rightarrow 0$  and  $M \rightarrow \infty$ , 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  $(X_0^1, \dots, X_0^N)$  with components which are independent and identically distributed with common law  $\rho_0$  (which is the law of  $X_0$ ), and  $(\Delta^1, \dots, \Delta^N)$  which is a  $N \times d$ -dimensional standard Gaussian random variable independent of  $(X_0^1, \dots, X_0^N)$ . Then we construct the particle system  $\vec{X}_t^{\mathcal{P}, M} = (X_t^{\mathcal{P}, M, 1}, \dots, X_t^{\mathcal{P}, M, N})$ :

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

where

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

is the empirical measure of  $\vec{X}_t^{\mathcal{P},M}$  (with  $\delta_x(dv)$  the Dirac measure),  $N_{\hat{\rho}(\vec{X}_t^{\mathcal{P},M})}^i(dv, dz, dr)$ ,  $i = 1, \dots, N$  are Poisson point measures that are independent each other conditionally to  $\vec{X}_t^{\mathcal{P},M}$  and independent of  $(X_0^1, \dots, X_0^N, \Delta^1, \dots, \Delta^N)$  with intensity  $\hat{\rho}(\vec{X}_t^{\mathcal{P},M})(dv)\mu(dz)dr$ . It is clear that  $\vec{X}_t^{\mathcal{P},M}$  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

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

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

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

Then we have

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

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

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

then we get moreover by Romberg method that

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

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  $\tilde{\Delta}$  a  $d$ -dimensional standard Gaussian random variable independent of  $\vec{X}_t^{\mathcal{P},M}$ . For any  $\varepsilon > 0$ , we take

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

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

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

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

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

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

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

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  $\delta$  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 Poisson 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 than  $M$  (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  $\mu$  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  $X_t^M \rightarrow X_t$ , 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 Wasserstein 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  $X_t^{\mathcal{P},M,i}$ ,  $i = 1, \dots, N$ . 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  $X_t^{\mathcal{P},M} \rightarrow X_t$  in total variation distance. We also prove that the Euler scheme  $X_t^{\mathcal{P}} \rightarrow X_t$  in total variation distance in a similar way. Furthermore, we obtain an algorithm based on the particle system  $X_t^{\mathcal{P},M,i}$ ,  $i = 1, \dots, N$  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  $\beta$ , we denote  $|\beta|$  to be the length of  $\beta$ . 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  $\|f\|_{l,\infty} := \sum_{|\beta| \leq l} \|\partial^\beta f\|_\infty$  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

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

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

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

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

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

where  $\rho_t(dv) = \mathbb{P}(X_t \in dv)$  is the law of  $X_t$ ,  $N_{\rho_t}$  is a Poisson point measure on the state space  $\mathbb{R}^d \times \mathbb{R}^d$  with intensity measure  $\hat{N}_{\rho_t}(dv, dz, dr) = \rho_t(dv) \mu(dz) dr$ ,  $X_0$  is the initial random variable with law  $\rho_0$  independent of the Poisson point measure  $N_{\rho_t}$ ,  $\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ , and  $b : [0, T] \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ ,  $c : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ .

*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  $\beta_1, \beta_2$ , there exists a non-negative function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  depending on  $\beta_1, \beta_2$  such that we have

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

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

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

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

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

*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  $\Phi(r, v, z, \omega, \rho) : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \Omega \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}_+$  and  $\varphi(r, v, \omega, \rho) : [0, T] \times \mathbb{R}^d \times \Omega \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}_+$  be two functions such that

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

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

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

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

*Proof.* By compensating  $N_\rho$  and using Burkholder inequality and (15), we have

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

□

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  $\check{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that  $\int_{\mathbb{R}^d} |\check{c}(z)|^p \mu(dz) := \check{c}_p < \infty, \forall p \geq 1$ , and

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

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

*Remark.* We need this hypothesis to prove the regularity of the inverse tangent flow (see Section 4.2 (131)).

**Hypothesis 2.3 (Ellipticity)** There exists a non-negative function  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  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  $\mu$ .

a) We assume that there exists a  $\theta > 0$  such that

$$\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \nu \left\{ \underline{c} \geq \frac{1}{u} \right\} := \theta > 0, \quad (18)$$

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 (138)).

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

b) We assume that  $\mu$  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  $\mathcal{P} = \{0 = r_0 < r_1 < \dots < r_{n-1} < r_n = T\}$  of the interval  $[0, T]$ , we define  $\tau(r) = r_k$  when  $r \in [r_k, r_{k+1})$ , and we consider the equation

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

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  $[0, T]$ ,

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

b) The solution of the following weak equation exists.

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

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  $\theta$  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  $p_t$ :*

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

*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  $\varepsilon, d$  and  $T$  such that for every partition  $\mathcal{P}$  of  $[0, T]$  with  $|\mathcal{P}| \leq 1$ , when  $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$ ,*

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

*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

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

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

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

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  $\Delta$ . We remark that  $\Delta$  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  $I_1 = B_1$ ,  $I_k = B_k \setminus B_{k-1}$  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:

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

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

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

and recall the notation  $\theta$  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  $\varepsilon, d$  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  $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$ ,*

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

*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  $\rho_{\tau(r)}^{\mathcal{P},M}$  and  $\rho_{\tau(T_k^k)-}^{\mathcal{P},M}$  directly in (26), so we construct the particle system as follows in order to obtain an explicit scheme of simulation. For a random vector  $X = (X^1, \dots, X^N)$ ,  $X^i \in \mathbb{R}^d$ ,  $i = 1, \dots, N$  with a fixed dimension  $N$ , we associate the (random) empirical measure

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

where  $\delta_x(dv)$  is the Dirac measure. Now we consider an initial vector  $(X_0^1, \dots, X_0^N)$  with components which are independent and identically distributed with common law  $\rho_0$  (we recall that  $\rho_0$  is the law of  $X_0$  in (14)), and we consider  $(\Delta^1, \dots, \Delta^N)$  which is a  $N \times d$ -dimensional standard Gaussian random variable independent of  $(X_0^1, \dots, X_0^N)$ . Then we construct the particle system  $\vec{X}_t^{\mathcal{P},M} = (X_t^{\mathcal{P},M,1}, \dots, X_t^{\mathcal{P},M,N})$ :

$$\begin{aligned} X_t^{\mathcal{P},M,i} &= X_0^i + a_T^M \Delta^i + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(r)}^{\mathcal{P},M})) dr \\ &+ \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}^d} c_M(\tau(r), v, z, X_{\tau(r)-}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(r)-}^{\mathcal{P},M})) N_{\hat{\rho}(\vec{X}_{\tau(r)-}^{\mathcal{P},M})}^i(dv, dz, dr), \quad i = 1, \dots, N, \end{aligned} \quad (30)$$

where  $N_{\hat{\rho}(\vec{X}_t^{\mathcal{P},M})}^i(dv, dz, dr)$ ,  $i = 1, \dots, N$  are Poisson point measures that are independent each other conditionally to  $\vec{X}_t^{\mathcal{P},M}$  and independent of  $(X_0^1, \dots, X_0^N, \Delta^1, \dots, \Delta^N)$  with intensity  $\hat{\rho}(\vec{X}_t^{\mathcal{P},M})(dv)\mu(dz)dr$ . 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  $I_1 = B_1$ ,  $I_k = B_k \setminus B_{k-1}$  for  $k \geq 2$ . Now for  $i = 1, \dots, N$ ,  $k \in \mathbb{N}$ , we take  $(J_t^{k,i})_{t \in [0,T]}$  a Poisson process of intensity  $\mu(I_k)$ . We denote by  $(T_l^{k,i})_{l \in \mathbb{N}}$  the jump times of  $(J_t^{k,i})_{t \in [0,T]}$  and we consider some sequences of independent random variables  $Z_l^{k,i} \sim \mathbb{1}_{I_k}(z) \frac{\mu(dz)}{\mu(I_k)}$  and  $U_l^{k,i}$  uniformly distributed on  $\{1, \dots, N\}$ , for all  $i = 1, \dots, N$ ,  $k, l \in \mathbb{N}$ . Moreover,  $(J_t^{k,i})_{t \in [0,T]}$ ,  $Z_l^{k,i}$ ,  $U_l^{k,i}$ ,  $\Delta^i$ ,  $X_0^i$ ,  $i = 1, \dots, N$ ,  $k, l \in \mathbb{N}$  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.

$$\begin{aligned} X_t^{\mathcal{P},M,i} &= X_0^i + a_T^M \Delta^i + \int_0^t b(\tau(r), X_{\tau(r)}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(r)}^{\mathcal{P},M})) dr \\ &+ \sum_{k=1}^M \sum_{l=1}^{J_t^{k,i}} c(\tau(T_l^{k,i}), X_{\tau(T_l^{k,i})-}^{\mathcal{P},M,U_l^{k,i}}, Z_l^{k,i}, X_{\tau(T_l^{k,i})-}^{\mathcal{P},M,i}, \hat{\rho}(\vec{X}_{\tau(T_l^{k,i})-}^{\mathcal{P},M})). \end{aligned} \quad (31)$$

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

We denote

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

and we consider the  $d$ -dimensional regularization kernels

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

We recall the notations  $\varepsilon_M$  in (27) and  $\theta$  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

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

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

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

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

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

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

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

**Theorem 2.4.** We suppose **Hypothesis 2.1**, **Hypothesis 2.2**, **Hypothesis 2.3** and **Hypothesis 2.4** hold true. We take supplementally  $\tilde{\Delta}$  to be a  $d$ -dimensional standard Gaussian random variable independent of  $\vec{X}_t^{\mathcal{P}, M}$ . 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

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

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

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

ii) (**Romberg**) We take

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

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

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

*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  $\delta, N$ , 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  $\delta$  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  $\mu$  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

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

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

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

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

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

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  $\sigma'$  is the differential of  $\sigma$  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,  $\underline{c}(z) = \underline{\sigma} \times \frac{1}{|z|^{\alpha}}$ , then

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

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  $a_T^M$ ,  $\Delta$  and  $c_M$  defined in Section 2.4).

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

We notice that we keep  $\rho_r$  (the law of  $X_r$ ) instead of taking  $\rho_r^M$  (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

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

where  $\Pi(\rho_1, \rho_2)$  is the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  with marginals  $\rho_1$  and  $\rho_2$ . 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  $\rho_{t-}$  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  $\rho_{t-}$  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  $\Phi : [0, 1] \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}^d$  such that for any  $\rho \in \mathcal{P}_1(\mathbb{R}^d)$ , any bounded and measurable function  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ , 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  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

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

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  $d\omega\mu(dz)dr$ . 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  $\bar{\Phi}(r, w, z, \omega, \rho) : [0, T] \times [0, 1] \times \mathbb{R}^d \times \Omega \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}_+$  and  $\bar{\varphi}(r, w, \omega, \rho) : [0, T] \times [0, 1] \times \Omega \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow \mathbb{R}_+$  be two non-negative functions.

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

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

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  $p \geq 2$ ,

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

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  $\varepsilon_*$ , for every  $M$  such that  $\varepsilon_M \leq 1$  and  $|\bar{c}(z)|^2 \mathbb{1}_{\{|z|>M\}} \leq 1$ , we have*

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

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

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

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

Firstly, by **Hypothesis 2.1**,

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

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

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

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

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

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

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

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

Using **Hypothesis 2.1**,

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

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

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

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

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

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.  $\square$

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  $((J_t^k)_{t \in [0, T]}, (Z_i^k)_{k \in \mathbb{N}}, X_0, \Delta)$  constructed in Section 2.4, we take moreover  $(W_i^k)_{k \in \mathbb{N}}$  a sequence of independent random variables which are uniformly distributed on  $[0, 1]$  and independent of  $((J_t^k)_{t \in [0, T]}, (Z_i^k)_{k \in \mathbb{N}}, X_0, \Delta)$ . Then we have

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

$$x_t^{\mathcal{P}, M} = X_0 + a_T^M \Delta + \int_0^t b(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M}) dr + \sum_{k=1}^M \sum_{i=1}^{J_t^k} c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau(T_i^k)-}^{\mathcal{P}, M}). \quad (52)$$

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  $C_p^\infty$  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  $(\Omega, \mathcal{F}, \mathbb{P})$ , 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  $D : \mathcal{S} \rightarrow \bigcap_{p=1}^\infty L^p(\Omega; \mathcal{H})$  which is a linear application which satisfies

a)

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

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

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

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

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

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

Then

$$DF = \sum_{i=1}^{\infty} D_i F \times e_i \quad \text{and} \quad D^q F = \sum_{i_1, \dots, i_q} D_{i_1, \dots, i_q} F \times \otimes_{j=1}^q e_{i_j}.$$

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

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

And we denote

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

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  $L : \mathcal{S} \rightarrow \mathcal{S}$  which is a linear operator satisfying

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

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

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

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

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

**Definition 3.1.** If  $D^q : \mathcal{S} \subset L^2(\Omega) \rightarrow L^2(\Omega; \mathcal{H}^{\otimes q}), \forall q \geq 1$ , 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  $l \geq 1, F \in \mathcal{S}$ ,

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

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

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

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  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $C^l(\mathbb{R}^d)$  function ( $l$ -times differentiable function), then there is a constant  $C_l$  dependent on  $l$  such that

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

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

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

For  $l = 0$ , we have

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

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

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

and

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

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  $\|F - F_n\|_p \rightarrow 0$ ,  $\|F_m - F_n\|_{q,p} \rightarrow 0$  and  $\|LF_m - LF_n\|_{q-2,p} \rightarrow 0$ . Since  $D^q$  and  $L$  are closable, we can define

$$D^q F = \lim_{n \rightarrow \infty} D^q F_n \quad \text{in } L^p(\Omega; \mathcal{H}^{\otimes q}), \quad LF = \lim_{n \rightarrow \infty} LF_n \quad \text{in } L^p(\Omega). \quad (62)$$

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.  $\square$

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

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

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

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

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

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

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

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

the lowest eigenvalue of the covariance matrix  $\sigma_F$ . 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  $p \geq 1$ ,

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

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

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

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  $F_n \rightarrow G$  weakly in the Hilbert space  $\mathcal{H}_l$ . This means that for every  $Q \in \mathcal{H}_l$ ,  $\langle F_n, Q \rangle_{L,l,2} \rightarrow \langle G, Q \rangle_{L,l,2}$ . 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  $\gamma_i^n \geq 0, i = n, \dots, m_n$  and  $\sum_{i=n}^{m_n} \gamma_i^n = 1$ , such that

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

In particular we have

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

Also, we notice that by *i*),

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

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

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

By passing to a subsequence, we have  $|G_n - F|_l + |LG_n - LF|_{l-2} \rightarrow 0$  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,  $Y_n \rightarrow Y$  almost surely, and for any  $\tilde{q} \in [\bar{p}, p]$ ,

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

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

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

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

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

so

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

Now we check that

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

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  $G_n = \sum_{i=n}^{m_n} \gamma_i^n \times F_i$ , we observe that

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

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

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

So we conclude that  $\|DF - DF_n\|_{L^2(\Omega; \mathcal{H})} \rightarrow 0$ . Thus, by (65),  $\mathbb{E}|\sqrt{\lambda(F)} - \sqrt{\lambda(F_n)}| \rightarrow 0$ . 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

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

**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

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

Then it follows by (64) that

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

□



### 3.1.1 Main consequences

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

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

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

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

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

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

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

with

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

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

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

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

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

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

Then for every  $f \in C_p^1(\mathbb{R}^d)$  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

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

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

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

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

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  $F = (F_1, \dots, F_d) \in \mathcal{D}_\infty^d$ . 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  $\beta$  with  $|\beta| = q$  and any function  $f \in C_b^q(\mathbb{R}^d)$ ,

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

(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  $\beta$  with  $|\beta| = q$  and any function  $f \in C_b^q(\mathbb{R}^d)$ ,

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

*Remark.* In (A), we assume the non-degeneracy condition for  $F$ , so we can control the weight  $H_\beta$  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  $F = (F_1, \dots, F_d) \in \mathcal{D}_\infty^d$ . 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

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

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

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

We consider the  $d$ -dimensional regularization kernels

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

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  $\beta$ , we suppose that  $F$  satisfies  $(\mathbf{B}_{2+|\beta|})$ . Then for any function  $f \in C_b^{2+|\beta|}(\mathbb{R}^d)$ ,

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

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

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

(B) *iii)* We suppose that  $F$  satisfies  $(\mathbf{B}'_2)$ . We fix  $\rho > 0$  and we take some  $G \in \mathcal{D}_\infty^d$  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  $p, q, \rho$  and  $d$  such that for any  $\eta > 0$  and  $\delta > 0$ , for any function  $f \in C_b^2(\mathbb{R}^d)$ ,

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

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

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

*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  $\varphi_\delta$  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  $\alpha = (\alpha_1, \dots, \alpha_q)$ . Notice that if  $F$  satisfies (B<sub>q</sub>) then (recall that  $\partial^\alpha g = \partial^\alpha \partial^\beta f$ )

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

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  $\delta$  and for  $a\delta$ , we obtain

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

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 (B'<sub>q</sub>), then

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

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

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

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

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

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  $\eta$ , we write

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

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

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

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

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

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

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

Then we have

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

Similarly, we also have

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

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  $\delta$  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),

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

We conclude together with (80) and (81).  $\square$

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  $(\mathbf{B}_{2+d})$ . Then, for every  $x$ ,*

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

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

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

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

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

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

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

In a similar way (70) gives (84).  $\square$

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

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

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  $F, G \in \mathcal{D}_\infty^d$ . One may find  $p \in \mathbb{N}$ ,  $C \in \mathbb{R}_+$  (depending on  $r, l$  and  $\varepsilon$ ) such that

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

and

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

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  $F, G \in \mathcal{D}_\infty^d$ . One may find  $p \in \mathbb{N}$ ,  $C \in \mathbb{R}_+$  (depending on  $r$  and  $\varepsilon$ ) such that

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

*Remark.* The result in **Proposition 3.6.1** is better than proposition 3.11 and remark 3.14 in [7]. We get  $\|DF - DG\|_{L^2(\Omega; \mathcal{H})}^2$  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 2.1~2.4**.

## 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.

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

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

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

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

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

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

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  $Z_i^k = (Z_{i,1}^k, \dots, Z_{i,d}^k), k, i \in \mathbb{N}$  and  $\Delta = (\Delta_1, \dots, \Delta_d)$ . We denote by  $C_{\mathcal{G},p}$  the space of the functions  $f : \Omega \times \mathbb{R}^{m \times m' \times d+d} \rightarrow \mathbb{R}$  such that for each  $\omega$ , the function  $(z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d) \mapsto f(\omega, z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$  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  $(z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$ , the function  $\omega \mapsto f(\omega, z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$  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

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

*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  $0 < t \leq T$ ,  $x_t^{\mathcal{P},M}$  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  $DF = (D^Z F, D^\Delta F)$ , where

$$\begin{aligned} D_{(\bar{k}, \bar{i}, \bar{j})}^Z F &= \xi_{\bar{i}}^{\bar{k}} \frac{\partial f}{\partial z_{\bar{i}, \bar{j}}^{\bar{k}}}(\omega, (Z_i^k)_{\substack{1 \leq k \leq m' \\ 1 \leq i \leq m}}, \Delta), \quad \bar{k}, \bar{i} \in \mathbb{N}, \bar{j} \in \{1, \dots, d\}, \\ D_{\bar{j}}^\Delta F &= \frac{\partial f}{\partial \delta_{\bar{j}}}(\omega, (Z_i^k)_{\substack{1 \leq k \leq m' \\ 1 \leq i \leq m}}, \Delta), \quad \bar{j} \in \{1, \dots, d\}. \end{aligned} \quad (93)$$

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

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

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

We recall the function  $h$  given in **Hypothesis 2.4 b)**. We also define the Ornstein-Uhlenbeck operator  $LF = L^Z F + L^\Delta F$  with

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

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  $p \geq 1, l \geq 0$ , there exists a constant  $C_{l,p}$  depending on  $l, p, d$  and  $T$ , such that for any  $0 < t \leq T$ ,*

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

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

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

**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  $\theta$  defined in (18)), we have (recalling by (55) that  $\sigma_F$  denotes the covariance matrix of  $F$ )

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

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

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  $0 < t \leq T$ .

*Remark.* We are not able to prove that  $\sup_{\mathcal{P}} \sup_M \mathbb{E}(1/\det \sigma_{x_t^{\mathcal{P},M}})^p \leq C_p$ , since the tangent flow of the Euler scheme is not invertible (see (130), the inverse tangent flow plays an important role in our proof). This is why we need **Proposition 3.6.1** instead of **Lemma 3.6**. Fortunately, thanks to (79) (inspired from [11] p14), we are able to keep the same speed of convergence  $1 - \varepsilon$  as if the tangent flow of the Euler scheme were invertible.

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  $T, d, \varepsilon_*$  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  $\tilde{\Pi}_t^{\mathcal{P},M}(dv_5, dv_6)$  to be the optimal  $W_1$ -coupling of  $\rho_{\tau(t)-}^{\mathcal{P},M}(dv_5)$  and  $\rho_{t-}(dv_6)$ , that is

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

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

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

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):

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



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

with  $\rho_t^{\mathcal{P},M}$  the law of  $X_t^{\mathcal{P},M}$  (see (25)) and  $\rho_t$  the law of  $X_t$  (see (14)). One can check that  $\tilde{x}_t$  and  $\tilde{x}_t^{\mathcal{P},M}$  have the same law as  $x_t$  and  $X_t^{\mathcal{P},M}$  respectively. We remark that  $\tilde{x}_t$ , and  $\tilde{x}_t^{\mathcal{P},M}$  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

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

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

### 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  $\theta$  defined in (18)),

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

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

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

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  $\Delta$ , 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  $M \rightarrow \infty$ , we give below in **Lemma 3.11** that  $\bar{x}_t^{\mathcal{P},M} \rightarrow \bar{x}_t^{\mathcal{P}}$  and  $\bar{x}_t^M \rightarrow \bar{x}_t$  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.

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

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):

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

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

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

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

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  $\rho_r$  instead of  $\rho_r^M$  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

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

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

as  $M \rightarrow \infty$ .

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

We notice that  $\bar{x}_t^{\mathcal{P},M}$  and  $\bar{x}_t^M$  are simple functionals (belong to  $S^d$ ) 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  $\varepsilon_*, d, T$  such that for every  $|\mathcal{P}| \leq 1$ , we have

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

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

*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.  $\square$

Then applying **Lemma 3.3**, by passing to the limit  $M \rightarrow \infty$ , 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  $\varepsilon_*, d, T$  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  $\|D\bar{x}_t^{\mathcal{P}} - D\bar{x}_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \leq C|\mathcal{P}|^{\frac{1}{2+\varepsilon_*}}$ .  $\square$

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  $t > \frac{8d}{\bar{\varepsilon}}(\frac{8}{\bar{\varepsilon}} + 1)$ , we have

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

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

### 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  $t > \frac{8d}{\bar{\varepsilon}}(2+d)$ , **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

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

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

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

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

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

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  $\delta$  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  $\varepsilon, d, T$  such that when  $t > \frac{8d}{\theta}(\frac{8}{\varepsilon} + 1)$ ,

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

with  $\rho_t$  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

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

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  $\delta, \eta$  and  $\rho$ . In order to keep the notations clear, we denote temporarily 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

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

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

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

*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

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

Now we optimize over  $N$ . We take  $N$  such that

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

so

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

Then we have

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

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

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

where  $\tilde{\Delta}$  is a  $d$ -dimensional standard Gaussian random variable independent of  $X_t^{\mathcal{P},M,i}, i = 1, \dots, N$ , 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  $\mathbb{E}|x_t^M - x_t| \rightarrow 0$  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

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

and

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

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

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

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

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

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

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

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

$$LZ_{i,j}^k = -\partial_{z_{i,j}^k} (\xi_i^k)^2 - \xi_i^k D_{(k,i,j)}^Z \ln[h(Z_i^k)].$$

We observe that  $|\xi_i^k| \leq 1$ , and we have  $|\partial_{z_{i,j}^k} (\xi_i^k)^2| = 2\Psi_k(Z_i^k)\partial_{z_{i,j}^k} \Psi_k(Z_i^k)$  is bounded by a universal constant (see (92)). These lead to

$$|LZ_{i,j}^k|_l \leq C_l(1 + |D_{(k,i,j)}^Z \ln[h(Z_i^k)]|_l).$$

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),

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

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

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

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

and by **Hypothesis 2.1** that for  $l = 0$ ,  $\mathbb{E} \sup_{0 < t \leq T} |x_t^{\mathcal{P},M}|^p \leq C_{0,p}$ . 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  $D_j^\Delta \Delta = e_j$ , where  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  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  $\mathbb{E} \sup_{0 < t \leq T} |x_t^{\mathcal{P}, M}|_{1, l}^p \leq C_{l, p}(1 + A_1 + A_2)$ , with

$$\begin{aligned} A_1 &= \mathbb{E} \int_0^T |b(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M})|_{1, l}^p dr, \\ A_2 &= \mathbb{E} \left( \sum_{k=1}^M \sum_{i=1}^{J_T^k} |c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau(T_i^k)-}^{\mathcal{P}, M})|_{1, l}^p \right). \end{aligned}$$

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

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

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

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

Hence, using (45) with

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

**Hypothesis 2.1** and the recurrence hypothesis, it follows that

$$\begin{aligned} A_2 &\leq C_{l, p} \mathbb{E} \left[ \sum_{k=1}^M \sum_{i=1}^{J_T^k} \bar{c}(Z_i^k) (1 + |x_{\tau(T_i^k)-}^{\mathcal{P}, M}|_{1, l} + |x_{\tau(T_i^k)-}^{\mathcal{P}, M}|_{1, l-1}^l)^p \right] \\ &\leq C_{l, p} \mathbb{E} \left[ \int_0^T \int_{[0, 1] \times \mathbb{R}^d} \bar{c}(z) (1 + |x_{\tau(r)-}^{\mathcal{P}, M}|_{1, l} + |x_{\tau(r)-}^{\mathcal{P}, M}|_{1, l-1}^l) \mathcal{N}(dw, dz, dr) \right]^p \\ &\leq C_{l, p} [1 + \int_0^T \mathbb{E} |x_{\tau(r)-}^{\mathcal{P}, M}|_{1, l}^p dr + \int_0^T \mathbb{E} |x_{\tau(r)-}^{\mathcal{P}, M}|_{1, l-1}^p dr] \\ &\leq C_{l, p} [1 + \int_0^T \mathbb{E} |x_{\tau(r)}^{\mathcal{P}, M}|_{1, l}^p dr]. \end{aligned} \tag{121}$$

Combining (120) and (121), one has

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



So we conclude by Gronwall lemma that

$$\sup_{\mathcal{P}, M} \mathbb{E} \sup_{0 < t \leq T} |x_t^{\mathcal{P}, M}|_{1,1}^p \leq C_{l,p}. \quad (122)$$

**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  $\Delta$  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

$$\begin{aligned} S_1 &= \mathbb{E} \int_0^T |Lb(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M})|^p dr, \\ S_2 &= \mathbb{E} \left( \sum_{k=1}^M \sum_{i=1}^{J_T^k} |Lc(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau(T_i^k)-}^{\mathcal{P}, M})| \right)^p. \end{aligned}$$

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

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

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

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

Therefore, using (45) with

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

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

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

Combining (123) and (124), one has

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

Applying Gronwall lemma, we obtain

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

**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  $L\Delta = \Delta$ ,  $D_j^\Delta \Delta = e_j$ , where  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  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  $\Delta$  has finite moments of any order, we write  $\mathbb{E} \sup_{0 \leq t \leq T} |Lx_t^{\mathcal{P}, M}|^p \leq C_{l,p}(1 + B_1 + B_2)$ , with

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

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

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

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

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

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

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

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

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

□

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  $\mathbb{E}|x_t^M - x_t| \rightarrow 0$  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  $k_0, i_0 \in \mathbb{N}, j \in \{1, \dots, d\}$ ,

$$\begin{aligned} D_{(k_0, i_0, j)}^Z x_t^M &= \int_{T_{i_0}^{k_0}}^t \nabla_x b(r, x_r^M, \rho_r) D_{(k_0, i_0, j)}^Z x_r^M dr \\ &+ \mathbb{1}_{\{0 < T_{i_0}^{k_0} \leq t\}} \mathbb{1}_{\{1 \leq k_0 \leq M\}} \xi_{i_0}^{k_0} \partial_{z_j} c(T_{i_0}^{k_0}, \eta_{T_{i_0}^{k_0}}^2(W_{i_0}^{k_0}), Z_{i_0}^{k_0}, x_{T_{i_0}^{k_0}-}^M, \rho_{T_{i_0}^{k_0}-}) \\ &+ \sum_{k=1}^M \sum_{T_{i_0}^{k_0} < T_i^k \leq t} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}) D_{(k_0, i_0, j)}^Z x_{T_i^k-}^M, \end{aligned} \quad (128)$$

$$D_j^\Delta x_t^M = a_T^M e_j + \int_0^t \nabla_x b(r, x_r^M, \rho_r) D_j^\Delta x_r^M dr + \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}) D_j^\Delta x_{T_i^k-}^M, \quad (129)$$

where  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  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

$$Y_t^M = I_d + \int_0^t \nabla_x b(r, x_r^M, \rho_r) Y_r^M dr + \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}) Y_{T_i^k-}^M.$$

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

$$\tilde{Y}_t^M = I_d - \int_0^t \tilde{Y}_r^M \nabla_x b(r, x_r^M, \rho_r) dr - \sum_{k=1}^M \sum_{i=1}^{J_t^k} \tilde{Y}_{T_i^k-}^M \nabla_x c(I_d + \nabla_x c)^{-1}(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}). \quad (130)$$

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

$$\mathbb{E} \left( \sup_{0 < t \leq T} (\|Y_t^M\|^p + \|\tilde{Y}_t^M\|^p) \right) \leq C_p < \infty. \quad (131)$$

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

$$D_{(k,i,j)}^Z x_t^M = \mathbb{1}_{\{0 < T_i^k \leq t\}} \mathbb{1}_{\{1 \leq k \leq M\}} \xi_i^k Y_t^M \tilde{Y}_{T_i^k}^M \partial_{z_j} c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}), \quad (132)$$

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

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

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

By (132),

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

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

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

We denote

$$\chi_t^M = \sum_{k=1}^M \sum_{i=1}^{J_t^k} |\xi_i^k|^2 \underline{c}(Z_i^k). \quad (133)$$

By (131),  $(\mathbb{E} \sup_{0 \leq t \leq T} \|\tilde{Y}_t^M\|^{4dp} \|Y_t^M\|^{4dp})^{1/2} \leq C_{d,p} < \infty$ , so that using Schwartz inequality, we have

$$\mathbb{E} \left| \frac{1}{\det \sigma_{x_t^M}} \right|^p \leq \mathbb{E} (|\lambda_t^M|^{-dp}) \leq C (\mathbb{E} (|\chi_t^M + |a_T^M|^2|^{-2dp}))^{\frac{1}{2}}. \quad (134)$$

**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

$$\mathbb{E} \left( \frac{1}{|\chi_t^M + |a_T^M|^2|^{2dp}} \right) = \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \mathbb{E} (e^{-s(\chi_t^M + |a_T^M|^2)}) ds. \quad (135)$$

**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  $\xi_t^k = \Psi_k(Z_t^k)$  (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  $\Theta_M(dz, dr) = \sum_{k=1}^M \Lambda_k(dz, dr)$ . 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  $\Psi(z) = \sum_{k=1}^{\infty} |\Psi_k(z)|^2 \mathbb{1}_{I_k}(z)$ . 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  $\Theta$  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

$$\begin{aligned}
\mathbb{E}(e^{-s(\chi_t^M + |a_T^M|^2)}) &= \mathbb{E}(e^{-s\chi_t^M}) \times e^{-s|a_T^M|^2} \\
&\leq \exp(-t \int_{B_M} (1 - e^{-s\varrho(z)})\nu(dz)) \times \exp(-t \int_{B_M^c} (1 - e^{-s\varrho(z)})\nu(dz)) \\
&= \exp(-t \int_{\mathbb{R}^d} (1 - e^{-s\varrho(z)})\nu(dz)),
\end{aligned} \tag{136}$$

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

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

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

$$\lim_{u \rightarrow \infty} \frac{1}{\ln u} \mathcal{M}(f \geq \frac{1}{u}) > \frac{p}{t}, \tag{137}$$

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

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

$$\lim_{u \rightarrow \infty} \frac{1}{\ln u} \nu(\varrho \geq \frac{1}{u}) = \theta > 0. \tag{138}$$

Then for every  $p \geq 1$ ,  $0 < t \leq T$ , when  $\theta > \frac{2dp}{t}$  (i.e.  $t > \frac{2dp}{\theta}$ ), we deduce from (134), (135), (136) and **Lemma 4.3** that

$$\begin{aligned}
\sup_M \mathbb{E} \left| \frac{1}{\det \sigma_{x_t^M}} \right|^p &\leq \sup_M \mathbb{E}(|\lambda_t^M|^{-dp}) \leq C \sup_M (\mathbb{E}(|\chi_t^M + |a_T^M|^2|^{-2dp}))^{\frac{1}{2}} \\
&\leq C \left( \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \mathbb{E}(e^{-s(\chi_t^M + |a_T^M|^2)}) ds \right)^{\frac{1}{2}} \\
&\leq C \left( \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \exp(-t \int_{\mathbb{R}^d} (1 - e^{-s\varrho(z)})\nu(dz) ds) \right)^{\frac{1}{2}} < \infty.
\end{aligned} \tag{139}$$

**Proof of ii):** We recall in **Lemma 2.6 i)** that  $\mathbb{E}|x_t^M - x_t| \rightarrow 0$ , 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).  $\square$

## 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  $\mathcal{P} = \{0 = r_0 < r_1 < \dots < r_{n-1} < r_n = T\}$ ,  $M \in \mathbb{N}$ , any  $k_0, i_0 \in \mathbb{N}, j \in \{1, \dots, d\}$ ,

$$\begin{aligned} D_{(k_0, i_0, j)}^Z x_t^{\mathcal{P}, M} &= \int_0^t \nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M}) D_{(k_0, i_0, j)}^Z x_{\tau(r)}^{\mathcal{P}, M} dr \\ &+ \mathbb{1}_{\{0 < T_{i_0}^{k_0} \leq t\}} \mathbb{1}_{\{1 \leq k_0 \leq M\}} \xi_{i_0}^{k_0} \partial_{z_j} c(\tau(T_{i_0}^{k_0}), \eta_{T_{i_0}^{k_0}}^1(W_{i_0}^{k_0}), Z_{i_0}^{k_0}, x_{\tau^{\mathcal{P}}(T_{i_0}^{k_0})-}^{\mathcal{P}, M}, \rho_{\tau^{\mathcal{P}}(T_{i_0}^{k_0})-}^{\mathcal{P}, M}) \\ &+ \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}) D_{(k_0, i_0, j)}^Z x_{\tau(T_i^k)-}^{\mathcal{P}, M}, \end{aligned} \quad (140)$$

$$\begin{aligned} D_j^\Delta x_t^{\mathcal{P}, M} &= a_T^M e_j + \int_0^t \nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M}) D_j^\Delta x_{\tau(r)}^{\mathcal{P}, M} dr \\ &+ \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}) D_j^\Delta x_{\tau(T_i^k)-}^{\mathcal{P}, M}, \end{aligned} \quad (141)$$

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

$$\begin{aligned} D_{(k_0, i_0, j)}^Z x_t^M &= \int_0^t \nabla_x b(r, x_r^M, \rho_r) D_{(k_0, i_0, j)}^Z x_r^M dr \\ &+ \mathbb{1}_{\{0 < T_{i_0}^{k_0} \leq t\}} \mathbb{1}_{\{1 \leq k_0 \leq M\}} \xi_{i_0}^{k_0} \partial_{z_j} c(T_{i_0}^{k_0}, \eta_{T_{i_0}^{k_0}}^2(W_{i_0}^{k_0}), Z_{i_0}^{k_0}, x_{T_{i_0}^{k_0}-}^M, \rho_{T_{i_0}^{k_0}-}^M) \\ &+ \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}^M) D_{(k_0, i_0, j)}^Z x_{T_i^k-}^M, \end{aligned} \quad (142)$$

$$D_j^\Delta x_t^M = a_T^M e_j + \int_0^t \nabla_x b(r, x_r^M, \rho_r) D_j^\Delta x_r^M dr + \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}^M) D_j^\Delta x_{T_i^k-}^M. \quad (143)$$

For  $u \in l_2$ , we will use the notation  $|u|_{l_2}^2 = |u|_{(\bullet, \circ, \diamond)}^2 = \sum_{k=1}^\infty \sum_{i=1}^\infty \sum_{j=1}^d |u_{(k, i, j)}|^2$ . We write  $\mathbb{E}|D_{(\bullet, \circ, \diamond)}^Z x_t^{\mathcal{P}, M} - D_{(\bullet, \circ, \diamond)}^Z x_t^M|_{l_2}^2 \leq C[H_1 + H_2 + H_3]$ , with

$$H_1 = \mathbb{E} \left| \int_0^t \nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P}, M}, \rho_{\tau(r)}^{\mathcal{P}, M}) D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P}, M} dr - \int_0^t \nabla_x b(r, x_r^M, \rho_r) D_{(\bullet, \circ, \diamond)}^Z x_r^M dr \right|_{l_2}^2,$$

$$\begin{aligned} H_2 &= \mathbb{E} \left| \mathbb{1}_{\{0 < T_\bullet \leq t\}} \mathbb{1}_{\{1 \leq \bullet \leq M\}} (\partial_{z_\circ} c(\tau(T_\bullet^\circ), \eta_{T_\bullet^\circ}^1(W_\bullet^\circ), Z_\bullet^\circ, x_{\tau^{\mathcal{P}}(T_\bullet^\circ)-}^{\mathcal{P}, M}, \rho_{\tau^{\mathcal{P}}(T_\bullet^\circ)-}^{\mathcal{P}, M}) \right. \\ &\quad \left. - \partial_{z_\circ} c(T_\bullet^\circ, \eta_{T_\bullet^\circ}^2(W_\bullet^\circ), Z_\bullet^\circ, x_{T_\bullet^\circ-}^M, \rho_{T_\bullet^\circ-}^M)) \right|_{l_2}^2, \end{aligned}$$

$$\begin{aligned} H_3 &= \mathbb{E} \left| \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P}, M}) D_{(\bullet, \circ, \diamond)}^Z x_{\tau(T_i^k)-}^{\mathcal{P}, M} \right. \\ &\quad \left. - \sum_{k=1}^M \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}^M) D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^M \right|_{l_2}^2. \end{aligned}$$

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

$$\begin{aligned}
H_1 &\leq C[\mathbb{E} \int_0^t |\nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M}) - \nabla_x b(r, x_r^M, \rho_r)|^2 |D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr \\
&\quad + \mathbb{E} \int_0^t |\nabla_x b(\tau(r), x_{\tau(r)}^{\mathcal{P},M}, \rho_{\tau(r)}^{\mathcal{P},M})|^2 |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr] \\
&\leq C[\mathbb{E} \int_0^t [|\mathcal{P}|^2 + |x_{\tau(r)}^{\mathcal{P},M} - x_r^M|^2 + (W_1(\rho_{\tau(r)}^{\mathcal{P},M}, \rho_r))^2] |D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr \\
&\quad + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr].
\end{aligned}$$

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

$$\begin{aligned}
H_1 &\leq C[|\mathcal{P}|^2 + \int_0^t (\mathbb{E} |x_{\tau(r)}^{\mathcal{P},M} - x_r^M|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (W_{2+\varepsilon_*}(\rho_{\tau(r)}^{\mathcal{P},M}, \rho_r))^2 dr \\
&\quad + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr] \\
&\leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr]. \tag{144}
\end{aligned}$$

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

$$\begin{aligned}
H_2 &= \mathbb{E} \sum_{k=1}^M \sum_{i=1}^{J_t^k} \sum_{j=1}^d |\partial_{z_j} c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}) - \partial_{z_j} c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-})|^2 \\
&\leq C \mathbb{E} \sum_{k=1}^M \sum_{i=1}^{J_t^k} [|\bar{c}(Z_i^k)|^2 [|\tau^{\mathcal{P}}(T_i^k) - T_i^k| + |\eta_{T_i^k}^1(W_i^k) - \eta_{T_i^k}^2(W_i^k)| + |x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M} - x_{T_i^k-}^M| + W_1(\rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}, \rho_{T_i^k-})]^2 \\
&\leq C \mathbb{E} \int_0^t \int_{[0,1] \times \mathbb{R}^d} |\bar{c}(z)|^2 [|\tau^{\mathcal{P}}(r) - r|^2 + |\eta_r^1(w) - \eta_r^2(w)|^2 \\
&\quad + |x_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M} - x_{r-}^M|^2 + (W_1(\rho_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M}, \rho_{r-}))^2] \mathcal{N}(dw, dz, dr) \\
&= C \mathbb{E} \int_0^t \int_{[0,1] \times \mathbb{R}^d} |\bar{c}(z)|^2 [|\tau^{\mathcal{P}}(r) - r|^2 + |\eta_r^1(w) - \eta_r^2(w)|^2 \\
&\quad + |x_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M} - x_{r-}^M|^2 + (W_1(\rho_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M}, \rho_{r-}))^2] dw \mu(dz) dr.
\end{aligned}$$

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

$$\begin{aligned}
H_2 &\leq C[|\mathcal{P}|^2 + \int_0^t (\mathbb{E} |x_{\tau(r)}^{\mathcal{P},M} - x_r^M|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (\mathbb{E} |x_{\tau(r)}^{\mathcal{P},M} - x_r^M|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t (W_{2+\varepsilon_*}(\rho_{\tau(r)}^{\mathcal{P},M}, \rho_r))^2 dr] \\
&\leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}]. \tag{145}
\end{aligned}$$

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

$$\begin{aligned}
H_{3,1} &= \mathbb{E} \left( \sum_{k=1}^M \sum_{i=1}^{J_t^k} |\nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}) \right. \\
&\quad \left. - \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^M, \rho_{T_i^k-}) ||D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^M|_{l_2}^2 \right), \\
H_{3,2} &= \mathbb{E} \left( \sum_{k=1}^M \sum_{i=1}^{J_t^k} |\nabla_x c(\tau(T_i^k), \eta_{T_i^k}^1(W_i^k), Z_i^k, x_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}, \rho_{\tau^{\mathcal{P}}(T_i^k)-}^{\mathcal{P},M}) ||D_{(\bullet, \circ, \diamond)}^Z x_{\tau(T_i^k)-}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^M|_{l_2}^2 \right).
\end{aligned}$$



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

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

we get

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

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

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

where the last inequality is a consequence of **Lemma 2.6** and (50).

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

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

we have

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

Therefore,

$$H_3 \leq C[H_{3,1} + H_{3,2}] \leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr]. \quad (146)$$

Combining (144), (145) and (146),

$$\mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_t^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_t^M|_{l_2}^2 \leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(r)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr].$$

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

$$\bar{\Phi}(r, w, z, \omega, \rho) = \nabla_x c(\tau(r), \eta_r^1(w), z, x_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M}, \rho_{\tau^{\mathcal{P}}(r)-}^{\mathcal{P},M}) D_{(k_0, i_0, j)}^Z x_{\tau(r)-}^{\mathcal{P},M}$$

that

$$\mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_{\tau(t)}^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_t^{\mathcal{P},M}|_{l_2}^2 \leq C|\mathcal{P}|, \quad (147)$$

so

$$\mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_t^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_t^M|_{l_2}^2 \leq C[(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_r^{\mathcal{P},M} - D_{(\bullet, \circ, \diamond)}^Z x_r^M|_{l_2}^2 dr].$$

We conclude by Gronwall lemma that  $\mathbb{E}|D_{(\bullet, \circ, \diamond)}^Z x_t^{P, M} - D_{(\bullet, \circ, \diamond)}^Z x_t^M|_{l_2}^2 \leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}$ . Finally, by a similar argument,  $\mathbb{E}|D_{(\diamond)}^\Delta x_t^{P, M} - D_{(\diamond)}^\Delta x_t^M|_{\mathbb{R}^d}^2 \leq C(|\mathcal{P}| + \varepsilon_M)^{\frac{2}{2+\varepsilon_*}}$ , 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

$$\|Dx_t^{M_1} - Dx_t^{M_2}\|_{L^2(\Omega; l^2 \times \mathbb{R}^d)} \leq (\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{1}{2+\varepsilon_*}}. \quad (148)$$

In fact, if  $(Dx_t^M)_{M \in \mathbb{N}}$  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  $F_M = X_t^M$  and  $F = X_t$ , we know that there exists a convex combination  $\sum_{M'=M}^{m_M} \gamma_{M'}^M \times F_t^{M'}$ , with  $\gamma_{M'}^M \geq 0$ ,  $M' = M, \dots, m_M$  and  $\sum_{M'=M}^{m_M} \gamma_{M'}^M = 1$ , such that

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

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

$$\left\| \sum_{M'=M}^{m_M} \gamma_{M'}^M \times DF_t^{M'} - Y \right\|_{L^2(\Omega; l^2 \times \mathbb{R}^d)} \leq \sum_{M'=M}^{m_M} \gamma_{M'}^M \|Dx_t^{M'} - Y\|_{L^2(\Omega; l^2 \times \mathbb{R}^d)} \rightarrow 0.$$

So  $Y = Dx_t$  and we conclude by passing to the limit  $M_2 \rightarrow \infty$  in (148).

Now we prove (148). We recall the equation (142) and we write  $\mathbb{E}|D_{(\bullet, \circ, \diamond)}^Z x_t^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_t^{M_2}|_{l_2}^2 \leq C[O_1 + O_2 + O_3]$ , with

$$\begin{aligned} O_1 &= \mathbb{E} \left| \int_0^t \nabla_x b(r, x_r^{M_1}, \rho_r) D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} dr - \int_0^t \nabla_x b(r, x_r^{M_2}, \rho_r) D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2} dr \right|_{l_2}^2, \\ O_2 &= \mathbb{E} \left| \mathbb{1}_{\{0 < T_\bullet \leq t\}} (\mathbb{1}_{\{1 \leq \bullet \leq M_1\}} \partial_{z_\diamond} c(T_\bullet^\diamond, \eta_{T_\bullet^\diamond}^2(W_\bullet^\diamond), Z_\bullet^\diamond, x_{T_\bullet^\diamond-}^{M_1}, \rho_{T_\bullet^\diamond-}) \right. \\ &\quad \left. - \mathbb{1}_{\{1 \leq \bullet \leq M_2\}} \partial_{z_\diamond} c(T_\bullet^\diamond, \eta_{T_\bullet^\diamond}^2(W_\bullet^\diamond), Z_\bullet^\diamond, x_{T_\bullet^\diamond-}^{M_2}, \rho_{T_\bullet^\diamond-})) \right|_{l_2}^2, \\ O_3 &= \mathbb{E} \left| \sum_{k=1}^{M_1} \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_1}, \rho_{T_i^k-}) D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_1} \right. \\ &\quad \left. - \sum_{k=1}^{M_2} \sum_{i=1}^{J_t^k} \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_2}, \rho_{T_i^k-}) D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_2} \right|_{l_2}^2. \end{aligned}$$

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

$$\begin{aligned} O_1 &\leq C \left[ \mathbb{E} \int_0^t |\nabla_x b(r, x_r^{M_1}, \rho_r) - \nabla_x b(r, x_r^{M_2}, \rho_r)|^2 |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr \right. \\ &\quad \left. + \mathbb{E} \int_0^t |\nabla_x b(r, x_r^{M_1}, \rho_r)|^2 |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr \right] \\ &\leq C \left[ \mathbb{E} \int_0^t |x_r^{M_1} - x_r^{M_2}|^2 |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr \right]. \end{aligned}$$

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

$$\begin{aligned} O_1 &\leq C \left[ \int_0^t (\mathbb{E} |x_r^{M_1} - x_r^{M_2}|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr \right] \\ &\leq C [(\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr]. \end{aligned} \quad (149)$$

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

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

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

$$\begin{aligned}
O_2 &\leq C[\int_0^t (\mathbb{E} |x_r^{M_1} - x_r^{M_2}|^{2+\varepsilon_*})^{\frac{2}{2+\varepsilon_*}} dr + \varepsilon_{M_1 \wedge M_2}] \\
&\leq C(\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}.
\end{aligned} \tag{150}$$

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

$$\begin{aligned}
O_{3,1} &= \mathbb{E}(\sum_{k=M_1 \wedge M_2}^{M_1 \vee M_2} \sum_{i=1}^{J_t^k} |\nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_1 \vee M_2}, \rho_{T_i^k-})| |D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_1 \vee M_2}|_{l_2})^2, \\
O_{3,2} &= \mathbb{E}(\sum_{k=1}^{M_1 \wedge M_2} \sum_{i=1}^{J_t^k} |\nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_1}, \rho_{T_i^k-}) \\
&\quad - \nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_2}, \rho_{T_i^k-})| |D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_2}|_{l_2})^2, \\
O_{3,3} &= \mathbb{E}(\sum_{k=1}^{M_1 \wedge M_2} \sum_{i=1}^{J_t^k} |\nabla_x c(T_i^k, \eta_{T_i^k}^2(W_i^k), Z_i^k, x_{T_i^k-}^{M_1}, \rho_{T_i^k-})| |D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_{T_i^k-}^{M_2}|_{l_2})^2.
\end{aligned}$$

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

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

we get

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

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

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

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

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

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

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

we have

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

Therefore,

$$O_3 \leq C[O_{3,1} + O_{3,2} + O_{3,3}] \leq C[(\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}} + \int_0^t \mathbb{E} |D_{(\bullet, \circ, \diamond)}^Z x_r^{M_1} - D_{(\bullet, \circ, \diamond)}^Z x_r^{M_2}|_{l_2}^2 dr]. \quad (151)$$

Combining (149), (150) and (151),

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

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

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

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

Then by a similar argument as above,  $\mathbb{E} |D_{(\diamond)}^\Delta x_t^{M_1} - D_{(\diamond)}^\Delta x_t^{M_2}|_{\mathbb{R}^d}^2 \leq C(\varepsilon_{M_1} + \varepsilon_{M_2})^{\frac{2}{2+\varepsilon_*}}$ , and we obtain (148).

**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. *Probability Theory and Related Fields*, **102**, (1995).
- [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: *Financial 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. Rüdiger, and P. Sundar: On uniqueness and stability for the Enskog equation (2020).
- [22] E. Gobet, G. Pagès, 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 SDE'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 and 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, Birkhäuser, 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  $\alpha$ -stable processes. *Ann. Probab.* Vol **42**, No.5, 1885-1910 (2014).

## Part III

.

Approximation for the invariant measure with applications  
for jump processes (convergence in total variation distance)



# Approximation for the invariant measure with applications for jump processes (convergence in total variation distance)

Vlad Bally\* and Yifeng Qin\*

\*Université Gustave Eiffel, LAMA (UMR CNRS, UPEMLV, UPEC), MathRisk INRIA, F-77454 Marne-la-Vallée, France. Email address: bally@univ-mlv.fr

2023

**Abstract** In this paper, we establish an abstract framework for the approximation of the invariant probability measure for a Markov semigroup. Following Pagès and Panloup [40] we use an Euler scheme with decreasing step (unadjusted Langevin algorithm). Under some contraction property with exponential rate and some regularization properties, we give an estimate of the error in total variation distance. This abstract framework covers the main results in [40] and [14]. As a specific application we study the convergence in total variation distance to the invariant measure for jump type equations. The main technical difficulty consists in proving the regularization properties - this is done under an ellipticity condition, using Malliavin calculus for jump processes.

**Key words:** Invariant measure, Unadjusted Langevin algorithm, Euler scheme with decreasing steps, Total variation distance, Malliavin calculus, Regularization lemma, Jump process

## Contents

<b>1</b>	<b>Introduction</b>	<b>112</b>
<b>2</b>	<b>Approximation of the invariant measure: Abstract framework</b>	<b>115</b>
2.1	The semigroup and the invariant measure . . . . .	115
2.2	The Euler scheme . . . . .	116
2.3	Regularization properties . . . . .	118
<b>3</b>	<b>Abstract integration by parts framework</b>	<b>121</b>
3.1	Main consequences . . . . .	125
<b>4</b>	<b>Application for jump equations</b>	<b>127</b>
4.1	Basic notations and the main equation . . . . .	127
4.2	Hypotheses . . . . .	128
4.3	The truncated Euler scheme . . . . .	130
4.4	Some examples . . . . .	131
<b>5</b>	<b>Malliavin framework for jump equations</b>	<b>132</b>
5.1	Sobolev norms . . . . .	134
5.2	Covariance matrix . . . . .	135
5.3	Auxiliary results . . . . .	139
<b>6</b>	<b>Proof of Theorem 4.1</b>	<b>140</b>
6.1	Euler: condition (7) . . . . .	140
6.2	Lipschitz: condition (4) and the existence of an invariant measure . . . . .	140
6.3	Regularity: conditions (14), (15), (16) and (20) . . . . .	142

# 1 Introduction

The aim of this paper is to study the convergence to the invariant measure of a Markov process. We refer to [18], [35], [38] for the existence of an invariant probability measure for a general Markov process and to [41], [42] for some basic computation of the invariant probability measure for a Lévy process. Following the ideas from Pagès and Panloup [40] (see also Lamberton and Pagès [30] [31]) we use an Euler scheme with decreasing step (known in the literature as the unadjusted Langevin algorithm) in order to construct our algorithm (this has been studied in depth in [45]).

Our paper has two parts. In the first part we construct an abstract framework which is appropriate in order to state and discuss our approximation problem. We focus on the estimate of the error in total variation distance. And the main achievement is to give some sufficient regularization properties for the semigroup and for the Euler scheme, which allow to treat bounded and measurable test functions. Furthermore, in order to check such regularization properties, one has to use integration by parts techniques inspired from Malliavin calculus. We give a regularization lemma based on such arguments, which is the crucial step in our approach (it has its own interest, beyond the application in this particular framework). Let us mention that the abstract framework settled in our paper encompass the following recent results: in [40], the authors use unadjusted Langevin algorithm to approximate the invariant probability measure of a diffusion process and study the Wasserstein and total variation distance between them. In [14], the authors approximate the invariant probability measure of a Lévy process but only study the Wasserstein distance.

In the second part of the paper we illustrate our results in the case of jump type  $SDE$ 's. In order to do it we recall the Malliavin calculus for jump processes and prove estimates of the Sobolev norms and of the Malliavin covariance matrix for the solution of such equations. These estimates are rather long and technical, but at a certain extend they come back on results already obtained in [44]. Once these estimates are proved, we apply the abstract results from the first part and obtain the estimate of the error in total variation distance.

Let us present in more detail our results. We give in Section 2 the abstract framework of the approximation for the invariant probability measure. 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  $l$ . We consider a semigroup  $P_t, t \geq 0$  on the space  $\mathcal{M}_b(\mathbb{R}^d)$  of the bounded measurable functions on  $\mathbb{R}^d$  and assume that there exists at least one invariant probability measure  $\nu$  for the semigroup  $P_t, t \geq 0$ . We assume moreover the "exponential Lipschitz property": there exists two constants  $C_0 \geq 1$  and  $\rho > 0$  such that for every  $t > 0$  and every  $\varphi \in C_b^1(\mathbb{R}^d)$

$$(L_0) \quad \|\nabla P_t \varphi\|_\infty \leq C_0 \|\nabla \varphi\|_\infty e^{-\rho t}.$$

This immediately implies that  $\nu$  is unique.

In order to approximate the invariant measure  $\nu$ , we introduce an Euler scheme with decreasing time steps (unadjusted Langevin algorithm). For every  $\gamma > 0$  we give an operator  $\bar{P}_\gamma : C_b^\infty \rightarrow C_b^\infty$  such that  $\|\bar{P}_\gamma \varphi\|_\infty \leq \|\varphi\|_\infty$  and which approximates our semigroup in the following sense: for every  $\gamma > 0$

$$A(k_0, \alpha) \quad \|(P_\gamma - \bar{P}_\gamma)\varphi\|_\infty \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma^{1+\alpha}.$$

Here  $\alpha > 0$  is a given number,  $k_0 \in \mathbb{N}$  and  $\|\psi\|_{k_0, \infty} = \sum_{|\alpha| \leq k_0} \|\partial^\alpha \psi\|_\infty$ . We consider a decreasing sequence

of time steps  $\gamma_n \downarrow 0$  and define the time grid  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . We assume that

$$(\Gamma) \quad \sum_{i=1}^{\infty} \gamma_i = \lim_{n \rightarrow \infty} \Gamma_n = \infty.$$

We also introduce

$$\bar{\omega} = \bar{\omega}((\gamma_n)_{n \in \mathbb{N}}) = \overline{\lim}_{n \rightarrow \infty} \frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2} < \infty.$$

The typical example is  $\gamma_n = \frac{1}{n}$  and then  $\bar{\omega} = 1$ . In the following we denote  $\{\Gamma\} = \{\Gamma_n, n \in \mathbb{N}\}$ . And, for  $\Gamma_i \leq t < \Gamma_{i+1}$  we denote  $N(t) = i$  and  $\tau(t) = \Gamma_i$ . Then, for  $s \in \{\Gamma\}$  and  $t \in \{\Gamma\}$  we define the Euler scheme

$$\bar{P}_{s,t} = \prod_{i=N(s)}^{N(t)-1} \bar{P}_{\gamma_i}, \quad (1)$$

the product being understood in the sense of composition. This means that we travel from  $\tau(s)$  to  $\tau(t)$  by using the Euler scheme associated to the one step Euler scheme  $\bar{P}_\gamma$ .

So now we use the Euler scheme with decreasing time steps  $\bar{P}_{0,\Gamma_n}$  (given in (1)) to approximate the invariant probability measure  $\nu$ . Our aim is to estimate the total variation distance between them. To do so, we need some regularization properties. First we give the regularization hypothesis concerning the semigroup  $P_t$ :

$$\begin{aligned} R_P(k) \sup_{1 \leq t \leq 2} \|\nabla P_t \varphi\|_{k-1,\infty} &\leq C_k \|\varphi\|_\infty, \quad \text{and} \\ R'_P(k) \sup_{1 \leq t \leq 2} \|\nabla P_t \varphi\|_{k-1,\infty} &\leq C'_k \|\nabla \varphi\|_\infty. \end{aligned}$$

We also introduce the following variant of the Lipschitz property:

$$\bar{L}_k \|\nabla P_t \varphi\|_{k,\infty} \leq C_k \|\nabla \varphi\|_{k,\infty}, \quad 1 \geq t > 0.$$

We give now the regularization properties for the Euler scheme  $\bar{P}_{s,t}$ . To begin, we introduce some notations. We fix a super kernel  $\phi$  (see (18) for the precise definition), and, for  $\delta \in (0, 1]$  we denote  $\phi_\delta(y) = \frac{1}{\delta^d} \phi(\frac{y}{\delta})$ . Moreover, for a function  $\varphi$  we denote  $\varphi_\delta$  the regularization by convolution with the super kernel:  $\varphi_\delta = \varphi * \phi_\delta$ , with  $*$  denoting convolution. For  $\delta > 0, \eta > 0$ , and  $q, \kappa, p \in \mathbb{N}$  we denote

$$A_{q,\kappa,p}^{\delta,\eta}(h) = \frac{\delta^q}{\eta^{2q}} + \eta^{-p} h^p + \eta^\kappa, \quad h > 0.$$

Let  $\beta > 0$  and  $p \geq 1$  be fixed and we assume the following regularization property for the Euler scheme  $\bar{P}_{s,t}$ : we assume that for every  $q, \kappa \in \mathbb{N}$  there exists a constant  $C = C_{q,\kappa,p}$  such that for every  $\delta > 0, \eta > 0$ , every  $1 < t < r < t + 2$  and every bounded measurable function  $\varphi$

$$\begin{aligned} R_{\bar{P}}(p, \beta) \quad & \|\bar{P}_{t-1,t} P_{t,r} \varphi - \bar{P}_{t-1,t} P_{t,r} \varphi_\delta\|_\infty + \|\bar{P}_{t-1,t} \bar{P}_{t,r} \varphi - \bar{P}_{t-1,t} \bar{P}_{t,r} \varphi_\delta\|_\infty \\ & \leq C_{q,\kappa,p} \times A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) \|\varphi\|_\infty. \end{aligned}$$

Now we can give our main result (see **Proposition 2.1.1**). We assume that an invariant probability measure  $\nu$  exists for the semigroup  $P_t, t \geq 0$ . We construct an Euler scheme with decreasing time steps  $\bar{P}_{s,t}$  by (1). Suppose that  $(L_0)$  holds for some  $\rho$ ,  $A(k_0, \alpha)$  holds for some  $k_0, \alpha$  with  $\rho > \alpha \bar{\omega}$ ,  $R_P(k)$ ,  $R'_P(k)$  and  $\bar{L}_k$  hold for every  $k$ , and  $R_{\bar{P}}(p, \beta)$  holds true for some  $p, \beta$ . Then the invariant probability measure  $\nu$  is unique and for any  $\varepsilon > 0$ , for every  $x \in \mathbb{R}^d$  and  $n$  large enough,

$$d_{TV}(\bar{P}_{0,\Gamma_n}(x, \cdot), \nu) \leq C_\varepsilon (\gamma_n^{(p\beta) \wedge \alpha})^{-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\rho \Gamma_n}.$$

We remark that we get the same speed of convergence as in [40] and [14], but in a more general framework.

We notice that we need some regularization properties (see  $R_P(k)$ ,  $R'_P(k)$  and  $R_{\bar{P}}(p, \beta)$ ). In order to obtain these properties, we introduce in Section 3 an abstract framework built on a particular case of the Dirichlet form theory (see [4] and [7]) in which such a property may be obtained by using some integration by parts techniques. Those techniques are very similar to the standard Malliavin calculus but are presented in a more general framework which goes beyond the sole case of the Wiener space. In particular, we aim at providing a minimalist setting leading to our regularization lemma. Our unified framework includes the standard Malliavin calculus and different known versions: the calculus based on the splitting method developed and used in [5], [6], [8] as well as the  $\Gamma$ -calculus in [4]. We also mention that our approach

applies in the case of the Malliavin calculus for jump type processes as settled by [12] and in the "lent particle" approach for Poisson point measures developed by [13].

In Section 4, we apply the results in Section 2 for jump processes. So we consider the  $d$ -dimensional stochastic differential equation with jumps as follows:

$$X_t = x + \int_0^t b(X_r)dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{r-})N(dz, dr), \quad (2)$$

where  $N(dz, dr)$  is a Poisson point measure on the state space  $\mathbb{R}^d$  with intensity measure  $\widehat{N}(dz, dr) = \mu(dz)dr$ ,  $x$  is the initial value,  $\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ , and  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ . Some basic background of jump processes can be found in [15], [19], [46], [47] and [3].

We need to give sufficient conditions to ensure the existence of an invariant probability measure for the jump equation (2). We recall by [18] the classical results of the existence of an invariant probability measure for a general Markov process. Recently, [33] gives some specific criterias for the existence of an invariant probability measure of a jump process and also discuss some ergodicity properties. Here we suppose that (**Hypothesis 2.5**)

$$\begin{aligned} i) \quad & \langle x - y, b(x) - b(y) \rangle \leq -\bar{b} |x - y|^2 \\ ii) \quad & |c(z, x) - c(z, y)| \leq \bar{c}(z) |x - y| \end{aligned}$$

and

$$iii) \quad 2\bar{b} - \int_{\mathbb{R}^d} (2\bar{c}(z) + \bar{c}^2(z))\mu(dz) := \theta > 0.$$

Our conditions are based on [18] and are essentially the same as the conditions in [33]. Indeed, the conditions above implies that for some  $\bar{\beta}, \bar{\alpha} > 0$  and a Lyapunov function  $V(x) = |x|^2$ , we have  $LV \leq \bar{\beta} - \bar{\alpha}V$ , with  $L$  denoting the infinitesimal operator of (2). This guarantees the existence of an invariant probability measure  $\nu$ .

Moreover, in order to apply the Malliavin framework in Section 3 and obtain regularization properties, we assume (see **Hypothesis 2.4 b**)) that the measure  $\mu$  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. We also need some regularity and ellipticity conditions on the coefficients (see **Hypothesis 2.1~2.3** for details). We mention that for every multi-indices  $\beta_1, \beta_2$ , we assume that there exists a non-negative function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that

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

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  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that for every  $\zeta \in \mathbb{R}^d$ ,

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

Now we construct the Euler scheme. We take a partition with decreasing time steps  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{n-1} < \Gamma_n < \dots\}$  with the time steps  $\gamma_n = \Gamma_n - \Gamma_{n-1}$ ,  $n \in \mathbb{N}$  verifying some suitable conditions (see Section 4.3 for details). For  $\Gamma_n \leq t < \Gamma_{n+1}$  we denote  $\tau(t) = \Gamma_n$ . We consider the Euler scheme:

$$X_t^{\mathcal{P}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}})dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{\tau(r)-}^{\mathcal{P}})N(dz, dr).$$

Some results concerning the convergence of the Euler scheme of a jump equation can be found for example in [43], [22], [24], [23], [21], [25] and [2].

Since  $\mu(\mathbb{R}^d) = \infty$  (which is a consequence of **Hypothesis 2.4 a**)), we have infinitely many jumps. So we construct the truncated Euler scheme in order to have finite numbers of jumps for the sake of simulation and Malliavin calculus. For  $m \in \mathbb{N}$ , we denote  $B_m = \{z \in \mathbb{R}^d : |z| \leq m\}$  and denote

$$\varepsilon_m := \int_{\{|z| > m\}} |\bar{c}(z)|^2 \mu(dz) + \left| \int_{\{|z| > m\}} \bar{c}(z) \mu(dz) \right|^2.$$

For every  $\gamma > 0$ , we define the truncation function  $M(\gamma) \in \mathbb{N}$  to be the smallest integer such that

$$\varepsilon_{M(\gamma)} \leq \gamma^2.$$

For  $\Gamma_n < t \leq \Gamma_{n+1}$ , we denote  $M_{\mathcal{P}}(t) = M(\gamma_{n+1})$ . Now we cancel the "big jumps" (the jumps of size  $|z| > M_{\mathcal{P}}(t)$ ):

$$X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \int_0^t \int_{B_{M_{\mathcal{P}}}(r)} c(z, X_{\tau(r)-}^{\mathcal{P}, M_{\mathcal{P}}}) N(dz, dr). \quad (3)$$

We remark that the solution of the equation (3) can be constructed in an explicit way.

Then we apply the abstract framework in Section 2 for  $X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}$  and obtain the following main result (see **Theorem 4.1**): An invariant probability measure  $\nu$  of the jump equation (2) exists and is unique, and for any  $\varepsilon > 0$ , there exists a constant  $C_\varepsilon$  such that for every  $x \in \mathbb{R}^d$  and  $n$  large enough, we have

$$d_{TV}(\mathcal{L}(X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}), \nu) \leq C_\varepsilon (\gamma_n^{1-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\frac{\varepsilon}{2} \Gamma_n}),$$

with  $\mathcal{L}(X)$  denoting the law of a random variable  $X$ . We notice that we obtain the same speed of convergence as in [40] but [40] concern the diffusion process driven by a Brownian motion while here we consider the jump process. Comparing with the results in [14], we also obtain the same speed of convergence but [14] only deals with the Wasserstein distance while in our paper, we deal with the total variation distance.

## 2 Approximation of the invariant measure: Abstract framework

### 2.1 The semigroup and the invariant measure

We consider a semigroup  $P_t, t \geq 0$  on the space  $\mathcal{M}_b(\mathbb{R}^d)$  of the bounded measurable functions on  $\mathbb{R}^d$ . 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  $l$ . We will use the following two hypotheses:

(I) We assume that there exists at least one invariant distribution for the semigroup  $P_t, t \geq 0$ .

Moreover we assume the following "exponential Lipschitz property": we assume that there exists two constants  $C_0 \geq 1$  and  $\rho > 0$  such that for every  $t > 0$  and every  $\varphi \in C_b^1(\mathbb{R}^d)$

$$(L_0) \quad \|\nabla P_t \varphi\|_\infty \leq C_0 \|\nabla \varphi\|_\infty e^{-\rho t}. \quad (4)$$

We also denote by  $\mathcal{P}_1$  the space of the probability measures on  $\mathbb{R}^d$  which have finite moment of order one  $\int_{\mathbb{R}^d} |x| \nu(dx) < \infty$ . This is a Banach space under the Wasserstein distance  $W_1$ :

$$W_1(\nu, \mu) = \sup \left\{ \left| \int_{\mathbb{R}^d} \varphi d(\nu - \mu) \right| : \|\nabla \varphi\|_\infty \leq 1 \right\}.$$

**Proposition 2.0.1.** *Suppose that the semigroup  $P_t, t \geq 0$  has at least an invariant probability measure  $\nu$  and that (4) holds true. Then the invariant probability measure is unique and moreover, for every  $x \in \mathbb{R}^d$*

$$W_1(\nu, P_t(x, \cdot)) \leq C \int_{\mathbb{R}^d} |x - y| \nu(dy) \times e^{-\rho t}. \quad (5)$$

**Proof. Step 1** We will prove that for sufficiently large  $t$ , the application  $\nu \mapsto \nu P_t$  is a strict contraction on the Wasserstein space: using (4),

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(y) d(\nu P_t - \mu P_t)(dy) \right| &= \left| \int_{\mathbb{R}^d} P_t \varphi(x) d(\nu(x) - \mu(x)) \right| \\ &\leq \|\nabla P_t \varphi\|_\infty W_1(\nu, \mu) \\ &\leq C_0 \|\nabla \varphi\|_\infty e^{-\rho t} W_1(\nu, \mu). \end{aligned}$$

This means that, for large  $t$

$$W_1(\nu P_t, \mu P_t) \leq C_0 e^{-\rho t} W_1(\nu, \mu) \leq \frac{1}{2} W_1(\nu, \mu)$$

and this guarantees the uniqueness of the invariant measure.

**Step 2** Since  $\nu$  is an invariant measure

$$\int_{\mathbb{R}^d} \varphi(z) \nu(dz) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} P_t(z, dy) \varphi(y) \nu(dz)$$

which gives, for every fixed  $x \in \mathbb{R}^d$  ( $\nu$  is a probability)

$$\begin{aligned} \int_{\mathbb{R}^d} \varphi(z) \nu(dz) - \int_{\mathbb{R}^d} P_t(x, dy) \varphi(y) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} (P_t(z, dy) - P_t(x, dy)) \varphi(y) \nu(dz) \\ &= \int_{\mathbb{R}^d} (P_t \varphi(z) - P_t \varphi(x)) \nu(dz) \end{aligned} \quad (6)$$

so that

$$\begin{aligned} \left| \int_{\mathbb{R}^d} \varphi(z) \nu(dz) - \int_{\mathbb{R}^d} P_t(x, dy) \varphi(y) \right| &\leq \|\nabla P_t \varphi\|_{\infty} \int_{\mathbb{R}^d} |x - z| \nu(dz) \\ &\leq C_0 e^{-\rho t} \|\nabla \varphi\|_{\infty} \int_{\mathbb{R}^d} |x - z| \nu(dz) \end{aligned}$$

which yields (5).  $\square$

## 2.2 The Euler scheme

We introduce now an Euler scheme with decreasing steps. First, for every  $\gamma > 0$  we give an operator  $\bar{P}_{\gamma} : C_b^{\infty} \rightarrow C_b^{\infty}$  such that  $\|\bar{P}_{\gamma} \varphi\|_{\infty} \leq \|\varphi\|_{\infty}$  and which approximates our semigroup in the following sense: for every  $\gamma > 0$

$$A(k_0, \alpha) \quad \|(P_{\gamma} - \bar{P}_{\gamma}) \varphi\|_{\infty} \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma^{1+\alpha}. \quad (7)$$

Here  $\alpha > 0$  is a given number,  $k_0 \in \mathbb{N}$  and

$$\|\psi\|_{k_0, \infty} = \sum_{|\alpha| \leq k_0} \|\partial^{\alpha} \psi\|_{\infty}.$$

Moreover, we consider a decreasing sequence of time steps  $\gamma_n \downarrow 0$  and define the time grid  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . We assume that

$$(\Gamma) \quad \sum_{i=1}^{\infty} \gamma_i = \lim_{n \rightarrow \infty} \Gamma_n = \infty. \quad (8)$$

We also introduce

$$\bar{\omega} = \bar{\omega}((\gamma_n)_{n \in \mathbb{N}}) = \overline{\lim_{n \rightarrow \infty} \frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2}} < \infty.$$

The typical example is  $\gamma_n = \frac{1}{n}$  and then  $\bar{\omega} = 1$ . In the following we denote  $\{\Gamma\} = \{\Gamma_n, n \in \mathbb{N}\}$ . And, for  $\Gamma_i \leq t < \Gamma_{i+1}$  we denote

$$N(t) = i \quad \text{and} \quad \tau(t) = \Gamma_i.$$

In particular, for  $t = \Gamma_i \in \{\Gamma\}$  we have  $N(t) = i$  such that  $t = \Gamma_{N(t)}$ . Then, for  $s \in \{\Gamma\}$  and  $t \in \{\Gamma\}$  we define the Euler scheme

$$\bar{P}_{s,t} = \prod_{i=N(s)}^{N(t)-1} \bar{P}_{\gamma_i} \quad (9)$$

the product being understood in sense of composition. This means that we travel from  $\tau(s)$  to  $\tau(t)$  by using the Euler scheme associated to the one step Euler scheme  $\bar{P}_\gamma$ . In the appendix 7.1 we will prove the following lemma (which is a slight generalisation of the lemma given by Pages and Panloup [40]): for every  $\rho > \alpha\bar{\omega}$ , there exists  $n_\rho$  and  $C_\rho$  such that for  $n \geq n_\rho$

$$\sum_{i=1}^n \gamma_i^{1+\alpha} e^{-\rho(\Gamma_n - \Gamma_i)} \leq C_\rho \gamma_n^\alpha. \quad (10)$$

Moreover, there exists  $n_*$  such that, for  $n_* \leq i \leq n$

$$\gamma_i \leq e^{2\bar{\omega}(\Gamma_n - \Gamma_i)} \gamma_n. \quad (11)$$

Notice that  $P_t, t \geq 0$  is a homogeneous semigroup, and we may define  $P_{s,t} = P_{t-s} = P_{0,t-s}$ . In contrast,  $\bar{P}_{s,t}, s < t$ , is not homogeneous: we do not have  $\bar{P}_{s,t} = \bar{P}_{0,t-s}$ . This is due to the fact that the greed  $\Gamma_i, i \in \mathbb{N}$  is not uniform.

Finally we assume the following stronger variant of the Lipschitz property  $L_0$ :

$$(L_{k_0}) \quad \|\nabla P_t \varphi\|_{k_0, \infty} \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} e^{-\rho t} \quad (12)$$

where  $k_0$  is the one from  $A(k_0, \alpha)$ .

**Proposition 2.0.2.** *Suppose that (7) and (12) hold true with  $\rho > \alpha\bar{\omega}$ . Then for  $N(t) > n_\rho + 1$ , we have*

$$\|(P_{s,t} - \bar{P}_{s,t})\varphi\|_\infty \leq C_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma_{N(t)}^\alpha. \quad (13)$$

**Proof** We use (7) first and (12) then

$$\begin{aligned} \|(P_{s,t} - \bar{P}_{s,t})\varphi\|_\infty &\leq \sum_{i=N(s)}^{N(t)-1} \|\bar{P}_{s, \Gamma_{i-1}}(\bar{P}_{\gamma_i} - P_{\gamma_i})P_{\Gamma_i, t}\varphi\|_\infty \\ &\leq \sum_{i=N(s)}^{N(t)-1} \|(\bar{P}_{\gamma_i} - P_{\gamma_i})P_{\Gamma_i, t}\varphi\|_\infty \\ &\leq C_{k_0} \sum_{i=N(s)}^{N(t)-1} \|\nabla P_{\Gamma_i, t}\varphi\|_{k_0, \infty} \gamma_i^{1+\alpha} \\ &\leq C'_{k_0} \sum_{i=N(s)}^{N(t)-1} \|\nabla \varphi\|_{k_0, \infty} \gamma_i^{1+\alpha} e^{-\rho(\Gamma_{N(t)} - \Gamma_i)} \\ &\leq C''_{k_0} \|\nabla \varphi\|_{k_0, \infty} \gamma_{N(t)}^\alpha. \end{aligned}$$

For the last inequality we have used (10).  $\square$

*Remark.* Suppose that (7) and (12) hold with  $k_0 = 0$ . We also suppose that an invariant probability measure  $\nu$  of the semigroup  $P_t, t \geq 0$  exists and that (4) holds true. Then **Proposition 2.0.1** and **Proposition 2.0.2** give that for every  $x \in \mathbb{R}^d$ , we have

$$W_1(\nu, \bar{P}_{0,t}(x, \cdot)) \leq C(\gamma_{N(t)}^\alpha) + \int_{\mathbb{R}^d} |x - y| \nu(dy) \times e^{-\rho t}.$$

For this result, we do not need any regularization properties. In order to obtain the result for the total variation distance, we give some regularization properties in the next subsection.

## 2.3 Regularization properties

In this section we will assume that the semigroup and the Euler scheme have some regularization properties which allow to obtain convergence in total variation distance.

First we give the regularization hypothesis concerning the semigroup:

$$R_P(k) \sup_{1 \leq t \leq 2} \|\nabla P_t \varphi\|_{k-1, \infty} \leq C_k \|\varphi\|_{\infty}, \quad \text{and} \quad (14)$$

$$R'_P(k) \sup_{1 \leq t \leq 2} \|\nabla P_t \varphi\|_{k-1, \infty} \leq C'_k \|\nabla \varphi\|_{\infty}, \quad (15)$$

Such a regularization property is proved using the integration by parts formula in Malliavin calculus. Moreover, we suppose that we have the following variant of the Lipschitz property:

$$\begin{aligned} \bar{L}_k \quad i) \quad & \|\nabla P_t \varphi\|_{\infty} \leq C_k \|\nabla \varphi\|_{k, \infty} e^{-\rho t}, \quad t \geq 1, \\ ii) \quad & \|\nabla P_t \varphi\|_{k, \infty} \leq C_k \|\nabla \varphi\|_{k, \infty}, \quad 1 \geq t > 0. \end{aligned} \quad (16)$$

Notice that  $\bar{L}_k, i)$  is weaker than  $L_0$  (see (4)) because we have  $\|\nabla \varphi\|_{k, \infty}$  instead of  $\|\nabla \varphi\|_{\infty}$ . However, if the regularization property  $R'_P(k)$  holds then  $\bar{L}_k, i)$  implies  $L_0$  (for  $t \geq 1$ ). Indeed,  $\bar{L}_k$  gives

$$\begin{aligned} \|\nabla P_t \varphi\|_{\infty} &= \|\nabla(P_{t-1} P_1 \varphi)\|_{\infty} \leq C \|\nabla P_1 \varphi\|_{k, \infty} e^{-\rho(t-1)} \\ &\leq C \|\nabla \varphi\|_{\infty} e^{-\rho(t-1)}, \end{aligned}$$

the last inequality being the consequence of  $R'_P(k)$ . In particular, if an invariant probability measure  $\nu$  exists, then it is unique and we have (5).

*Remark.* We also notice that  $R'_P(k+1)$  and  $\bar{L}_k$  imply  $L_k$ . Indeed, for  $t \leq 1$ ,  $\bar{L}_k$  ii) gives

$$\|\nabla P_t \varphi\|_{k, \infty} \leq C_k \|\nabla \varphi\|_{k, \infty} \leq e^{\rho} C_k \|\nabla \varphi\|_{k, \infty} e^{-\rho t}$$

and for  $t \geq 1$

$$\begin{aligned} \|\nabla P_t \varphi\|_{k, \infty} &= \|\nabla(P_1 P_{t-1} \varphi)\|_{k, \infty} \leq C \|\nabla P_{t-1} \varphi\|_{\infty} \\ &\leq C \|\nabla \varphi\|_{k, \infty} e^{-\rho(t-1)}. \end{aligned}$$

Moreover, for  $t \geq 1$ ,  $\bar{L}_k$  and  $R_P(k+1)$  give

$$d_{TV}(P_t(x, \cdot), \nu) \leq C \left( \int_{\mathbb{R}^d} |x - y| d\nu(y) \right) e^{-\rho t}, \quad (17)$$

where  $d_{TV}$  denotes the total variation distance:

$$d_{TV}(\mu, \nu) = \sup_{\|f\|_{\infty} \leq 1} \left| \int_{\mathbb{R}^d} f(x) \mu(dx) - \int_{\mathbb{R}^d} f(x) \nu(dx) \right|.$$

Indeed,

$$\begin{aligned} |P_t \varphi(x) - P_t \varphi(y)| &= |P_{t-1} P_1 \varphi(x) - P_{t-1} P_1 \varphi(y)| \\ &\leq C_k \|\nabla P_1 \varphi\|_{k, \infty} e^{-\rho(t-1)} |x - y| \\ &\leq C_k C_{k+1} e^{\rho} \|\varphi\|_{\infty} e^{-\rho t} |x - y|. \end{aligned}$$

Then we come back to (6) and we obtain

$$\left| \int_{\mathbb{R}^d} \varphi(z) \nu(dz) - \int_{\mathbb{R}^d} P_t(x, dy) \varphi(y) \right| \leq C \|\varphi\|_{\infty} \int_{\mathbb{R}^d} e^{-\rho t} |x - y| \nu(dy)$$



so (17) is proved.  $\square$

We give now the regularization properties for the Euler scheme; this is a more delicate subject, because we have some difficulties in order to use directly the Malliavin calculus for the Euler scheme (the reason is that the decomposition using the inverse of the tangent flow does not work, and so the proof of the non degeneracy property is more difficult) .

We introduce some notations. We recall that a super kernel  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a function which belongs to the Schwartz space and such that for every multi-indexes  $\beta_1$  and  $\beta_2$ , one has

$$\int_{\mathbb{R}^d} \phi(x) dx = 1, \quad \int_{\mathbb{R}^d} y^{\beta_1} \phi(y) dy = 0 \quad \text{for } |\beta_1| \geq 1, \quad \int_{\mathbb{R}^d} |y|^m |\partial_{\beta_2} \phi(y)| dy < \infty \quad \text{for } m \in \mathbb{N}. \quad (18)$$

We fix a super kernel  $\phi$ . For  $\delta \in (0, 1]$ , we denote  $\phi_\delta(y) = \frac{1}{\delta^d} \phi(\frac{y}{\delta})$  and  $\varphi_\delta$  the regularization by convolution with a super kernel:

$$\varphi_\delta = \varphi * \phi_\delta, \quad (19)$$

with  $*$  denoting convolution.

As usual, for a multi-index  $\beta_1 = (\beta_1^1, \dots, \beta_1^m) \in \{1, \dots, d\}^m$ , one denotes  $|\beta_1| = m$  and  $y^{\beta_1} = \prod_{i=1}^m y_{\beta_1^i}$ . For  $\delta > 0, \eta > 0$ , and  $q, \kappa, p \in \mathbb{N}$  we denote

$$A_{q, \kappa, p}^{\delta, \eta}(h) = \frac{\delta^q}{\eta^{2q}} + \eta^{-p} h^p + \eta^\kappa, \quad h > 0.$$

Then we assume the following:

Let  $\beta > 0$  and  $p \geq 1$  be fixed. We assume that for every  $q, \kappa \in \mathbb{N}$  there exists a constant  $C = C_{q, \kappa, p}$  such that for every  $\delta > 0, \eta > 0$ , every  $1 < t < r < t + 2$  and every bounded measurable function  $\varphi$

$$\begin{aligned} R_{\overline{P}}(p, \beta) & \quad \left\| \overline{P}_{t-1, t} P_{t, r} \varphi - \overline{P}_{t-1, t} P_{t, r} \varphi_\delta \right\|_\infty + \left\| \overline{P}_{t-1, t} \overline{P}_{t, r} \varphi - \overline{P}_{t-1, t} \overline{P}_{t, r} \varphi_\delta \right\|_\infty \\ & \leq C_{q, \kappa, p} \times A_{q, \kappa, p}^{\delta, \eta}(\gamma_{N(t-1)}^\beta) \|\varphi\|_\infty. \end{aligned} \quad (20)$$

This represents the "regularization property for  $\overline{P}_{t-1, t}$ ". In order to prove it, one employs **Lemma 3.5** (see (38)) in Section 3.1.

As a consequence of these properties, we obtain the following lemma. We recall  $n_\rho$  and  $n_*$  in (10) and (11).

**Lemma 2.1.** *We fix  $\beta > 0$  and  $p \geq 1$ . Suppose that (7) (12) hold with  $\rho > \alpha\overline{\omega}$ , and  $R_{\overline{P}}(p, \beta)$  (see (20)) holds. Then, for every  $\varepsilon > 0$  there exists a constant  $C_\varepsilon \geq 1$  such that for every  $s < t - 1 < r < t + 2$  with  $N(r) > n_\rho + 1$  and  $N(t - 1) > n_*$ , and for every bounded measurable function  $\varphi$*

$$\left\| \overline{P}_{s, t} (\overline{P}_{t, r} - P_{t, r}) \varphi \right\|_\infty \leq C_\varepsilon \|\varphi\|_\infty \gamma_{N(t)}^{((p\beta) \wedge \alpha) - \varepsilon}. \quad (21)$$

**Proof** We use (20) and (11) in order to get

$$\begin{aligned} \left\| \overline{P}_{s, t} (\overline{P}_{t, r} - P_{t, r}) \varphi \right\|_\infty & \leq \left\| \overline{P}_{t-1, t} (\overline{P}_{t, r} - P_{t, r}) \varphi \right\|_\infty \\ & \leq C_{q, \kappa, p} \|\varphi\|_\infty \times A_{q, \kappa, p}^{\delta, \eta}(\gamma_{N(t-1)}^\beta) + b_\delta \\ & \leq C_{q, \kappa, p} \|\varphi\|_\infty \times A_{q, \kappa, p}^{\delta, \eta}(\gamma_{N(t)}^\beta) + b_\delta \end{aligned}$$

with

$$\begin{aligned} b_\delta & = \left\| \overline{P}_{t-1, t} (\overline{P}_{t, r} - P_{t, r}) \varphi_\delta \right\|_\infty \leq \left\| (\overline{P}_{t, r} - P_{t, r}) \varphi_\delta \right\|_\infty \\ & \leq C \|\nabla \varphi_\delta\|_{k_0, \infty} \gamma_{N(r)}^\alpha \leq \frac{C}{\delta^{1+k_0}} \|\varphi\|_\infty \gamma_{N(t)}^\alpha. \end{aligned}$$

Here we used (13) and  $\gamma_{N(r)} \leq \gamma_{N(t)}$ . We conclude that

$$\left\| \overline{P}_{s, t} (\overline{P}_{t, r} - P_{t, r}) \varphi \right\|_\infty \leq C_{q, \kappa, p} \|\varphi\|_\infty \times (A_{q, \kappa, p}^{\delta, \eta}(\gamma_{N(t)}^\beta) + \frac{1}{\delta^{1+k_0}} \gamma_{N(t)}^\alpha).$$

**Optimization** For some fixed  $\alpha, \beta, p, k_0, \varepsilon$ , we optimize over  $\delta, \eta, \kappa, q$ . Let  $\Delta = \gamma_{N(t)}^\beta$ . First we choose  $\eta = \Delta^{\frac{p}{p+\kappa}}$  so that  $\eta^{-p} \Delta^p = \eta^\kappa$ . Then

$$A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) = \frac{\delta^q}{\Delta^{\frac{2pq}{p+\kappa}}} + 2\Delta^{\frac{p\kappa}{p+\kappa}}.$$

Take now  $\delta = \Delta^{\frac{3p}{p+\kappa}}$  so that

$$A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) = \Delta^{\frac{pq}{p+\kappa}} + 2\Delta^{\frac{p\kappa}{p+\kappa}}.$$

With this choice

$$\begin{aligned} A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) + \frac{\gamma_{N(t)}^\alpha}{\delta^{1+k_0}} &= \Delta^{\frac{pq}{p+\kappa}} + 2\Delta^{\frac{p\kappa}{p+\kappa}} + \Delta^{-\frac{3p(1+k_0)}{p+\kappa}} \gamma_{N(t)}^\alpha \\ &= \gamma_{N(t)}^{\frac{pq\beta}{p+\kappa}} + 2\gamma_{N(t)}^{\frac{p\kappa\beta}{p+\kappa}} + \gamma_{N(t)}^{-\frac{3p(1+k_0)\beta}{p+\kappa}} \times \gamma_{N(t)}^\alpha \end{aligned}$$

We need

$$\begin{aligned} i) \quad & \frac{3p(1+k_0)\beta}{p+\kappa} < \varepsilon, \\ ii) \quad & \frac{\kappa}{p+\kappa} \geq 1 - \varepsilon \\ iii) \quad & \frac{q}{p+\kappa} \geq 1 - \varepsilon. \end{aligned}$$

We first choose  $\kappa(\varepsilon)$  such that *i*) and *ii*) hold true. Then we choose  $q(\varepsilon)$  such that  $\frac{q(\varepsilon)}{p+\kappa(\varepsilon)} \geq 1 - \varepsilon$ . With this choice we have

$$\begin{aligned} \|\bar{P}_{s,t}(\bar{P}_{t,r} - P_{t,r})\varphi\|_\infty &\leq C_{q,\kappa,p} \|\varphi\|_\infty \times (A_{q,\kappa,p}^{\delta,\eta}(\gamma_{N(t)}^\beta) + \gamma_{N(t)}^{-\frac{3p(1+k_0)\beta}{p+\kappa}} \gamma_{N(t)}^\alpha) \\ &\leq C'_{q(\varepsilon),\kappa(\varepsilon),p} \|\varphi\|_\infty \times (\gamma_{N(t)}^{p\beta(1-\varepsilon)} + \gamma_{N(t)}^{\alpha-\varepsilon}) \\ &\leq C'_{q(\varepsilon),\kappa(\varepsilon),p} \|\varphi\|_\infty \times \gamma_{N(t)}^{((p\beta)\wedge\alpha)-\varepsilon}, \end{aligned}$$

with  $\bar{\varepsilon} = p\beta\varepsilon \vee \varepsilon$ .  $\square$

We give now **the main result**. We recall  $n_\rho$  and  $n_*$  in (10) and (11).

**Proposition 2.1.1.** *Let  $\beta > 0$  and  $p \geq 1$  be fixed. Suppose that (7) holds for some  $\alpha, k_0$ , (14), (15), (16) hold for every  $k$  and some  $\rho$  with  $\rho > \alpha\bar{\omega}$ , and  $R_{\bar{p}}(p, \beta)$  (see (20)) holds. For every  $\varepsilon > 0$  and every measurable and bounded function  $\varphi$ , for  $n$  large enough such that  $N(\Gamma_n - 3) > n_*$  and  $N(\Gamma_n - 2) > n_\rho + 1$ , we have*

$$\|(\bar{P}_{0,\Gamma_n} - P_{0,\Gamma_n})\varphi\|_\infty \leq C_\varepsilon \|\varphi\|_\infty \gamma_n^{((p\beta)\wedge\alpha)-\varepsilon}. \quad (22)$$

Moreover, if an invariant probability measure  $\nu$  exists, then the invariant probability measure  $\nu$  is unique and for every  $x \in \mathbb{R}^d$ , we have

$$d_{TV}(\bar{P}_{0,\Gamma_n}(x, \cdot), \nu) \leq C_\varepsilon (\gamma_n^{((p\beta)\wedge\alpha)-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\rho\Gamma_n}). \quad (23)$$

**Proof** We fix  $i < n$  such that  $1 < \Gamma_i$  and  $\Gamma_i + 1 \leq \Gamma_n \leq \Gamma_i + 2$  and we write

$$\begin{aligned} & \|(\bar{P}_{0,\Gamma_n} - P_{0,\Gamma_n})\varphi\|_\infty \\ & \leq \|(\bar{P}_{0,\Gamma_i} \bar{P}_{\Gamma_i,\Gamma_n} - \bar{P}_{0,\Gamma_i} P_{\Gamma_i,\Gamma_n})\varphi\|_\infty + \|(\bar{P}_{0,\Gamma_i} P_{\Gamma_i,\Gamma_n} - P_{0,\Gamma_i} P_{\Gamma_i,\Gamma_n})\varphi\|_\infty \\ & = : A + B. \end{aligned}$$

First, since  $\Gamma_i > 1$ , using (21) with  $s = 0, t = \Gamma_i$  and  $r = \Gamma_n$  we obtain

$$A \leq C_\varepsilon \|\varphi\|_\infty \times \gamma_i^{((p\beta)\wedge\alpha)-\varepsilon} \leq C_\varepsilon \|\varphi\|_\infty \times \gamma_n^{((p\beta)\wedge\alpha)-\varepsilon},$$

where in the last inequality, we have used (11).

Moreover, we recall that (15) and (16) imply (12). So using (13) and the regularization property (14) (notice that  $\Gamma_n - \Gamma_i \geq 1$ ) we obtain

$$B \leq C \|\nabla P_{\Gamma_i, \Gamma_n} \varphi\|_{k_0, \infty} \gamma_i^\alpha \leq C \|\varphi\|_\infty \gamma_i^\alpha \leq C \|\varphi\|_\infty \gamma_n^\alpha,$$

the last inequality being obtained by (11) (because  $\Gamma_n - \Gamma_i \leq 2$ ).

Finally, in order to obtain (23) we use (17). The uniqueness of the invariant probability measure  $\nu$  comes directly from **Proposition 2.0.1**.  $\square$

### 3 Abstract integration by parts framework

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

We denote  $C_p^\infty(\mathbb{R}^d)$  to be the space of smooth functions which, together with all the derivatives, have polynomial growth. We also denote  $C_p^q(\mathbb{R}^d)$  to be the space of  $q$ -times differentiable functions which, together with all the derivatives, have polynomial growth.

We consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , and a linear 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", usually denoted by  $\mathcal{D}_\infty$  (see [37]).

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

a)

$$D_h F := \langle DF, h \rangle_{\mathcal{H}} \in \mathcal{S}, \text{ for any } h \in \mathcal{H}, \quad (24)$$

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

$$D\phi(F) = \sum_{i=1}^d \partial_i \phi(F) DF_i, \quad (25)$$

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

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

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

Then

$$DF = \sum_{i=1}^{\infty} D_i F \times e_i \quad \text{and} \quad D^q F = \sum_{i_1, \dots, i_q} D_{i_1, \dots, i_q} F \times \otimes_{j=1}^q e_{i_j}.$$

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

$$\sigma_F = (\sigma_F^{i,j})_{i,j=1, \dots, d}, \quad \text{with} \quad \sigma_F^{i,j} = \langle DF_i, DF_j \rangle_{\mathcal{H}}. \quad (26)$$

And we denote

$$\Sigma_p(F) = \mathbb{E}(1/\det \sigma_F)^p. \quad (27)$$

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

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

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

$$\mathbb{E}\langle DF, DG \rangle_{\mathcal{H}} = \mathbb{E}(FLG) = \mathbb{E}(GLF), \quad (28)$$

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

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

As an immediate consequence of the duality formula (28), we know that  $L : \mathcal{S} \subset L^2(\Omega) \rightarrow L^2(\Omega)$  is closable. But it is not clear that  $D$  is also closable. We have to assume this and to check it for each particular example.

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

*Remark.* The bilinear forms  $\Gamma(F, G) = \langle DF, DG \rangle_{\mathcal{H}}$  is called "carré du champ" operator in the theory of Dirichlet form. And  $\mathcal{E}(F, G) = \mathbb{E}(\Gamma(F, G))$  is the Dirichlet form associated to  $\Gamma$ . So our Integration by Parts framework appears as a particular case of the  $\Gamma$ -calculus, presented in [4] and [7].

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

$$|F|_{1,l} = \sum_{q=1}^l |D^q F|_{\mathcal{H}^{\otimes q}}, \quad |F|_l = |F| + |F|_{1,l}, \quad (29)$$

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

$$\begin{aligned} \|F\|_{l,p} &= (\mathbb{E} |F|_l^p)^{1/p}, \quad \|F\|_p = (\mathbb{E} |F|^p)^{1/p}, \\ \|F\|_{L,l,p} &= \|F\|_{l,p} + \|LF\|_{l-2,p}. \end{aligned} \quad (30)$$

With these notations, we have the following lemma from [9] (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  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is a  $C^l(\mathbb{R}^d)$  function ( $l$ -times differentiable function), then there is a constant  $C_l$  dependent on  $l$  such that

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

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

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

For  $l = 0$ , we have

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

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

$$\mathcal{D}_{l,p} = \overline{\mathcal{S}}^{\|\circ\|_{L,l,p}}, \quad (31)$$

and

$$\mathcal{D}_\infty = \bigcap_{l=1}^{\infty} \bigcap_{p=1}^{\infty} \mathcal{D}_{l,p}, \quad \mathcal{H}_l = \mathcal{D}_{l,2}. \quad (32)$$

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  $\|F - F_n\|_p \rightarrow 0$ ,  $\|F_n - F_m\|_{q,p} \rightarrow 0$  and  $\|LF_n - LF_m\|_{q-2,p} \rightarrow 0$ . Since  $D^q$  and  $L$  are closable, we can define

$$D^q F = \lim_{n \rightarrow \infty} D^q F_n \quad \text{in} \quad L^p(\Omega; \mathcal{H}^{\otimes q}), \quad LF = \lim_{n \rightarrow \infty} LF_n \quad \text{in} \quad L^p(\Omega). \quad (33)$$

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.  $\square$

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  $F \in L^1(\Omega; \mathbb{R}^d)$  and let  $F_n \in \mathcal{S}^d, n \in \mathbb{N}$  such that*

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

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

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

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

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

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

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

*the lowest eigenvalue of the covariance matrix  $\sigma_F$ . 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  $p \geq 1$ ,*

$$iv) \quad \sup_n \mathbb{E}(\lambda^{-p}(F_n)) \leq Q_p < \infty. \quad (34)$$

*Then we have*

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

**(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*

$$v) \quad \sup_n \|DF_n - D\bar{F}_n\|_{L^2(\Omega; \mathcal{H})} \leq \bar{\varepsilon}, \quad (35)$$

*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. We recall the notations in Section 3. 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  $F_n \rightarrow G$  weakly in the Hilbert space  $\mathcal{H}_l$ . This means that for every  $Q \in \mathcal{H}_l$ ,  $\langle F_n, Q \rangle_{L,l,2} \rightarrow \langle G, Q \rangle_{L,l,2}$ . 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  $\gamma_i^n \geq 0, i = n, \dots, m_n$  and  $\sum_{i=n}^{m_n} \gamma_i^n = 1$ , such that

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

In particular we have

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

Also, we notice that by *i*),

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

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

$$\|G_n\|_{L,l,p} \leq \sum_{i=n}^{m_n} \gamma_i^n \|F_i\|_{L,l,p} \leq K_{l,p}.$$

Then a standard argument gives, for every  $\bar{p} \in [1, p)$ ,

$$\|F\|_{L,l,\bar{p}} \leq K_{l,\bar{p}}.$$

**Proof of (B)** We consider for a moment some general  $F, G \in \mathcal{D}_{\infty}^d$ . Notice that  $\langle \sigma(F)\zeta, \zeta \rangle = |\langle DF, \zeta \rangle|_{\mathcal{H}}^2$ , so  $\lambda(F) = \inf_{|\zeta|=1} |\langle DF, \zeta \rangle|_{\mathcal{H}}^2$ . It is easy to check that

$$|\sqrt{\lambda(F)} - \sqrt{\lambda(G)}| \leq |D(F - G)|_{\mathcal{H}}. \quad (36)$$

We now come back to our framework. Recalling that  $G_n = \sum_{i=n}^{m_n} \gamma_i^n \times F_i$ , we observe that

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

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  $\|DG_n - DF\|_{L^2(\Omega; \mathcal{H})} \rightarrow 0$ . So we conclude that  $\|DF - DF_n\|_{L^2(\Omega; \mathcal{H})} \rightarrow 0$ . Thus, by (36),  $\mathbb{E} |\sqrt{\lambda(F)} - \sqrt{\lambda(F_n)}| \rightarrow 0$ . This gives that there exists a subsequence (also denote by  $n$ ) such that  $\sqrt{\lambda(F_n)}$  converges

to  $\sqrt{\lambda(\bar{F})}$  almost surely, and consequently  $|\lambda(F_n)|^{-p}$  converges to  $|\lambda(F)|^{-p}$  almost surely. Since we have (34),  $(|\lambda(F_n)|^{-p})_{n \in \mathbb{N}}$  is uniformly integrable. It follows that

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

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

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

Then it follows by (35) that

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

□

### 3.1 Main consequences

We will use the abstract framework presented above for the IbP framework  $(\mathcal{D}_\infty, D, L)$ , with  $D$  and  $L$  defined in (33). We recall the notations  $\|F\|_{L, l, p}$  in (30),  $\Sigma_p(F)$  in (27) and  $\sigma_F$  in (26). For any  $\eta > 0$ , we take  $\Upsilon_\eta(x) : (0, \infty) \rightarrow \mathbb{R}$  to be a function of class  $C_b^\infty$  such that

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

We remark that  $\sigma_F$  is invertible on the set  $\{\Upsilon_\eta(\det \sigma_F) > 0\}$ . We give the following lemma, which is stated in lemma 2.4 of [7] and is proved in the Appendix of [6], based on some integration by parts formula.

**Lemma 3.4.** *Let  $F = (F_1, \dots, F_d) \in \mathcal{D}_\infty^d$  and  $G \in \mathcal{D}_\infty$ . We fix  $q \in \mathbb{N}$ .*

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

$$(\mathbf{B}_q) \quad |\mathbb{E}(\partial^\beta f(F)G)| \leq C_q \|f\|_\infty, \quad \forall |\beta| = q. \quad (37)$$

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

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

*Remark.* In (A), we assume the non-degeneracy condition for  $F$ , so we can give the estimate based on the standard integration by parts formula. In (B), we no longer suppose non-degeneracy condition for  $F$ , so we can only obtain an estimate based on a localized form of integration by parts formula.

*Remark.* If the property  $(\mathbf{B}_q)$  (respectively  $(\mathbf{B}'_q)$ ) holds for a random variable  $F$ , then it also holds for  $F + x$  for every  $x$  in  $\mathbb{R}^d$ , with the same constant  $C_q$  (respectively  $C'_q$ ). In order to see this, given a test function  $f$ , one defines  $f_x(y) = f(x + y)$  so that  $f(F + x) = f_x(F)$ . And one notice that the infinite norm of  $f_x$  is the same as the infinite norm of  $f$ .

We give now a regularization lemma which plays a crucial role in our paper. We consider the  $d$ -dimensional super kernel  $\phi_\delta$  in (18) and (19) and we denote

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

Then we have the following regularization lemma.

**Lemma 3.5.** *We fix some  $q, d \in \mathbb{N}$  and  $\kappa, p \geq 1$ . We suppose that  $F \in \mathcal{D}_\infty^d$  such that  $\|F\|_{L, q+2, (4d+1)q} < \infty$ . We also consider an auxiliary random variable  $Q \in \mathcal{D}_\infty^d$  such that  $\Sigma_\kappa(Q) < \infty$ . Then there exists a constant  $C$  depending on  $p, q, \kappa$  and  $d$  (but not on  $Q$ ) such that for any  $\eta > 0$  and  $\delta > 0$ , for any function  $f \in C_b^q(\mathbb{R}^d)$ , we have*

$$|\mathbb{E}(f(F)) - \mathbb{E}(f_\delta(F))| \leq C \|f\|_\infty \times \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p) + \eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}) \right). \quad (38)$$

*Remark.* We remark that we do not assume the non-degeneracy condition for  $F$ , but we need to assume that we have another random variable  $Q$  which is non-degenerated such that  $\det \sigma_Q$  is close to  $\det \sigma_F$ . Then we obtain the regularization lemma (38). The regularization lemma here is originally from the paper [7].

*Remark.* If the property (38) holds for a random variable  $F$ , then it also holds for  $F + x$  for every  $x$  in  $\mathbb{R}^d$ , with the same constant  $C$ .

*Proof.* We denote

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

with  $y^\alpha = \prod_{i=1}^q y_{\alpha_i}$  for  $\alpha = (\alpha_1, \dots, \alpha_q)$ . Notice that if  $F$  satisfies  $(B'_q)$  with  $G = 1$ , then

$$|\mathbb{E}(R_q(\delta, F) \Upsilon_\eta(\det \sigma_F))| \leq C'_q \frac{\|f\|_\infty}{\eta^{2q}} \int_{\mathbb{R}^d} dy \phi_\delta(y) |y|^q = C'_q \int_{\mathbb{R}^d} \phi(y) |y|^q dy \|f\|_\infty \frac{\delta^q}{\eta^{2q}}. \quad (39)$$

We use a development in Taylor series of order  $q$  in order to get

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

Here we have used the property of a super kernel:  $\int_{\mathbb{R}^d} y^\beta \phi(y) dy = 0$ ,  $\forall |\beta| \leq q$ . Using (39), we have

$$|\mathbb{E}(f(F) \Upsilon_\eta(\det \sigma_F)) - \mathbb{E}(f_\delta(F) \Upsilon_\eta(\det \sigma_F))| \leq C \|f\|_\infty \frac{\delta^q}{\eta^{2q}}. \quad (40)$$

Following the idea from [11] p14, we denote

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

For an arbitrary  $\eta$ , we write

$$\mathbb{P}(\det \sigma_F < \eta) \leq \mathbb{P}(\det \sigma_F < \eta, |R| < \frac{1}{4}) + \mathbb{P}(|R| \geq \frac{1}{4}). \quad (41)$$

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

$$\mathbb{P}(\det \sigma_F < \eta, |R| < \frac{1}{4}) \leq \mathbb{P}(\det \sigma_Q < 2\eta) \leq 2^\kappa \eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}). \quad (42)$$



For any  $\eta > 0$ ,  $\kappa \in \mathbb{N}$ , we write

$$\begin{aligned} \mathbb{P}(|R| \geq \frac{1}{4}) &= \mathbb{P}(|\det \sigma_F - \det \sigma_Q| \geq \frac{1}{4} \det \sigma_Q) \\ &\leq \mathbb{P}(\det \sigma_Q \leq \eta) + \mathbb{P}(|\det \sigma_F - \det \sigma_Q| > \frac{1}{4} \eta) \\ &\leq C(\eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}) + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p)). \end{aligned} \quad (43)$$

So we conclude that

$$\mathbb{P}(\det \sigma_F < \eta) \leq C(\eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}) + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p)). \quad (44)$$

Then we have

$$|\mathbb{E}((1 - \Upsilon_\eta(\det \sigma_F))f(F))| \leq \|f\|_\infty \mathbb{P}(\det \sigma_F < \eta) \leq C\|f\|_\infty (\eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}) + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p)). \quad (45)$$

Similarly, we also have

$$|\mathbb{E}((1 - \Upsilon_\eta(\det \sigma_F))f_\delta(F))| \leq C\|f\|_\infty (\eta^\kappa \mathbb{E}(|\det \sigma_Q|^{-\kappa}) + \eta^{-p} \mathbb{E}(|\det \sigma_F - \det \sigma_Q|^p)). \quad (46)$$

We conclude by combining (40), (45) and (46).  $\square$

## 4 Application for jump equations

### 4.1 Basic notations and the main equation

To begin, we introduce some notations which will be used in the following sections. For a multi-index  $\beta$ , we denote  $|\beta|$  to be the length of  $\beta$ . 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  $\|f\|_{l,\infty} := \sum_{|\beta| \leq l} \|\partial^\beta f\|_\infty$  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

$$W_1(\rho_1, \rho_2) = \sup_{\text{Lip}(f) \leq 1} \left| \int_{\mathbb{R}^d} f(x) \rho_1(dx) - \int_{\mathbb{R}^d} f(x) \rho_2(dx) \right|, \quad (47)$$

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

$$d_{TV}(\rho_1, \rho_2) = \sup_{\|f\|_\infty \leq 1} \left| \int_{\mathbb{R}^d} f(x) \rho_1(dx) - \int_{\mathbb{R}^d} f(x) \rho_2(dx) \right|. \quad (48)$$

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  $G$ ). We refer to [49] and [34] the basic properties of these distances. 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

$$X_t = x + \int_0^t b(X_r) dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{r-}) N(dz, dr), \quad (49)$$

where  $N(dz, dr)$  is a Poisson point measure on the state space  $\mathbb{R}^d$  with intensity measure  $\hat{N}(dz, dr) = \mu(dz)dr$ ,  $x$  is the initial value,  $\mu$  is a positive  $\sigma$ -finite measure on  $\mathbb{R}^d$ , and  $b : \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ .

## 4.2 Hypotheses

Here we give our hypotheses.

**Hypothesis 2.1 (Regularity)** We assume that the function  $x \mapsto b(x)$  is infinitely differentiable with bounded derivatives of any orders. We also assume that the function  $(z, x) \mapsto c(z, x)$  is infinitely differentiable and for every multi-indices  $\beta_1, \beta_2$ , there exists a function  $\bar{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  depending on  $\beta_1, \beta_2$  such that we have

$$\sup_{x \in \mathbb{R}^d} (|c(z, x)| + |\partial_z^{\beta_2} \partial_x^{\beta_1} c(z, x)|) \leq \bar{c}(z), \quad \forall z \in \mathbb{R}^d, \quad (50)$$

$$\text{with } \int_{\mathbb{R}^d} |\bar{c}(z)|^p \mu(dz) := \bar{c}_p < \infty, \quad \forall p \geq 1. \quad (51)$$

*Remark.* We will use several times the following consequence of (51) and of Burkholder inequality (see for example the Theorem 2.11 in [28], see also [29]): Let  $\Phi(s, z, \omega) : [0, T] \times \mathbb{R}^d \times \Omega \rightarrow \mathbb{R}_+$  and  $\varphi(s, \omega) : [0, T] \times \Omega \rightarrow \mathbb{R}_+$  be two non-negative functions. The Burkholder inequality states that for any  $p \geq 2$ , we have

$$\begin{aligned} & \mathbb{E} \left| \int_0^t \int_{\mathbb{R}^d} \Phi(s, z, \omega) N(dz, ds) \right|^p \\ & \leq C \left[ \mathbb{E} \left( \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)|^2 \mu(dz) ds \right)^{\frac{p}{2}} + \mathbb{E} \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)|^p \mu(dz) ds \right. \\ & \quad \left. + \mathbb{E} \left| \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)| \mu(dz) ds \right|^p \right]. \end{aligned} \quad (52)$$

If we have

$$|\Phi(s, z, \omega)| \leq |\bar{c}(z)| |\varphi(s, \omega)|,$$

then for any  $p \geq 2$ ,

$$\mathbb{E} \left| \int_0^t \int_{\mathbb{R}^d} \Phi(s, z, \omega) N(dz, ds) \right|^p \leq C \mathbb{E} \int_0^t |\varphi(s, \omega)|^p ds, \quad (53)$$

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

*Proof.* By compensating  $N$  and using Burkholder inequality and (51), we have

$$\begin{aligned} & \mathbb{E} \left| \int_0^t \int_{\mathbb{R}^d} \Phi(s, z, \omega) N(dz, ds) \right|^p \\ & \leq C \left[ \mathbb{E} \left( \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)|^2 \mu(dz) ds \right)^{\frac{p}{2}} + \mathbb{E} \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)|^p \mu(dz) ds \right. \\ & \quad \left. + \mathbb{E} \left| \int_0^t \int_{\mathbb{R}^d} |\Phi(s, z, \omega)| \mu(dz) ds \right|^p \right] \\ & \leq C \mathbb{E} \int_0^t |\varphi(s, \omega)|^p ds. \end{aligned}$$

□

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  $\check{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that  $\int_{\mathbb{R}^d} |\check{c}(z)|^p \mu(dz) := \check{c}_p < \infty, \forall p \geq 1$ , and

$$\|\nabla_x c(z, x)(I_d + \nabla_x c(z, x))^{-1}\| \leq \check{c}(z), \quad \forall x \in \mathbb{R}^d, z \in \mathbb{R}^d,$$

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

*Remark.* We need this hypothesis to prove the regularity of the inverse tangent flow (see Section 5.2).

**Hypothesis 2.3 (Ellipticity)** There exists a non-negative function  $\underline{c} : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that for every  $x \in \mathbb{R}^d, z \in \mathbb{R}^d, \zeta \in \mathbb{R}^d$ , we have

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

*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  $\mu$ .

a) We assume that

$$\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \bar{\mu} \left\{ \underline{c} \geq \frac{1}{u} \right\} = \infty, \quad (54)$$

with

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

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 5.2).

*Remark.* If  $\mu(\mathbb{R}^d) < \infty$ , then  $\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \bar{\mu} \left\{ \underline{c} \geq \frac{1}{u} \right\} = 0$ . So (54) implies that  $\mu(\mathbb{R}^d) = \infty$ .

b) We assume that  $\mu$  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.

**Hypothesis 2.5**

We give some conditions which ensure the existence and uniqueness of the invariant measure and the "exponential Lipschitz property" (4).

Suppose that

$$\begin{aligned} i) \quad \langle x - y, b(x) - b(y) \rangle &\leq -\bar{b} |x - y|^2 \\ ii) \quad |c(z, x) - c(z, y)| &\leq \bar{c}(z) |x - y| \end{aligned} \quad (55)$$

and

$$iii) \quad 2\bar{b} - \int_{\mathbb{R}^d} (2\bar{c}(z) + \bar{c}^2(z)) \mu(dz) := \theta > 0. \quad (56)$$

**Hypothesis 2.6**

We assume that  $\mathcal{P}$  is a partition with decreasing time steps:  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{n-1} < \Gamma_n < \dots\}$ . We denote  $\gamma_n = \Gamma_n - \Gamma_{n-1}$ ,  $n \in \mathbb{N}$  and assume that  $\gamma_n \downarrow 0$ . We also introduce

$$\bar{\omega} = \overline{\lim}_{n \rightarrow \infty} \frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2},$$

and assume that  $\bar{\omega} < \frac{\theta}{2}$ , with  $\theta$  given in (56).

*Remark.* A typical example is  $\gamma_n = \frac{1}{n}$  and so  $\bar{\omega} = 1$ .

### 4.3 The truncated Euler scheme

Now we construct the Euler scheme. For some technical reasons, we take a general partition  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{n-1} < \Gamma_n < \dots\}$  (without assuming **Hypothesis 2.6** at this moment). We denote  $\gamma_n = \Gamma_n - \Gamma_{n-1}$ ,  $n \in \mathbb{N}$  and denote  $|\mathcal{P}| := \max_{n \in \mathbb{N}} (\Gamma_{n+1} - \Gamma_n)$ . We assume that  $|\mathcal{P}| \leq 1$ , and

$$\sum_{i=1}^{\infty} \gamma_i = \lim_{n \rightarrow \infty} \Gamma_n = \infty.$$

For  $\Gamma_n \leq t < \Gamma_{n+1}$  we denote  $N(t) = n$  and  $\tau(t) = \Gamma_n$ . We consider the Euler scheme:

$$X_t^{\mathcal{P}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}}) dr + \int_0^t \int_{\mathbb{R}^d} c(z, X_{\tau(r)-}^{\mathcal{P}}) N(dz, dr). \quad (57)$$

Since we have  $\mu(\mathbb{R}^d) = \infty$  (which is a consequence of (54)), 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. We construct the truncated Euler scheme as below. To begin, we give some notations.

We denote

$$\varepsilon_m := \int_{\{|z| > m\}} |\bar{c}(z)|^2 \mu(dz) + \left| \int_{\{|z| > m\}} \bar{c}(z) \mu(dz) \right|^2, \quad \forall m \in \mathbb{N}. \quad (58)$$

For every  $\gamma > 0$ , we define the truncation function  $M(\gamma) \in \mathbb{N}$  to be the smallest integer such that

$$\varepsilon_{M(\gamma)} \leq \gamma^2. \quad (59)$$

For  $m \in \mathbb{N}$ , we denote  $B_m = \{z \in \mathbb{R}^d : |z| \leq m\}$ . For  $\Gamma_{n-1} < t \leq \Gamma_n$ , we denote  $M_{\mathcal{P}}(t) = M(\gamma_n)$ . We remark that we have  $\lim_{\gamma \rightarrow 0} M(\gamma) = \infty$  and for  $\Gamma_{n-1} < t \leq \Gamma_n$ , we have  $M_{\mathcal{P}}(t) = M(\gamma_n) \geq M(|\mathcal{P}|) \rightarrow \infty$ , as  $|\mathcal{P}| \rightarrow 0$ . Now we discard the "big jumps" (the jumps of size  $|z| > M_{\mathcal{P}}(t)$ ):

$$X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \int_0^t \int_{B_{M_{\mathcal{P}}(r)}} c(z, X_{\tau(r)-}^{\mathcal{P}, M_{\mathcal{P}}}) N(dz, dr). \quad (60)$$

The advantage of considering  $X_t^{\mathcal{P}, M_{\mathcal{P}}}$  is that we may represent it by means of compound Poisson processes. For  $k \in \mathbb{N}$ , we denote  $I_1 = B_1$ ,  $I_k = B_k \setminus B_{k-1}$  for  $k \geq 2$  and take  $(J_t^k)_{t \geq 0}$  a Poisson process of intensity  $\mu(I_k)$ . We denote by  $(T_i^k)_{i \in \mathbb{N}}$  the jump times of  $(J_t^k)_{t \geq 0}$  and we consider a sequences of independent random variables  $Z_i^k \sim \mathbb{1}_{I_k}(z) \frac{\mu(dz)}{\mu(I_k)}$ ,  $k, i \in \mathbb{N}$ . Moreover,  $(J_t^k)_{t \geq 0}$  and  $(Z_i^k)_{k, i \in \mathbb{N}}$  are taken to be independent. Then we represent the jump's part of the equation (60) by compound Poisson processes. We write

$$\begin{aligned} X_t^{\mathcal{P}, M_{\mathcal{P}}} &= x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{k=1}^{\infty} \sum_{i=1}^{J_t^k} \mathbb{1}_{B_{M_{\mathcal{P}}(T_i^k)}}(Z_i^k) c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}) \\ &= x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{k=1}^{\infty} \sum_{i=1}^{J_t^k} \sum_{n=0}^{N(t)} \mathbb{1}_{B_{M(\gamma_{n+1})}}(Z_i^k) c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}). \end{aligned}$$

Since  $Z_i^k \in B_k \setminus B_{k-1}$ , it follows that  $Z_i^k \in B_{M(\gamma_{n+1})}$  is equivalent to  $k \leq M(\gamma_{n+1})$ . Then

$$X_t^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_0^t b(X_{\tau(r)}^{\mathcal{P}, M_{\mathcal{P}}}) dr + \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} c(Z_i^k, X_{\tau(T_i^k)-}^{\mathcal{P}, M_{\mathcal{P}}}). \quad (61)$$

We remark that the solution of the equation (61) can be constructed in an explicit way.

We recall the notation  $\theta$  in **Hypothesis 2.5**. We also recall  $n_{\rho} = n_{\frac{\theta}{2}}$  in (10) (with  $\rho = \frac{\theta}{2}$  in our case) and  $n_*$  in (11). We obtain the following error estimate for  $X_t^{\mathcal{P}, M_{\mathcal{P}}}$ , which represents the main result in our paper.

**Theorem 4.1.** Assume that **Hypothesis 2.1~2.5** hold and the partition  $\mathcal{P}$  satisfies **Hypothesis 2.6**. Then an invariant probability measure  $\nu$  exists and is unique, and for  $n > \max\{n_{\frac{\theta}{2}} + 3, n_* + 3\}$ , for any  $\varepsilon > 0$  there exists a constant  $C_\varepsilon$  such that

$$d_{TV}(\mathcal{L}(X_{\Gamma_n}^{\mathcal{P}, M_{\mathcal{P}}}), \nu) \leq C_\varepsilon(\gamma_n^{1-\varepsilon} + \int_{\mathbb{R}^d} |x - y| d\nu(y) e^{-\frac{\theta}{2}\Gamma_n}). \quad (62)$$

The proof of this theorem will be given in Section 6 by using some Malliavin integration by parts techniques introduced in Section 5.

In order to apply the Malliavin framework which will be presented in Section 5, we introduce additionally an auxiliary equation as follows (see (64) below).

For  $\Gamma_n < t \leq \Gamma_{n+1}$ , we define

$$a_t^{\mathcal{P}} = \left( \sum_{1 \leq i \leq n} \gamma_i \int_{\{|z| \geq M(\gamma_i)\}} \underline{c}(z) \mu(dz) + (t - \Gamma_n) \int_{\{|z| \geq M(\gamma_{n+1})\}} \underline{c}(z) \mu(dz) \right)^{\frac{1}{2}}, \quad (63)$$

where  $\underline{c}$  is given in **Hypothesis 2.3**. We notice that  $|a_t^{\mathcal{P}}| \leq \sqrt{t \times \varepsilon_{M(|\mathcal{P}|)}} \leq \sqrt{t} \times |\mathcal{P}|$ .

Now we cancel the big jumps in equation (49) and replace them by a ( $d$ -dimensional) Gaussian random variable  $\Delta$  which is independent of the Poisson point measure  $N(dz, ds)$ :

$$X_t^{M_{\mathcal{P}}} = x + a_t^{\mathcal{P}} \Delta + \int_0^t b(X_s^{M_{\mathcal{P}}}) ds + \int_0^t \int_{B_{M_{\mathcal{P}}}(s)} c(z, X_{s-}^{M_{\mathcal{P}}}) N(dz, ds). \quad (64)$$

We remark that  $\Delta$  is necessary in order to obtain the non degeneracy of the covariance matrix (see Section 5.2 for details).

Following the same idea as above, we represent the jump's parts of the equation (64) by compound Poisson processes:

$$X_t^{M_{\mathcal{P}}} = x + a_t^{\mathcal{P}} \Delta + \int_0^t b(X_s^{M_{\mathcal{P}}}) ds + \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} c(Z_i^k, X_{T_i^k-}^{M_{\mathcal{P}}}). \quad (65)$$

We sometimes write  $X_t^{\mathcal{P}, M_{\mathcal{P}}}(x)$  (resp.  $X_t^{M_{\mathcal{P}}}(x)$ ,  $X_t(x)$ ) instead of  $X_t^{\mathcal{P}, M_{\mathcal{P}}}$  (resp.  $X_t^{M_{\mathcal{P}}}$ ,  $X_t$ ) to stress the dependence on the initial value  $x$ .

## 4.4 Some examples

We give some typical examples to illustrate our main results.

**Example 1** We take  $h = 1$  so the measure  $\mu$  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

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

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

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

We notice that  $\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \bar{\mu}\{\underline{c} \geq \frac{1}{u}\} = \infty$  when  $0 < p < d$ . Therefore, when  $p \geq d$ , we can say nothing; when  $0 < p < d$ , the results in **Theorem 4.1** are true.

**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

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

so that

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

We notice that in this case,  $\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \bar{\mu}\{\underline{c} \geq \frac{1}{u}\} = \infty$ . Thus, the results in **Theorem 4.1** hold true.

**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

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

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  $\sigma'$  is the differential of  $\sigma$  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  $\mu(dz) = \mathbb{1}_{\{|z| \geq 1\}} \frac{1}{|z|^{1-\alpha}} dz$ . We only check **Hypothesis 2.4** here. In this case,  $\underline{c}(z) = \underline{\sigma} \times \frac{1}{|z|^4}$ , then

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

so that  $\lim_{u \rightarrow +\infty} \frac{1}{\ln u} \bar{\mu}\{\underline{c} \geq \frac{1}{u}\} = \infty$ . Thus we can apply **Theorem 4.1**.

## 5 Malliavin framework for jump equations

We take time  $t \in [0, 3]$  throughout this section and we use the notations from Section 4. We recall  $(X_t)_{t \in [0, 3]}$  in (49),  $(X_t^{\mathcal{P}, M_{\mathcal{P}}})_{t \in [0, 3]}$  in (60) and  $(X_t^{M_{\mathcal{P}}})_{t \in [0, 3]}$  in (64), where  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{N(3)} \leq 3\}$  is a general partition (which is not supposed to verify **Hypothesis 2.6**).

**Lemma 5.1.** *Suppose that **Hypothesis 2.1** holds true. Then we have the followings.*

i) *For every  $t \in [0, 3]$ , we have*

$$\mathbb{E}|X_t^{\mathcal{P}, M_{\mathcal{P}}} - X_t| \rightarrow 0, \text{ as } |\mathcal{P}| \rightarrow 0;$$

ii) *For every fixed  $t \in [0, 3]$  and every  $p \geq 2$ , we have*

$$\mathbb{E}|X_t^{M_{\mathcal{P}}} - X_t|^p \rightarrow 0, \text{ as } |\mathcal{P}| \rightarrow 0;$$

iii) *For every fixed  $t \in [0, 3]$  and every multi-index  $\beta$ , we have*

$$\mathbb{E}|\partial_x^\beta X_t^{M_{\mathcal{P}}} - \partial_x^\beta X_t| \rightarrow 0, \text{ as } |\mathcal{P}| \rightarrow 0.$$

*Proof.* The proof of this lemma is standard and straightforward by Gronwall lemma and Buckholder inequality. So we leave it out.  $\square$

Now we use Malliavin calculus for  $X_t^{\mathcal{P}, M_{\mathcal{P}}}$ ,  $X_t^{M_{\mathcal{P}}}$  and  $X_t$ . There are several approaches given in [12], [20], [26], [27], [37], [48] and [50] for example. Here we give a framework analogous to [9].

To begin we define a regularization function.

$$a(y) = 1 - \frac{1}{1 - (4y - 1)^2} \quad \text{for } y \in [\frac{1}{4}, \frac{1}{2}), \quad (66)$$

$$\psi(y) = \mathbb{1}_{\{|y| \leq \frac{1}{4}\}} + \mathbb{1}_{\{\frac{1}{4} < |y| \leq \frac{1}{2}\}} e^{a(|y|)}. \quad (67)$$

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

$$\Psi_k(y) = \psi(|y| - (k - \frac{1}{2})), \quad \forall k \in \mathbb{N}. \quad (68)$$

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

$$\sup_{k \in \mathbb{N}} \|\Psi_k\|_{l, \infty} \leq C_l < \infty. \quad (69)$$

We focus on  $X_t^{\mathcal{P}, M_{\mathcal{P}}}(x)$  and  $X_t^{M_{\mathcal{P}}}(x)$  (solutions of (61) and (65)) which are functions of random variables  $T_i^k$ ,  $Z_i^k$  and  $\Delta$ .

Now we introduce the space of simple functionals  $\mathcal{S}$ . We take  $\mathcal{G} = \sigma(T_i^k : k, i \in \mathbb{N})$  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  $Z_i^k = (Z_{i,1}^k, \dots, Z_{i,d}^k)$ ,  $k, i \in \mathbb{N}$  and  $\Delta = (\Delta_1, \dots, \Delta_d)$ . We denote by  $C_{\mathcal{G}, p}$  the space of the functions  $f : \Omega \times \mathbb{R}^{m \times m' \times d + d} \rightarrow \mathbb{R}$  such that for each  $\omega$ , the function  $(z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d) \mapsto f(\omega, z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$  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  $(z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$ , the function  $\omega \mapsto f(\omega, z_{1,1}^1, \dots, z_{m,d}^{m'}, \delta_1, \dots, \delta_d)$  is  $\mathcal{G}$ -measurable. And we consider the weights

$$\xi_i^k = \Psi_k(Z_i^k). \quad (70)$$

Then we define the space of simple functionals

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

*Remark.* Take  $m' = \max_{t \leq 3} M_{\mathcal{P}}(t)$  and  $m = \max_{k \leq m'} J_t^k$ . Then  $X_t^{M_{\mathcal{P}}}(x)$  (solution of (65)) is a function of  $T_i^k$ ,  $Z_i^k$  and of  $\Delta$ , with  $k \leq m'$  and  $i \leq m$ . So it is a simple functional (the same for  $X_t^{\mathcal{P}, M_{\mathcal{P}}}(x)$  (solution of (61))).

On the space  $\mathcal{S}$ , for  $t \geq 1$ , we define the derivative operator  $DF = (D^Z F, D^\Delta F)$ , where

$$\begin{aligned} D_{(\bar{k}, \bar{i}, \bar{j})}^Z F &= \xi_{\bar{i}}^{\bar{k}} \frac{\partial f}{\partial z_{\bar{i}, \bar{j}}^{\bar{k}}}(\omega, (Z_i^k)_{\substack{1 \leq k \leq m' \\ 1 \leq i \leq m}}, \Delta), \quad \bar{k}, \bar{i} \in \mathbb{N}, \bar{j} \in \{1, \dots, d\}, \\ D_{\tilde{j}}^\Delta F &= \frac{\partial f}{\partial \delta_{\tilde{j}}}(\omega, (Z_i^k)_{\substack{1 \leq k \leq m' \\ 1 \leq i \leq m}}, \Delta), \quad \tilde{j} \in \{1, \dots, d\}. \end{aligned} \quad (71)$$

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

$$\langle DF, DG \rangle_{l_2 \times \mathbb{R}^d} = \sum_{j=1}^d D_j^\Delta F \times D_j^\Delta G + \sum_{k=1}^\infty \sum_{i=1}^\infty \sum_{j=1}^d D_{(k,i,j)}^Z F \times D_{(k,i,j)}^Z G. \quad (72)$$

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

We recall that  $\mu(dz) = h(z)dz$  with  $h \in C^\infty(\mathbb{R}^d)$  (see **Hypothesis 2.4 b**). We define the Ornstein-Uhlenbeck operator  $LF = L^Z F + L^\Delta F$  with

$$\begin{aligned} L^Z F &= - \sum_{k=1}^{m'} \sum_{i=1}^m \sum_{j=1}^d (\partial_{z_{i,j}^k} (\xi_i^k D_{(k,i,j)}^Z F) + D_{(k,i,j)}^Z F \times D_{(k,i,j)}^Z \ln[h(Z_i^k)]), \\ L^\Delta F &= \sum_{j=1}^d D_j^\Delta F \times \Delta_j - \sum_{j=1}^d D_j^\Delta D_j^\Delta F. \end{aligned} \quad (73)$$

One can check that the triplet  $(S, D, L)$  is consistent with the IbP framework given in Section 3.1. In particular the duality formula (28) 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 (32)).

We recall  $X_t^{\mathcal{P}, M_{\mathcal{P}}}(x)$  in (61),  $X_t^{M_{\mathcal{P}}}(x)$  in (65) and  $X_t(x)$  in (49). We denote

$$F_t^{\mathcal{P}, M_{\mathcal{P}}}(x) = X_t^{\mathcal{P}, M_{\mathcal{P}}}(x) - x, F_t^{M_{\mathcal{P}}}(x) = X_t^{M_{\mathcal{P}}}(x) - x \text{ and } F_t(x) = X_t(x) - x \quad (74)$$

In the following subsections, we will give some lemmas concerning the Sobolev norms and the covariance matrices. We recall (see (26)) that  $\sigma_F$  denotes the covariance matrix of  $F$ , and recall the Sobolev norms defined in (29) and (30).

## 5.1 Sobolev norms

We recall the notations  $F_t^{\mathcal{P}, M_{\mathcal{P}}}(x)$ ,  $F_t^{M_{\mathcal{P}}}(x)$  and  $F_t(x)$  in (74).

**Lemma 5.2.** *Assuming **Hypothesis 2.1** and **Hypothesis 2.4 b)**, for all  $p \geq 1, l \geq 0$ , there exists a constant  $C_{l,p}$  depending on  $l, p, d$ , such that for any  $t \in [0, 3]$ ,*

$$i) \quad \sup_{\mathcal{P}} \sup_x (\|F_t^{\mathcal{P}, M_{\mathcal{P}}}(x)\|_{L,l,p} + \|F_t^{M_{\mathcal{P}}}(x)\|_{L,l,p}) \leq C_{l,p}.$$

Moreover,  $F_t(x)$  belongs to  $\mathcal{D}_\infty^d$  and

$$ii) \quad \sup_x \|F_t(x)\|_{L,l,p} \leq C_{l,p}.$$

For all  $p, q \geq 1, l \geq 0$ , there exists a constant  $C_{l,p,q}$  depending on  $l, p, q, d$ , such that for every multi-index  $\beta$  with  $|\beta| = q$ , we also have

$$iii) \quad \sup_x \|\partial_x^\beta (X_t(x))\|_{l,p} \leq C_{l,p,q}.$$

*Remark.* Since  $Dx = 0, \forall x \in \mathbb{R}^d$ , we also have

$$\sup_{\mathcal{P}} \sup_x (\mathbb{E}|X_t^{\mathcal{P}, M_{\mathcal{P}}}(x)|_{1,l}^p + \mathbb{E}|X_t^{M_{\mathcal{P}}}(x)|_{1,l}^p + \mathbb{E}|X_t(x)|_{1,l}^p) \leq C_{l,p}.$$

*Proof.* We first notice that for any  $l, p$ ,  $\sup_{\mathcal{P}} \sup_x (\|F_t^{\mathcal{P}, M_{\mathcal{P}}}(x)\|_{L,l,p} + \|F_t^{M_{\mathcal{P}}}(x)\|_{L,l,p}) \leq C_{l,p}$ . This is a slight variant of the proof of Lemma 3.7 i) in [44]. The difference is that the truncation function  $M$  is constant in [44] while here it depends on the time. But this does not change anything. In a similar way, for every multi-index  $\beta$  with  $|\beta| = q$ , we have  $\sup_{\mathcal{P}} \sup_x \|\partial_x^\beta (X_t^{M_{\mathcal{P}}}(x))\|_{l,p} \leq C_{l,p,q}$ .

Afterwards, we consider an increasing sequence of partition  $\mathcal{P}_n, n \in \mathbb{N}, (\mathcal{P}_n \subset \mathcal{P}_{n+1})$ , such that  $|\mathcal{P}_n| \downarrow 0$ . In particular,  $\forall t, M_{\mathcal{P}_n}(t) \uparrow \infty$ . Noticing by **Lemma 5.1 ii)** that  $\mathbb{E}|F_t^{M_{\mathcal{P}_n}} - F_t| \rightarrow 0$  as  $n \rightarrow \infty$ , and applying **Lemma 3.3 (A)** with  $F_n = F_t^{M_{\mathcal{P}_n}}$  and  $F = F_t$ , we get that  $F_t$  belongs to  $\mathcal{D}_\infty^d$  and  $\sup_x \|F_t(x)\|_{L,l,p} \leq C_{l,p}$ .

Furthermore, noticing by **Lemma 5.1 iii)** that  $\mathbb{E}|\partial_x^\beta X_t^{M_{\mathcal{P}_n}} - \partial_x^\beta X_t| \rightarrow 0$  as  $n \rightarrow \infty$ , and applying **Lemma 3.3 (A)** with  $F_n = \partial_x^\beta X_t^{M_{\mathcal{P}_n}}$  and  $F = \partial_x^\beta X_t$ , we obtain that  $\partial_x^\beta X_t$  belongs to  $\mathcal{D}_\infty^d$  and  $\sup_x \|\partial_x^\beta (X_t(x))\|_{l,p} \leq C_{l,p,q}$ .

□



## 5.2 Covariance matrix

**Lemma 5.3.** Assume that **Hypothesis 2.1, 2.2, 2.3 and 2.4** hold true. We denote the lowest eigenvalue of the Malliavin covariance matrix  $\sigma_{X_t^{M\mathcal{P}}}$  by  $\lambda_t^{M\mathcal{P}}$ . Then for every  $p \geq 1$ ,  $1 \leq t \leq 3$ , we have

$$\begin{aligned} i) \quad & \sup_{\mathcal{P}} \sup_x \mathbb{E}(1/\det \sigma_{X_t^{M\mathcal{P}}}(x))^p \leq \sup_{\mathcal{P}} \sup_x \mathbb{E}(|\lambda_t^{M\mathcal{P}}|^{-dp}) \leq C_p, \\ ii) \quad & \sup_x \mathbb{E}(1/\det \sigma_{X_t(x)})^p \leq C_p, \end{aligned}$$

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

*Remark.* We recall the notations  $F_t^{M\mathcal{P}}(x) = X_t^{M\mathcal{P}}(x) - x$  and  $F_t(x) = X_t(x) - x$ . Since  $Dx = 0$ ,  $\forall x \in \mathbb{R}^d$ , the above results are equivalent to

$$\begin{aligned} i) \quad & \sup_{\mathcal{P}} \sup_x \mathbb{E}(1/\det \sigma_{F_t^{M\mathcal{P}}}(x))^p \leq \sup_{\mathcal{P}} \sup_x \mathbb{E}(|\lambda_t^{M\mathcal{P}}|^{-dp}) \leq C_p, \\ ii) \quad & \sup_x \mathbb{E}(1/\det \sigma_{F_t(x)})^p \leq C_p. \end{aligned}$$

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

**Step 1** We notice by the definition (71) that for any  $k_0, i_0 \in \mathbb{N}, j \in \{1, \dots, d\}$ ,

$$\begin{aligned} D_{(k_0, i_0, j)}^Z X_t^{M\mathcal{P}} &= \int_{T_{i_0}^{k_0}}^t \nabla_x b(X_r^{M\mathcal{P}}) D_{(k_0, i_0, j)}^Z X_r^{M\mathcal{P}} dr \\ &+ \sum_{n=0}^{N(t)} \mathbb{1}_{\{\Gamma_n < T_{i_0}^{k_0} \leq \Gamma_{n+1} \wedge t\}} \mathbb{1}_{\{1 \leq k_0 \leq M(\gamma_{n+1})\}} \xi_{i_0}^{k_0} \partial_{z_{i_0, j}^{k_0}} c(Z_{i_0}^{k_0}, X_{T_{i_0}^{k_0}-}^{M\mathcal{P}}) \\ &+ \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n \vee T_{i_0}^{k_0} < T_i^k \leq \Gamma_{n+1} \wedge t\}} \nabla_x c(Z_i^k, X_{T_i^k-}^{M\mathcal{P}}) D_{(k_0, i_0, j)}^Z X_{T_i^k-}^{M\mathcal{P}}, \end{aligned} \quad (75)$$

$$D_j^\Delta X_t^{M\mathcal{P}} = a_t^{\mathcal{P}} e_j + \int_0^t \nabla_x b(X_r^{M\mathcal{P}}) D_j^\Delta X_r^{M\mathcal{P}} dr + \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \nabla_x c(Z_i^k, X_{T_i^k-}^{M\mathcal{P}}) D_j^\Delta X_{T_i^k-}^{M\mathcal{P}}, \quad (76)$$

where  $e_j = (0, \dots, 0, 1, 0, \dots, 0)$  with value 1 at the  $j$ -th component.

Now we introduce  $(Y_t^{M\mathcal{P}})_{t \geq 0}$  (this is so-called the tangent flow) which is the matrix solution of the linear equation

$$Y_t^{M\mathcal{P}} = I_d + \int_0^t \nabla_x b(X_r^{M\mathcal{P}}) Y_r^{M\mathcal{P}} dr + \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \nabla_x c(Z_i^k, X_{T_i^k-}^{M\mathcal{P}}) Y_{T_i^k-}^{M\mathcal{P}}.$$

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

$$\tilde{Y}_t^{M\mathcal{P}} = I_d - \int_0^t \tilde{Y}_r^{M\mathcal{P}} \nabla_x b(X_r^{M\mathcal{P}}) dr - \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \tilde{Y}_{T_i^k-}^{M\mathcal{P}} \nabla_x c(I_d + \nabla_x c)^{-1}(Z_i^k, X_{T_i^k-}^{M\mathcal{P}}). \quad (77)$$

*Remark.* We notice that  $Y_t^{M\mathcal{P}} = \nabla_x(X_t^{M\mathcal{P}}(x))$ . If instead we consider the gradient of the Euler scheme  $Y_t^{\mathcal{P}, M\mathcal{P}} = \nabla_x(X_t^{\mathcal{P}, M\mathcal{P}}(x))$ , the matrix  $Y_t^{\mathcal{P}, M\mathcal{P}}$  is not invertible, and this is a specific difficulty when we deal with the Euler scheme. This is why we have to work with  $X_t^{M\mathcal{P}}$  only.

Applying **Hypothesis 2.1** and **Hypothesis 2.2**, one also has

$$\mathbb{E}(\sup_{0 < t \leq 2} (\|Y_t^{M\mathcal{P}}\|^p + \|\tilde{Y}_t^{M\mathcal{P}}\|^p)) \leq C_p < \infty. \quad (78)$$

The proof of (78) is straightforward and we leave it out.

Then using the uniqueness of solution to the equation (75) and (76), one obtains

$$D_{(k,i,j)}^Z X_t^{M\mathcal{P}} = \sum_{n=0}^{N(t)} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \mathbb{1}_{\{1 \leq k \leq M(\gamma_{n+1})\}} \xi_i^k Y_t^{M\mathcal{P}} \tilde{Y}_{T_i^k}^{M\mathcal{P}} \partial_{z_{i,j}^k} c(Z_i^k, X_{T_i^k}^{M\mathcal{P}}), \quad (79)$$

and  $D_j^\Delta X_t^{M\mathcal{P}} = a_t^{\mathcal{P}} Y_t^{M\mathcal{P}} e_j$ .

We recall that we denote the lowest eigenvalue of the Malliavin covariance matrix  $\sigma_{X_t^{M\mathcal{P}}}$  by  $\lambda_t^{M\mathcal{P}}$ . Then we have (recalling the definitions (26) and (72))

$$\lambda_t^{M\mathcal{P}} = \inf_{|\zeta|=1} \langle \sigma_{X_t^{M\mathcal{P}}} \zeta, \zeta \rangle \geq \inf_{|\zeta|=1} \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \sum_{j=1}^d \langle D_{(k,i,j)}^Z X_t^{M\mathcal{P}}, \zeta \rangle^2 + \inf_{|\zeta|=1} \sum_{j=1}^d \langle D_j^\Delta X_t^{M\mathcal{P}}, \zeta \rangle^2.$$

By (79),

$$\begin{aligned} \lambda_t^{M\mathcal{P}} &\geq \inf_{|\zeta|=1} \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} \sum_{j=1}^d |\xi_i^k|^2 \langle \partial_{z_{i,j}^k} c(Z_i^k, X_{T_i^k}^{M\mathcal{P}}), (Y_t^{M\mathcal{P}} \tilde{Y}_{T_i^k}^{M\mathcal{P}})^* \zeta \rangle^2 \\ &\quad + \inf_{|\zeta|=1} \sum_{j=1}^d |a_t^{\mathcal{P}}|^2 \langle e_j, (Y_t^{M\mathcal{P}})^* \zeta \rangle^2, \end{aligned}$$

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

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

So we deduce that

$$\lambda_t^{M\mathcal{P}} \geq \inf_{|\zeta|=1} \left( \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} |\xi_i^k|^2 \underline{c}(Z_i^k) \|(Y_t^{M\mathcal{P}} \tilde{Y}_{T_i^k}^{M\mathcal{P}})^* \zeta\|^2 + |a_t^{\mathcal{P}}|^2 \inf_{|\zeta|=1} \|(Y_t^{M\mathcal{P}})^* \zeta\|^2 \right).$$

For every invertible matrix  $A$  and every vector  $y$ , one has  $|Ay| \geq \frac{1}{\|A^{-1}\|} |y|$ , so that

$$\begin{aligned} \lambda_t^{M\mathcal{P}} &\geq \left( \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} |\xi_i^k|^2 \underline{c}(Z_i^k) \|\tilde{Y}_t^{M\mathcal{P}}\|^{-2} \|Y_{T_i^k}^{M\mathcal{P}}\|^{-2} \right) + |a_t^{\mathcal{P}}|^2 \|\tilde{Y}_t^{M\mathcal{P}}\|^{-2} \\ &\geq \left( \inf_{0 < t \leq 2} \|\tilde{Y}_t^{M\mathcal{P}}\|^{-2} \|Y_t^{M\mathcal{P}}\|^{-2} \right) \left( \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} |\xi_i^k|^2 \underline{c}(Z_i^k) \right) + |a_t^{\mathcal{P}}|^2. \end{aligned}$$

We denote

$$\chi_t^{M\mathcal{P}} = \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \sum_{i=1}^{J_t^k} \mathbb{1}_{\{\Gamma_n < T_i^k \leq \Gamma_{n+1} \wedge t\}} |\xi_i^k|^2 \underline{c}(Z_i^k). \quad (80)$$

By (78),  $(\mathbb{E} \sup_{0 \leq t \leq 2} \|\tilde{Y}_t^{M\mathcal{P}}\|^{4dp} \|Y_t^{M\mathcal{P}}\|^{4dp})^{1/2} \leq C_{d,p} < \infty$ , so that using Schwartz inequality, we have

$$\mathbb{E} \left| \frac{1}{\det \sigma_{X_t^{M\mathcal{P}}}} \right|^p \leq \mathbb{E} (|\lambda_t^{M\mathcal{P}}|^{-dp}) \leq C (\mathbb{E} (|\chi_t^{M\mathcal{P}} + |a_t^{\mathcal{P}}|^2|^{-2dp}))^{\frac{1}{2}}. \quad (81)$$

**Step 2** Since it is not easy to compute  $\mathbb{E} (|\chi_t^{M\mathcal{P}} + |a_t^{\mathcal{P}}|^2|^{-2dp})$  directly, we make the following argument where the idea comes originally from [12]. 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

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

which, by taking expectation, gives

$$\mathbb{E} \left( \frac{1}{|\chi_t^{M\mathcal{P}} + |a_t^{\mathcal{P}}|^2|^{2dp}} \right) = \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \mathbb{E} (e^{-s(\chi_t^{M\mathcal{P}} + |a_t^{\mathcal{P}}|^2)}) ds. \quad (82)$$

**Step 3** Now we compute  $\mathbb{E} (e^{-s(\chi_t^{M\mathcal{P}} + |a_t^{\mathcal{P}}|^2)})$  for any  $s > 0$ . We recall that  $I_1 = B_1$ ,  $I_k = B_k - B_{k-1}$ ,  $k \geq 2$  (given in Section 4.3), and  $\xi_i^k = \Psi_k(Z_i^k)$  (see (70)). Then

$$\chi_t^{M\mathcal{P}} = \sum_{n=0}^{N(t)} \sum_{k=1}^{M(\gamma_{n+1})} \int_{\Gamma_n}^{\Gamma_{n+1} \wedge t} \int_{I_k} |\Psi_k(z)|^2 \underline{c}(z) N(dz, dr) = \int_0^t \int_{B_{M\mathcal{P}}(r)} \Psi(z) \underline{c}(z) N(dz, dr),$$

with

$$\Psi(z) = \sum_{k=1}^\infty |\Psi_k(z)|^2 \mathbb{1}_{I_k}(z) \geq \sum_{k=1}^\infty \mathbb{1}_{[k-\frac{3}{4}, k-\frac{1}{4}]}(|z|) \mathbb{1}_{I_k}(z).$$

Using Itô formula,

$$\begin{aligned} \mathbb{E}(e^{-s\chi_t^{M\mathcal{P}}}) &= 1 + \mathbb{E} \int_0^t \int_{B_{M\mathcal{P}}(r)} (e^{-s(\chi_{r-}^{M\mathcal{P}} + \Psi(z) \underline{c}(z))} - e^{-s\chi_{r-}^{M\mathcal{P}}}) \widehat{N}(dz, dr) \\ &= 1 - \sum_{n=0}^{N(t)} \left( \int_{\Gamma_n}^{\Gamma_{n+1} \wedge t} \mathbb{E}(e^{-s\chi_r^{M\mathcal{P}}}) dr \sum_{k=1}^{M(\gamma_{n+1})} \int_{I_k} (1 - e^{-s|\Psi_k(z)|^2 \underline{c}(z)}) \mu(dz) \right). \end{aligned}$$

Solving the above equation we obtain

$$\begin{aligned} \mathbb{E}(e^{-s\chi_t^{M\mathcal{P}}}) &= \exp \left( - \sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \sum_{k=1}^{M(\gamma_{n+1})} \int_{I_k} (1 - e^{-s|\Psi_k(z)|^2 \underline{c}(z)}) \mu(dz)) \right) \\ &\leq \exp \left( - \sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \sum_{k=1}^{M(\gamma_{n+1})} \int_{I_k} (1 - e^{-s \mathbb{1}_{[k-\frac{3}{4}, k-\frac{1}{4}]}(|z|) \underline{c}(z)}) \mu(dz)) \right) \\ &= \exp \left( - \sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \sum_{k=1}^{M(\gamma_{n+1})} \int_{I_k} (1 - e^{-s \underline{c}(z)}) \mathbb{1}_{[k-\frac{3}{4}, k-\frac{1}{4}]}(|z|) \mu(dz)) \right) \\ &= \exp \left( - \sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \int_{B_{M(\gamma_{n+1})}} (1 - e^{-s \underline{c}(z)}) \bar{\mu}(dz)) \right), \end{aligned}$$

with

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

On the other hand, we denote

$$\bar{\chi}_t^{M_P} = \int_0^t \int_{B_{M_P}^c(r)} \Psi(z) \underline{c}(z) N(dz, dr),$$

where  $B_m^c$  denote the complementary set of  $B_m$ . Then in the same way,

$$\mathbb{E}(e^{-s\bar{\chi}_t^{M_P}}) \leq \exp\left(-\sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \int_{B_{M(\gamma_{n+1})}^c} (1 - e^{-s\underline{c}(z)}) \bar{\mu}(dz))\right).$$

We recall by (63) that for  $\Gamma_n < t \leq \Gamma_{n+1}$ ,

$$a_t^P = \left( \sum_{1 \leq i \leq n} \gamma_i \int_{\{|z| \geq M(\gamma_i)\}} \underline{c}(z) \mu(dz) + (t - \Gamma_n) \int_{\{|z| \geq M(\gamma_{n+1})\}} \underline{c}(z) \mu(dz) \right)^{\frac{1}{2}}.$$

Then

$$a_t^P \geq \sqrt{\mathbb{E}\bar{\chi}_t^{M_P}}.$$

Using Jensen inequality for the convex function  $f(x) = e^{-sx}$ ,  $s, x > 0$ , we have

$$e^{-s|a_t^{M_P}|^2} \leq e^{-s\mathbb{E}\bar{\chi}_t^{M_P}} \leq \mathbb{E}(e^{-s\bar{\chi}_t^{M_P}}) \leq \exp\left(-\sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \int_{B_{M(\gamma_{n+1})}^c} (1 - e^{-s\underline{c}(z)}) \bar{\mu}(dz))\right).$$

So we deduce that

$$\begin{aligned} \mathbb{E}(e^{-s(\chi_t^{M_P} + |a_t^P|^2)}) &= \mathbb{E}(e^{-s\chi_t^{M_P}}) \times e^{-s|a_t^P|^2} \\ &\leq \exp\left(-\sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \int_{B_{M(\gamma_{n+1})}^c} (1 - e^{-s\underline{c}(z)}) \bar{\mu}(dz))\right) \\ &\times \exp\left(-\sum_{n=0}^{N(t)} (((\Gamma_{n+1} \wedge t) - \Gamma_n) \int_{B_{M(\gamma_{n+1})}^c} (1 - e^{-s\underline{c}(z)}) \bar{\mu}(dz))\right) \\ &= \exp\left(-t \int_{\mathbb{R}^d} (1 - e^{-s\underline{c}(z)}) \bar{\mu}(dz)\right), \end{aligned} \tag{83}$$

and the last term does not depend on  $M_P(t)$ .

**Step 4** Now we use the Lemma 14 from [9], which states the following.

**Lemma 5.4.** *We consider an abstract measurable space  $B$ , a  $\sigma$ -finite measure  $\mathcal{M}$  on this space and a non-negative measurable function  $f : B \rightarrow \mathbb{R}_+$  such that  $\int_B f d\mathcal{M} < \infty$ . For  $t > 0$  and  $p \geq 1$ , we note*

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

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

$$\lim_{u \rightarrow \infty} \frac{1}{\ln u} \mathcal{M}(f \geq \frac{1}{u}) > \frac{p}{t}, \tag{84}$$

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

We will use the above lemma for  $\mathcal{M}(dz) = \bar{\mu}(dz)$ ,  $f(z) = \underline{c}(z)$  and  $B = \mathbb{R}^d$ . Thanks to (54) in **Hypothesis 2.4**,

$$\lim_{u \rightarrow \infty} \frac{1}{\ln u} \bar{\mu}(\underline{c} \geq \frac{1}{u}) = \infty. \tag{85}$$

Then for every  $p \geq 1, 1 \leq t \leq 3$ , we deduce from (81),(82),(83) and **Lemma 5.4** that

$$\begin{aligned} \mathbb{E} \left| \frac{1}{\det \sigma_{X_t^{M_P}}} \right|^p &\leq \mathbb{E}(|\lambda_t^{M_P}|^{-dp}) \leq C(\mathbb{E}(|\chi_t^{M_P} + |a_t^P|^2|^{-2dp}))^{\frac{1}{2}} \\ &\leq C \left( \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \mathbb{E}(e^{-s(\chi_t^{M_P} + |a_t^P|^2)}) ds \right)^{\frac{1}{2}} \\ &\leq C \left( \frac{1}{\Gamma(2dp)} \int_0^\infty s^{2dp-1} \exp(-t \int_{\mathbb{R}^d} (1 - e^{-s\zeta(z)}) \bar{\mu}(dz) ds) \right)^{\frac{1}{2}} < \infty. \end{aligned} \quad (86)$$

**Proof of ii)** We consider an increasing sequence of partition  $\mathcal{P}_n, n \in \mathbb{N}, (\mathcal{P}_n \subset \mathcal{P}_{n+1})$ , such that  $|\mathcal{P}_n| \downarrow 0$ . In particular,  $\forall t, M_{\mathcal{P}_n}(t) \uparrow \infty$ .

We recall the notations  $F_t^{M_P}(x) = X_t^{M_P}(x) - x$  and  $F_t(x) = X_t(x) - x$ . We notice by **Lemma 5.1 ii)** that  $\mathbb{E}|F_t^{M_{\mathcal{P}_n}} - F_t| \rightarrow 0$  as  $n \rightarrow \infty$ , and by **Lemma 5.2** that  $\sup_n \sup_x \|F_t^{M_{\mathcal{P}_n}}(x)\|_{L,l,p} \leq C_{l,p}$ .

Moreover, by **Lemma 5.5 ii)** (given immediately below), we know that  $(DF_t^{M_{\mathcal{P}_n}})_{n \in \mathbb{N}}$  is a Cauchy sequence in  $L^2(\Omega; l^2 \times \mathbb{R}^d)$ . Then applying **Lemma 3.3 (B)** with  $F_n = F_t^{M_{\mathcal{P}_n}}$  and  $F = F_t$ , **Lemma 5.3 i)** implies **Lemma 5.3 ii)**.  $\square$

### 5.3 Auxiliary results

Besides the lemmas concerning the Sobolev norms and covariance matrices, we establish an auxiliary result. We recall  $\varepsilon_m$  given in (58).

**Lemma 5.5.** *We assume that **Hypothesis 2.1** and **Hypothesis 2.4 b)** hold true.*

i) *Then for any  $\varepsilon_0 > 0$ , there exists a constant  $C$  dependent on  $d, \varepsilon_0$  such that for every  $t \in [0, 3]$  and every stating point  $x \in \mathbb{R}^d$ , we have*

$$\mathbb{E}|\det \sigma_{X_t^{\mathcal{P}, M_P}} - \det \sigma_{X_t}|^{\frac{2}{1+\varepsilon_0}} \leq C \|DX_t^{\mathcal{P}, M_P} - DX_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}^{\frac{2}{1+\varepsilon_0}} \leq C(|\mathcal{P}| + \varepsilon_M(|\mathcal{P}|))^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}.$$

ii) *We consider an increasing sequence of partition  $\mathcal{P}_n, n \in \mathbb{N}, (\mathcal{P}_n \subset \mathcal{P}_{n+1})$ , such that  $|\mathcal{P}_n| \downarrow 0$ . In particular,  $\forall t, M_{\mathcal{P}_n}(t) \uparrow \infty$ . We denote*

$$F_n(x) = X_t^{M_{\mathcal{P}_n}}(x).$$

*Then for each  $x \in \mathbb{R}^d$ , the sequence  $DF_n(x), n \in \mathbb{N}$  is Cauchy in  $L^2(\Omega; l_2 \times \mathbb{R}^d)$ , uniformly with respect to  $x$ :*

$$\sup_x \|DF_n(x) - DF_m(x)\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \rightarrow 0, \text{ as } n, m \rightarrow \infty.$$

*Proof. Proof of i)*

By **Lemma 5.2**, we know that  $\|DX_t^{\mathcal{P}, M_P}\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}$  and  $\|DX_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}$  are bounded, uniformly with respect to  $x$ . Then using Hölder's inequality with conjugates  $1 + \varepsilon_0$  and  $\frac{1+\varepsilon_0}{\varepsilon_0}$ , we get

$$\mathbb{E}(|\det \sigma_{X_t^{\mathcal{P}, M_P}} - \det \sigma_{X_t}|^{\frac{2}{1+\varepsilon_0}}) \leq C \|DX_t^{\mathcal{P}, M_P} - DX_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}^{\frac{2}{1+\varepsilon_0}}. \quad (87)$$

Now we only need to prove that

$$\|DX_t^{\mathcal{P}, M_P} - DX_t\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}^{\frac{2}{1+\varepsilon_0}} \leq C(|\mathcal{P}| + \varepsilon_M(|\mathcal{P}|))^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}. \quad (88)$$

The proof of (88) is a slight variant of the proof of **Lemma 3.9 iii)** in the paper [44]. The difference is that the truncation function  $M$  is constant in [44] while here it may vary on different time intervals. We do not discuss in detail here. So we conclude that **Lemma 5.5 i)** holds.

**Proof of ii)**

We consider an increasing sequence of partition  $\mathcal{P}_n$ ,  $n \in \mathbb{N}$ , ( $\mathcal{P}_n \subset \mathcal{P}_{n+1}$ ), such that  $|\mathcal{P}_n| \downarrow 0$ . In particular,  $\forall t$ ,  $M_{\mathcal{P}_n}(t) \uparrow \infty$ . We need to prove that

$$\|DX_t^{M_{\mathcal{P}_n}} - DX_t^{M_{\mathcal{P}_m}}\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)} \rightarrow 0, \text{ as } n, m \rightarrow \infty. \quad (89)$$

The proof of (89) is a slight variant of the proof of (148) p.47-49 in [44], so we omit it.

□

## 6 Proof of Theorem 4.1

In this section, we give the proof of **Theorem 4.1**. We apply **Proposition 2.1.1** in Section 2. For a measurable function  $f$ , we denote  $\bar{P}_t f(x) = \mathbb{E}f(X_t^{\mathcal{P}, M_{\mathcal{P}}}(x))$  and  $P_t f(x) = \mathbb{E}f(X_t(x))$ . In the following subsections, we will check the conditions of **Proposition 2.1.1**.

### 6.1 Euler: condition (7)

For every  $\gamma > 0$ , we recall in (59) that we define  $M(\gamma) \in \mathbb{N}$  such that

$$\varepsilon_{M(\gamma)} \leq \gamma^2.$$

We recall the basic equation  $X_t$  (see (49)). We denote by  $\tilde{X}_t^{M_{\mathcal{P}}}$  the one step truncated Euler scheme:

$$\tilde{X}_t^{M_{\mathcal{P}}}(x) = x + \int_0^t \int_{B_{M(\gamma)}} c(z, x) dN(z, s) + \int_0^t b(x) ds.$$

Then,

$$\begin{aligned} \mathbb{E}|\tilde{X}_\gamma^{M_{\mathcal{P}}} - X_\gamma| &\leq \mathbb{E} \int_0^\gamma \int_{\{|z| \geq M(\gamma)\}} |c(z, X_s)| dN(z, s) + \mathbb{E} \int_0^\gamma \int_{B_{M(\gamma)}} |c(z, x) - c(z, X_s)| dN(z, s) \\ &\quad + \mathbb{E} \int_0^\gamma |b(x) - b(X_s)| ds \\ &\leq \gamma \int_{\{|z| \geq M(\gamma)\}} \bar{c}(z) \mu(dz) + C \int_0^\gamma \mathbb{E}|x - X_s| ds \\ &\leq \gamma \sqrt{\varepsilon_{M(\gamma)}} + C \times \gamma^2 \leq C \times \gamma^2 \end{aligned}$$

So

$$W_1(\tilde{X}_\gamma^{M_{\mathcal{P}}}, X_\gamma) \leq \mathbb{E}|\tilde{X}_\gamma^{M_{\mathcal{P}}} - X_\gamma| \leq C \times \gamma^2.$$

So we conclude that (7) holds for  $\alpha = 1$  and  $k_0 = 0$ .

### 6.2 Lipschitz: condition (4) and the existence of an invariant measure

We recall that  $X$  is the solution of the equation (49).

**Lemma 6.1.** *Suppose that **Hypothesis 2.5** (see (55) and (56)) holds.*

a) *Then, for a Lipschitz continuous function  $f$*

$$|\mathbb{E}(f(X_t(x))) - \mathbb{E}(f(X_t(y)))| \leq L_f e^{-\frac{\theta}{2}t} |x - y|,$$

*with  $L_f$  the Lipschitz constant of  $f$ .*

b) *Moreover, there exists at least one invariant probability.*

**Proof a)** We fix  $x, y \in \mathbb{R}^d$  and we construct on the same probability space, with the same Poisson point measure  $N$  the solution  $X_t^M(y)$  which starts from  $y$ . Then we denote

$$\begin{aligned} Y_t &= X_t(x) - X_t(y), \\ \Delta_t^c(z) &= c(z, X_{s-}(x)) - c(z, X_{s-}(y)) \\ \Delta_t^b &= b(X_{s-}(x)) - b(X_{s-}(y)) \end{aligned}$$

and we have

$$Y_t = x - y + \int_0^t \int_{\mathbb{R}^d} \Delta_s^c(z) dN(z, s) + \int_0^t \Delta_s^b ds.$$

Using Itô's formula for  $\Phi(t, u) = e^{\lambda t} |u|^2$  we obtain

$$\begin{aligned} \Phi(t, Y_t) &= |x - y|^2 + \lambda \int_0^t \Phi(s, Y_s) ds + \int_0^t 2e^{\lambda s} \langle Y_s, \Delta_s^b \rangle ds \\ &\quad + \int_0^t \int_{\mathbb{R}^d} (\Phi(s, Y_{s-} + \Delta_s^c(z)) - \Phi(s, Y_{s-})) dN(z, s) \\ &= |x - y|^2 + \lambda \int_0^t \Phi(s, Y_s) ds + \int_0^t 2e^{\lambda s} \langle Y_s, \Delta_s^b \rangle ds \\ &\quad + M_t + \int_0^t \int_{\mathbb{R}^d} (\Phi(s, Y_{s-} + \Delta_s^c(z)) - \Phi(s, Y_{s-})) d\mu(z) ds \end{aligned}$$

with  $M_t$  a martingale. Taking the expectation we get

$$e^{\lambda t} \mathbb{E} |Y_t|^2 \leq |x - y|^2 + \int_0^t e^{\lambda s} \mathbb{E}(\Psi_s) ds$$

with

$$\begin{aligned} \Psi_s &= \lambda |Y_s|^2 + 2 \langle Y_s, \Delta_s^b \rangle + \int_{\mathbb{R}^d} |Y_s + \Delta_s^c(z)|^2 - |Y_s|^2 \mu(dz) \\ &= \lambda |Y_s|^2 + 2 \langle Y_s, \Delta_s^b \rangle + \int_{\mathbb{R}^d} \langle \Delta_s^c(z), 2Y_s + \Delta_s^c(z) \rangle \mu(dz). \end{aligned}$$

We need to prove that  $\mathbb{E}(\Psi_s) \leq 0$ . We recall that we assume **Hypothesis 2.5 i)ii)** (see (55)). We also have

$$|\langle \Delta_s^c(z), 2Y_s + \Delta_s^c(z) \rangle| \leq (2\bar{c}(z) + \bar{c}^2(z)) |Y_s|^2,$$

so that

$$\Psi_s \leq |Y_s|^2 (\lambda + \int_{\mathbb{R}^d} (2\bar{c}(z) + \bar{c}^2(z)) \mu(dz) - 2\bar{b}).$$

Thanks to **Hypothesis 2.5 iii)** (see (56)), taking  $\lambda \leq \theta$ , we have

$$e^{\lambda t} \mathbb{E} |Y_t|^2 \leq |x - y|^2 + \int_0^t e^{\lambda s} \mathbb{E}(\Psi_s) ds \leq |x - y|^2$$

so that

$$\mathbb{E} |X_t(x) - X_t(y)|^2 \leq e^{-\theta t} |x - y|^2.$$

Then, for a Lipschitz continuous function  $f$ ,

$$|\mathbb{E}(f(X_t(x))) - \mathbb{E}(f(X_t(y)))| \leq L_f \mathbb{E} |X_t(x) - X_t(y)| \leq L_f e^{-\frac{\theta}{2}t} |x - y|.$$

b) We denote  $L$  to be the infinitesimal operator of (49). We take  $V(x) = |x|^2$  and we will prove that

$$LV \leq \bar{\beta} - \bar{\alpha}V$$

for some  $\bar{\beta}, \bar{\alpha} > 0$  (the Lyapunov mean reverting condition). This implies  $LV \leq C$  and  $\lim_{|x| \rightarrow \infty} LV(x) = -\infty$ . Then we use Theorem 9.3 and Lemma 9.7 from [18] (with  $\varphi = V$  and  $\psi = LV$ ) which guarantees existence of an invariant distribution. We have

$$\begin{aligned} LV(x) &= 2 \langle x, b(x) \rangle + \int_{\mathbb{R}^d} (V(x + c(z, x)) - V(x)) \mu(dz) \\ &\leq -2\bar{b} |x|^2 + \int_{\mathbb{R}^d} (2 \langle x, c(z, x) \rangle + |c(z, x)|^2) \mu(dz) \\ &\leq -2\bar{b} |x|^2 + (|x|^2 + 1) \int_{\mathbb{R}^d} \bar{c}(z) \mu(dz) + \int_{\mathbb{R}^d} \bar{c}^2(z) \mu(dz) \\ &= \int_{\mathbb{R}^d} (\bar{c}(z) + \bar{c}^2(z)) \mu(dz) - (2\bar{b} - \int_{\mathbb{R}^d} \bar{c}(z) \mu(dz)) |x|^2. \end{aligned}$$

□

### 6.3 Regularity: conditions (14), (15), (16) and (20)

Firstly, we deal with (14). Let  $t \in [1, 2]$ . For any  $k$  and any multi-index  $\beta_0$  with  $|\beta_0| = k$ , we write

$$\partial_x^{\beta_0} P_t \varphi(x) = \mathbb{E}[\partial_x^{\beta_0} (\varphi(X_t(x)))] = \sum_{|\alpha_0| \leq |\beta_0|} \mathbb{E}[(\partial^{\alpha_0} \varphi)(X_t(x)) \mathbf{P}_{\alpha_0}(x)],$$

with  $\mathbf{P}_{\alpha_0}(x)$  a polynomial of  $\partial_x^{\alpha_1} X_t(x)$ ,  $|\alpha_1| \leq |\beta_0|$ .

In the following, we use the results from Section 5. In **Lemma 5.2**, we prove that the Sobolev norms of each  $\partial_x^{\alpha_1} X_t(x)$  are bounded, uniformly with respect to  $x$ . It follows that this is also true for  $\mathbf{P}_{\alpha_0}(x)$ .

We denote that  $F_t(x) = X_t(x) - x$ . In **Lemma 5.2**, we have proved that the Sobolev norms of each  $F_t(x)$  are bounded, uniformly with respect to  $x$ . Moreover, in **Lemma 5.3**, we prove that  $F_t(x)$  is non-degenerated, uniformly with respect to  $x$ , that is  $\Sigma_p(F_t(x)) < \infty$  for each  $p$  (see (26)).

Then we use **Lemma 3.4 (A)** which asserts that  $(B_k)$  is true for  $F = F_t(x)$  and  $G = \mathbf{P}_{\alpha_0}(x)$ . By the remark of **Lemma 3.4**,  $(B_k)$  is also true for  $F = X_t(x) = F_t(x) + x$  and  $G = \mathbf{P}_{\alpha_0}(x)$ . This reads

$$|\mathbb{E}[(\partial^{\alpha_0} \varphi)(X_t(x)) \mathbf{P}_{\alpha_0}(x)]| \leq C \|\varphi\|_{\infty},$$

which gives (14).

In a similar way, we can obtain (15).

For (16), *i* is a direct consequence of (4) which has been proved in Section 6.2. For (16) *ii*), we take  $t \in (0, 1]$ . For any  $k$  and any multi-index  $\beta_0$  with  $|\beta_0| = k$ , we notice that

$$|\partial_x^{\beta_0} \nabla P_t \varphi(x)| = |\mathbb{E}[\partial_x^{\beta_0} (\nabla \varphi(X_t(x)))]| = \left| \sum_{|\alpha_0| \leq |\beta_0|} \mathbb{E}[(\partial^{\alpha_0} \nabla \varphi)(X_t(x)) \mathbf{P}_{\alpha_0}(x)] \right| \leq \|\nabla \varphi\|_{k, \infty} \sum_{|\alpha_0| \leq |\beta_0|} \mathbb{E}|\mathbf{P}_{\alpha_0}(x)|,$$

with  $\mathbf{P}_{\alpha_0}(x)$  a polynomial of  $\partial_x^{\alpha_1} X_t(x)$ ,  $|\alpha_1| \leq |\beta_0|$ . In [29], Kunita has shown in Theorem 3.4.1 and Theorem 3.4.2 the regularity of the flow associated with the jump-diffusion. So in our case, we have  $\mathbb{E}|\mathbf{P}_{\alpha_0}(x)| < \infty$  and thus (16) *ii*) holds true.

Now we prove (20). In order to prove (20), we need to represent  $\bar{P}_{s,t} \varphi(x)$  and  $P_{s,t} \varphi(x)$ . So we consider the following equations.

We denote  $X_{s,t}^{\mathcal{P}, M_{\mathcal{P}}}$  and  $X_{s,t}$  the solutions of the following equations respectively:

$$X_{s,t}^{\mathcal{P}, M_{\mathcal{P}}} = x + \int_s^t b(X_{s,\tau}^{\mathcal{P}, M_{\mathcal{P}}}) d\tau + \int_s^t \int_{B_{M_{\mathcal{P}}}(\tau)} c(z, X_{s,\tau}^{\mathcal{P}, M_{\mathcal{P}}}) N(dz, dr); \quad (90)$$

$$X_{s,t} = x + \int_s^t b(X_{s,r}) dr + \int_s^t \int_{\mathbb{R}^d} c(z, X_{s,r}) N(dz, dr). \quad (91)$$



We sometimes write  $X_{s,t}^{\mathcal{P},M_{\mathcal{P}}}(x)$  (and  $X_{s,t}(x)$ ) instead of  $X_{s,t}^{\mathcal{P},M_{\mathcal{P}}}$  (and  $X_{s,t}$ ) to stress the dependence on the initial value  $x$ . And we denote  $\bar{P}_{s,t}\varphi(x) = \mathbb{E}\varphi(X_{s,t}^{\mathcal{P},M_{\mathcal{P}}}(x))$  and  $P_{s,t}\varphi(x) = \mathbb{E}\varphi(X_{s,t}(x))$ .

Let  $1 < t < r < t+2$ . We recall that  $\mathcal{P} = \{0 = \Gamma_0 < \Gamma_1 < \dots < \Gamma_{l-1} < \Gamma_l < \dots\}$ ,  $\gamma_l = \Gamma_l - \Gamma_{l-1}$  and for  $\Gamma_l \leq t < \Gamma_{l+1}$ ,  $N(t) = l$ . We denote

$$F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x) = X_{t,r} \circ X_{t-1,t}^{\mathcal{P},M_{\mathcal{P}}}(x) - x \text{ and } F_{r-t+1}(x) = X_{t,r} \circ X_{t-1,t}(x) - x = X_{t-1,r}(x) - x. \quad (92)$$

We also denote  $|\mathcal{P}^{t-1,t}| := \max_{\substack{l \in \mathbb{N} \\ \Gamma_{l+1} > t-1, \Gamma_l < t}} ((\Gamma_{l+1} \wedge t) - (\Gamma_l \vee (t-1)))$ . Before we give the proof of (20), we state the following lemma concerning  $F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)$  and  $F_{r-t+1}(x)$  given in (92).

**Lemma 6.2.** *Under the Hypothesis 2.1~2.4, we have these results.*

i) For all  $p \geq 1, q \geq 0$ , there exists a constant  $C_{q,p}$  depending on  $q, p, d$ , such that  $F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)$  and  $F_{r-t+1}(x)$  belong to  $\mathcal{D}_{\infty}^d$  and

$$\sup_x \sup_{\mathcal{P}} \|F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x) + F_{r-t+1}(x)\|_{L,q,p} \leq C_{q,p}.$$

ii) For every  $p \geq 1$ , we have

$$\sup_{\mathcal{P}} \sup_x \mathbb{E}(1/\det \sigma_{F_{r-t+1}(x)})^p \leq C_p,$$

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

iii) For any  $\varepsilon_0 > 0$ , there exists a constant  $C$  dependent on  $d, \varepsilon_0$  such that

$$\sup_x \mathbb{E} |\det \sigma_{F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)} - \det \sigma_{F_{r-t+1}(x)}|^{\frac{2}{1+\varepsilon_0}} \leq C |\mathcal{P}^{t-1,t}|^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}.$$

*Proof.* Firstly, we will construct an approximation scheme for  $X_{t,r} \circ X_{t-1,t}^{\mathcal{P},M_{\mathcal{P}}}(x)$ . We take an integer  $N_0$  such that  $\frac{1}{2^{N_0}} \leq |\mathcal{P}|$ . For  $n > N_0$ , we take a "mixed partition"

$$\begin{aligned} \mathcal{P}_n &= \{t-1 < \Gamma_{N(t-1)+1} < \dots < \Gamma_{N(t)} \leq t \\ &< t + \frac{1}{2^n}(r-t) < t + \frac{2}{2^n}(r-t) < \dots < t + \frac{l}{2^n}(r-t) < t + \frac{l+1}{2^n}(r-t) < \dots < r\} \\ &:= \{t-1 = s_0 < s_1 < \dots < s_{n_0} = r\}. \end{aligned}$$

We remark that we take the partition  $\{\Gamma_l\}$  on  $[t-1, t]$  and take the partition  $\{\frac{l}{2^n}\}$  on  $[t, r]$ . We denote  $|\mathcal{P}_n| := \max_{k \in \{1, \dots, n_0\}} s_k - s_{k-1}$ . We construct  $M_{\mathcal{P}_n}(t) = M(s_{l+1} - s_l)$  when  $s_l < t \leq s_{l+1}$  with the truncation function  $M(\bullet)$  given in (59). And we denote  $\tau^{\mathcal{P}_n}(t) = s_l$  when  $s_l < t \leq s_{l+1}$ . Then we consider the truncated Euler scheme based on  $\mathcal{P}_n, M_{\mathcal{P}_n}$ :

$$X_{t-1,r}^{\mathcal{P}_n,M_{\mathcal{P}_n}} = x + \int_{t-1}^r b(X_{t-1,\tau^{\mathcal{P}_n}(s)}^{\mathcal{P}_n,M_{\mathcal{P}_n}})ds + \int_{t-1}^r \int_{B_{M_{\mathcal{P}_n}(s)}} c(z, X_{t-1,\tau^{\mathcal{P}_n}(s)}^{\mathcal{P}_n,M_{\mathcal{P}_n}})N(dz, ds).$$

We denote

$$F_{r-t+1}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x) = X_{t-1,r}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x) - x. \quad (93)$$

We notice that we can apply the results from Section 5 for  $F_{r-t+1}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x)$ ,  $F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)$  and  $F_{r-t+1}(x)$  defined in (92) and (93).

Since  $r-t+1 < 3$ , by **Lemma 5.2 i)**, the Sobolev norms of  $F_{r-t+1}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x)$  and  $F_{r-t+1}(x)$  are bounded, uniformly with respect to  $x$ . One can check that  $F_{r-t+1}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x) \rightarrow F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)$  in  $L^1(\Omega)$ , as  $n \rightarrow \infty$  (which is a variant of **Lemma 5.1 i)**). So we can apply **Lemma 3.3 (A)** for  $F_n = F_{r-t+1}^{\mathcal{P}_n,M_{\mathcal{P}_n}}(x)$  and  $F = F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)$  in order to get that  $F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x) \in \mathcal{D}_{\infty}^d$  and  $\sup_x \sup_{\mathcal{P}} \|F_{r-t+1}^{\mathcal{P},M_{\mathcal{P}}}(x)\|_{L,q,p} \leq C_{q,p}$ . Hence, **Lemma 6.2 i)** is proved.

Moreover, since  $r - t + 1 > 1$ , using **Lemma 5.3 ii)** we have  $\sup_{\mathcal{P}} \sup_x \mathbb{E}(1/\det \sigma_{F_{r-t+1}^{\mathcal{M}_{\mathcal{P}}}(x)})^p \leq C_p$ . So

**Lemma 6.2 ii)** is proved.

Finally, by **Lemma 5.5 i)** and recalling by (59) that  $\varepsilon_{M(\gamma)} \leq \gamma^2$ , we have

$$\|DF_{r-t+1}^{\mathcal{P}_n, \mathcal{M}_{\mathcal{P}_n}}(x) - DF_{r-t+1}(x)\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}^{\frac{2}{1+\varepsilon_0}} \leq C(|\mathcal{P}_n| + \varepsilon_M(|\mathcal{P}_n|))^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}} \leq C|\mathcal{P}^{t-1, t}|^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}},$$

where the last equality is true since  $\frac{1}{2^n} \leq |\mathcal{P}^{t-1, t}|$  for every  $n > N_0$ . Then we can apply **Lemma 3.3 (C)** for  $(F_n, \bar{F}_n) = (F_{r-t+1}^{\mathcal{P}_n, \mathcal{M}_{\mathcal{P}_n}}(x), F_{r-t+1}(x))$  and  $(F, \bar{F}) = (F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x), F_{r-t+1}(x))$ . So  $\sup_x \mathbb{E}|\det \sigma_{F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)} - \det \sigma_{F_{r-t+1}(x)}|^{\frac{2}{1+\varepsilon_0}} \leq C\|DF_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x) - DF_{r-t+1}(x)\|_{L^2(\Omega; l_2 \times \mathbb{R}^d)}^{\frac{2}{1+\varepsilon_0}} \leq C|\mathcal{P}^{t-1, t}|^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}$  and **Lemma 6.2 iii)** is proved.  $\square$

Then we can prove (20). By **Lemma 6.2 i)**, the Sobolev norms of  $F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)$  are bounded, uniformly with respect to  $x$ . Using **Lemma 6.2 ii)**, the covariance matrix of  $F_{r-t+1}(x)$  is non-degenerated. Then we are able to apply **Lemma 3.5** for  $F = F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)$  and  $Q = F_{r-t+1}(x)$  so (38) holds for  $F = F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)$  and  $Q = F_{r-t+1}(x)$ . Thanks to the remark of **Lemma 3.5**, (38) also holds for  $F = X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x) = F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x) + x$ ,  $Q = X_{t,r} \circ X_{t-1,t}(x) = F_{r-t+1}(x) + x$  and get

$$\begin{aligned} & \left| \mathbb{E}(f(X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) - \mathbb{E}(f_\delta(X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) \right| \\ & \leq C \|f\|_\infty \times \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-p} \mathbb{E}|\det \sigma_{X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)} - \det \sigma_{X_{t,r} \circ X_{t-1,t}(x)}|^p + \eta^\kappa \right), \end{aligned} \quad (94)$$

where we have also used the fact that  $\sup_{\mathcal{P}} \sup_x \mathbb{E}(1/\det \sigma_{X_{t,r} \circ X_{t-1,t}(x)})^\kappa \leq C_\kappa$  from **Lemma 6.2 ii)**.

We take  $p = \frac{2}{1+\varepsilon_0}$  for any small  $\varepsilon_0$ . Thanks to **Lemma 6.2 iii)**,

$$\sup_x \mathbb{E}|\det \sigma_{X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)} - \det \sigma_{X_{t,r} \circ X_{t-1,t}(x)}|^{\frac{2}{1+\varepsilon_0}} = \sup_x \mathbb{E}|\det \sigma_{F_{r-t+1}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)} - \det \sigma_{F_{r-t+1}(x)}|^{\frac{2}{1+\varepsilon_0}} \leq C|\mathcal{P}^{t-1, t}|^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}.$$

This implies that

$$\sup_x \mathbb{E}|\det \sigma_{X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x)} - \det \sigma_{X_{t,r} \circ X_{t-1,t}(x)}|^{\frac{2}{1+\varepsilon_0}} \leq C\gamma_{N(t-1)}^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}}.$$

Substituting into (94), we obtain

$$\sup_x \left| \mathbb{E}(f(X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) - \mathbb{E}(f_\delta(X_{t,r} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) \right| \leq C \|f\|_\infty \times \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-\frac{2}{1+\varepsilon_0}} \gamma_{N(t-1)}^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}} + \eta^\kappa \right).$$

By a similar argument, we have

$$\sup_x \left| \mathbb{E}(f(X_{t,r}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) - \mathbb{E}(f_\delta(X_{t,r}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}} \circ X_{t-1,t}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}(x))) \right| \leq C \|f\|_\infty \times \left( \frac{\delta^q}{\eta^{2q}} + \eta^{-\frac{2}{1+\varepsilon_0}} \gamma_{N(t-1)}^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)}} + \eta^\kappa \right).$$

So (20) holds for  $p = \frac{2}{1+\varepsilon_0}$  and  $\beta = \frac{1}{2+\varepsilon_0}$ .

Finally, we can apply **Proposition 2.1.1** for  $X_{0, \Gamma_n}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}$  and  $X_{0, \Gamma_n}$  with  $\alpha = 1$ ,  $k_0 = 0$ ,  $p = \frac{2}{1+\varepsilon_0}$ ,  $\beta = \frac{1}{2+\varepsilon_0}$  (for any small  $\varepsilon_0$ ), and obtain the following result: for every  $\varepsilon > 0$ , there exists a constant  $C$  such that

$$d_{TV}(X_{0, \Gamma_n}^{\mathcal{P}, \mathcal{M}_{\mathcal{P}}}, X_{0, \Gamma_n}) \leq C\gamma_n^{\frac{2}{(2+\varepsilon_0)(1+\varepsilon_0)} - \varepsilon} = C\gamma_n^{1-\varepsilon}, \quad (95)$$

with  $\bar{\varepsilon} = 1 - \frac{2-\varepsilon(2+\varepsilon_0)(1+\varepsilon_0)}{(2+\varepsilon_0)(1+\varepsilon_0)}$ .  
And moreover, we have

$$d_{TV}(\mathcal{L}(X_{\Gamma_n}^{\mathcal{P}, M^{\mathcal{P}}}), \nu) \leq C(\gamma_n^{1-\varepsilon} + \int_{\mathbb{R}^d} |x-y| d\nu(y) e^{-\frac{\rho}{2}\Gamma_n}),$$

where  $\nu$  is the unique invariant probability measure.

## 7 Appendix

### 7.1 The numerical lemma

In Section 2, we need to use the following numerical lemma.

**Lemma 7.1.** (A) Take an integer  $n_*$ . Let  $(\gamma_n)_{n \in \mathbb{N}}$  be a non-increasing positive sequence such that for  $n \geq n_*$ , we have

$$\frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2} \leq 2\bar{\omega}. \quad (96)$$

We denote  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . Then for every  $n_* \leq i \leq n$ , we have

$$\gamma_i \leq e^{2\bar{\omega}(\Gamma_n - \Gamma_i)} \times \gamma_n. \quad (97)$$

(B) We assume that  $(\gamma_n)_{n \in \mathbb{N}}$  is a non-increasing positive sequence verifying

$$\frac{\gamma_n - \gamma_{n+1}}{\gamma_{n+1}^2} \leq c_* < \frac{\rho}{\alpha}. \quad (98)$$

We denote  $\Gamma_n = \sum_{i=1}^n \gamma_i$ . Then

$$u_n := \sum_{i=1}^n \gamma_i^{1+\alpha} e^{-\rho(\Gamma_n - \Gamma_i)} \leq C\gamma_n^\alpha. \quad (99)$$

**Proof of (A)** Notice that (96) implies

$$\frac{\gamma_n}{\gamma_{n+1}} \leq 1 + 2\bar{\omega}\gamma_{n+1} \leq e^{2\bar{\omega}\gamma_{n+1}}.$$

Then

$$\frac{\gamma_i}{\gamma_n} = \prod_{k=i}^{n-1} \frac{\gamma_k}{\gamma_{k+1}} \leq \prod_{k=i}^{n-1} e^{2\bar{\omega}(\gamma_{k+1})} \leq e^{2\bar{\omega}(\Gamma_n - \Gamma_i)}.$$

**Proof of (B)** Notice that (98) implies

$$\frac{\gamma_n}{\gamma_{n+1}} \leq 1 + c_*\gamma_{n+1} \leq e^{c_*\gamma_{n+1}}.$$

Then we define  $v_n = u_n/\gamma_n^\alpha$  and we have the recurrence relation

$$v_{n+1} = \theta_n v_n + \gamma_{n+1}, \quad \theta_n = \frac{\gamma_n^\alpha}{\gamma_{n+1}^\alpha} \times e^{-\rho\gamma_{n+1}}.$$

Using the previous inequality we obtain

$$v_{n+1} \leq e^{(\alpha c_* - \rho)\gamma_{n+1}} v_n + \gamma_{n+1}$$

and further

$$\begin{aligned} e^{(\rho-\alpha c_*)\Gamma_{n+1}} v_{n+1} &\leq e^{(\rho-\alpha c_*)\Gamma_n} v_n + e^{(\rho-\alpha c_*)\Gamma_{n+1}} \gamma_{n+1} \\ &\leq e^{(\rho-\alpha c_*)\Gamma_n} v_n + C' e^{(\rho-\alpha c_*)\Gamma_n} \gamma_{n+1}, \end{aligned}$$

with  $C' = \sup_{k \geq 1} e^{(\rho-\alpha c_*)\gamma_k} = e^{(\rho-\alpha c_*)\gamma_1}$ . We use recursively this inequality and we obtain

$$\begin{aligned} e^{(\rho-\alpha c_*)\Gamma_{n+1}} v_{n+1} &\leq e^{(\rho-\alpha c_*)\Gamma_1} v_1 + C' \sum_{i=1}^n e^{(\rho-\alpha c_*)\Gamma_i} \gamma_{i+1} \\ &\leq e^{(\rho-\alpha c_*)\Gamma_1} v_1 + C' \int_0^{\Gamma_n} e^{(\rho-\alpha c_*)s} ds \\ &\leq e^{(\rho-\alpha c_*)\Gamma_1} v_1 + \frac{C'}{\rho - \alpha c_*} e^{(\rho-\alpha c_*)\Gamma_{n+1}}. \end{aligned}$$

That is

$$v_{n+1} \leq v_1 + \frac{C'}{\rho - \alpha c_*} \leq \gamma_1 + \frac{C'}{\rho - \alpha c_*}$$

which finally gives

$$u_{n+1} \leq (\gamma_1 + \frac{C'}{\rho - \alpha c_*}) \gamma_{n+1}^\alpha.$$

□

## References

- [1] A. Alfonsi, J. Corbetta, B. Jourdain: Evolution of the Wasserstein distance between the marginals of two Markov processes. *Bernoulli*, **24** (4A), pp.2461-2498. (hal-01390887) (2018).
- [2] 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).
- [3] D. Applebaum: *Lévy Processes and Stochastic Calculus* (2nd ed., Cambridge Studies in Advanced Mathematics). Cambridge: Cambridge University Press. (2009) doi:10.1017/CBO9780511809781.
- [4] D. Bakry, I. Gentil, M. Ledoux: Analysis and geometry of Markov diffusion operators. *Grundlehren der Mathematischen Wissenschaften* **348**, Springer, MR-3155209 (2014).
- [5] V. Bally and L. Caramellino: On the distances between probability density functions. *Electron. J. Probab.* **19**, no. 110, 1-33. MR-3296526 (2014).
- [6] V. Bally and L. Caramellino: Asymptotic development for the CLT in total variation distance. *Bernoulli* **22**, 2442–2485. MR-3498034 (2016).
- [7] V. Bally, L. Caramellino, G. Poly: Regularization lemmas and convergence in total variation. *Electron. J. Probab.* **25** 1-20. (2020).
- [8] V. Bally, L. Caramellino, G. Poly: Non universality for the variance of the number of real roots of random trigonometric polynomials. *Probab. Th. Rel. Fields.* **174**, 887-927. MR-3980307 (2019).
- [9] V. Bally, E. Clément: Integration by parts formula and applications to equations with jumps. *Probab. Th. Rel. Fields*, **151**, 613-657 (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] K. Bichteler, J. B. Gravereaux, J. Jacod: *Malliavin calculus for processes with jumps*. Gordon and Breach, (1987).
- [13] N. Bouleau, L. Denis: *Dirichlet forms methods for Poisson point measures and Lévy processes. With emphasis on the creation-annihilation techniques*. Probability Theory and Stochastic Modelling **76**, Springer. MR-3444890 (2015).
- [14] P. Chen, C. Deng, René L. Schilling, et. al.: Approximation of the invariant measure of stable SDEs by an Euler–Maruyama scheme. arXiv:2205.01342v2 [math.PR] (2023).
- [15] R. Cont, P. Tankov: *Finacial modelling with jump processes*. Chapman & Hall/CRC (2004).
- [16] A. Durmus, É. Moulines: High-dimensional Bayesian inference via the Unadjusted Langevin Algorithm. arXiv:1605.01559v4 (2018).
- [17] A. Durmus, É. Moulines: Nonasymptotic convergence analysis for the unadjusted Langevin algorithm. *The Annals of Applied Probability* Vol. **27**, No. 3, 1551–1587 (2017).
- [18] Stewart N. Ethier, Thomas G. Kurtz: *Markov Processes: characterization and convergence*. (1984).
- [19] N. Ikeda, S. Watanabe: *Stochastic differential equations and diffusion processes*. 2nd ed. Amsterdam, Netherlands, North Holland, (1989).
- [20] Y. Ishikawa: *Stochastic Calculus of Variations for Jump Processes*, Berlin, Boston: De Gruyter. (2013).
- [21] 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).
- [22] 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).
- [23] 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).
- [24] A. Kohatsu-Higa, P. Protter: The Euler scheme for SDE'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).
- [25] A. Kohatsu-Higa and P. Tankov: Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Processes and their Applications*, Vol. **120**, 2258-2285 (2010).
- [26] A. M. Kulik: Malliavin calculus for Lévy processes with arbitrary Lévy measures. *Theor. Probability ad Math. Statist.* No.**72**, 75-92 (2006).
- [27] 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).
- [28] 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).
- [29] H. Kunita: *Stochastic flows and jump-diffusions*. Springer,(2019).

- [30] D. Lamberton and G. Pagès: Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, **8**(3):367–405 (2002).
- [31] D. Lamberton and G. Pagès: Recursive computation of the invariant distribution of a diffusion: the case of a weakly mean reverting drift. *Stoch. Dyn.*, **3**(4):435–451 (2003).
- [32] B. Lapeyre, É. Pardoux and R. Sentis: *Méthodes de Monte-Carlo pour les équations de transport et de diffusion*. Springer-Verlag, Berlin, (1998).
- [33] E. Löcherbach, V. Rabet: Ergodicity for multidimensional jump diffusions with position dependent jump rate. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques* <hal-01144260> (2015).
- [34] E. Mariucci, M. Reiß: Wasserstein and total variation distance between marginals of Lévy processes. *Electronic Journal of Statistics* **12**, 2482-2514 (2018).
- [35] Sean P. Meyn and R. L. Tweedie: Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Advances in Applied Probability* Vol. **25**, No. 3, pp. 518-548 (1993).
- [36] W. Mou, N. Flammarion, M. J. Wainwright and P. L. Bartlett: Improved bounds for discretization of Langevin diffusions: Near-optimal rates without convexity. arXiv:1907.11331, (2019).
- [37] D. Nualart: *The Malliavin calculus and related topics*. Springer-Verlag, (2006).
- [38] G. Pagès: Sur quelques algorithmes récurrents pour les probabilités numériques. *Probability and Statistics* Vol. **5**, 141-170 (2001).
- [39] G. Pagès: *Numerical Probability: An introduction with applications to finance*. Universitext. Springer, Cham (2018).
- [40] G. Pagès, F. Panloup: Unadjusted Langevin algorithm with multiplicative noise: Total variation and Wasserstein bounds. (hal-03082311v3) (2022).
- [41] F. Panloup: Computation of the invariant measure for a Lévy driven SDE: Rate of convergence. *Stochastic Processes and their Applications*, Elsevier, **118** (8), pp.1351-1384. (2008).
- [42] F. Panloup: Recursive computation of the invariant measure of a stochastic differential equation driven by a Lévy process. *Annals of Applied Probability*, Institute of Mathematical Statistics (IMS), **18** (2), pp.379-426. (2008).
- [43] P. Protter and D. Talay: The Euler scheme for Lévy driven stochastic differential equations. *Ann. Probab.* Vol **25**, No.1, pg 393-423 (1997).
- [44] Y. Qin: Approximation schemes for McKean-Vlasov and Boltzmann type equations (error analysis in total variation distance). (arXiv:2212.07411v2) (2023).
- [45] Gareth O. Roberts and Richard L. Tweedie: Exponential convergence of Langevin distributions and their discrete approximations *Bernoulli* **2**(4), 341-363. (1996).
- [46] K. Sato. Lévy processes and infinitely divisible distributions. *Cambridge University press*, Cambridge (1999).
- [47] 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).
- [48] Y. Song and X. Zhang: Regularity of density for SDEs driven by degenerate Lévy noises. arXiv:1401.4624 (2014).
- [49] C. Villani: *Optimal Transport* Springer-Verlag, (2009).
- [50] X. Zhang: Densities for SDEs driven by degenerate  $\alpha$ -stable processes. *Ann. Probab.* Vol **42**, No.5, 1885-1910 (2014).