Limite de grande échelle de systèmes de particules en interaction avec sauts simultanés en régime diffusif

Large scale limits for interacting particle systems with simultaneous jumps in diffusive regime

Thèse de doctorat de l'Université Paris-Saclay

Ecole Doctorale de Mathématique Hadamard (EDMH) n° 574
Spécialité de doctorat : Mathématiques appliquées
Unité de recherche : Laboratoire de mathématiques et modélisation d'Evry (UEVE), UMR 8071 CNRS-INRA
Référent : Université d'Evry-Val d'Essonne

Thèse présentée et soutenue à Evry, le 30 juin 2021, par

Xavier ERNY

Composition du jury :

Stéphane Menozzi Professeur, Université d'Évry	Président du jury
François Delarue Professeur, Université Nice Sophia-Antipolis	Rapporteur & Examinateur
Alexander Veretennikov Professeur, Kharkevich Institute	Rapporteur & Examinateur
Emmanuelle Clément Maître de conférences, Université Gustave Eiffel	Examinatrice
Nicolas Fournier Professeur, Sorbonne Université	Examinateur
Sylvie Méléard Professeure, École Polytechnique	Examinatrice

Membres non délibérants :

Dasha Loukianova Maître de conférences, Université d'Évry	Directrice
Eva Löcherbach Professeure, Université Paris 1 Panthéon-Sorbonne	Codirectrice
Carl Graham Chargé de recherche, École Polytechnique	Invité

Thèse de doctorat

NNT : 2021UPASM012

université evry val-d'essonne

université PARIS-SACLAY

LaMME Laboratoire de Mathématiques et Modélisation d'Évry

Fondation mathématique FMJH Jacques Hadamard

Contents

Chapitre 1

Introduction (en français)

1.1 Résumé

Cette thèse est consacrée à l'étude de systèmes de particules en interaction dans différents modèles, où les particules interagissent, entre autre, à travers leurs sauts. Nous nous intéressons principalement à la question de la limite en grande échelle.

Plus précisément, dans chaque modèle, nous considérons pour chaque entier N, un système de N particules, où la dynamique de chaque particule est donnée par un processus stochastique solution d'une équation différentielle stochastique dirigée par des mouvements browniens et des mesures de Poisson. Dans un même système, les N équations ont les mêmes coefficients, de telle sorte que les interactions entre les particules d'un même système sont symétriques. On dit d'un tel système qu'il est échangeable. La question principale de la thèse est la convergence de ces systèmes quand le nombre de particules N tend vers l'infini. Nous détaillons dans chaque modèle le sens de cette convergence : il s'agit à la fois de la convergence des processus et de la convergence de la mesure empirique des systèmes. Pour un système échangeable, ces deux convergences sont équivalentes.

Pour décrire les interactions des systèmes que nous étudions nous avons besoin de la notion de processus ponctuel. En effet, pour chaque système de particules, nous associons à chaque processus stochastique un processus ponctuel dont l'intensité des sauts dépend du processus stochastique. Les équations différentielles stochastiques sont dirigées par ces processus ponctuels. Lorsqu'un processus ponctuel crée un point, cela modifie les valeurs de tous les processus stochastiques du système en les faisant "sauter" simultanément. C'est en ce sens que les particules interagissent.

Dans les équations, il y a donc un terme de saut d'interaction. Dans un système à N particules, ce terme s'écrit comme une somme de N processus ponctuels. Pour forcer ce terme à converger, il faut le normaliser. La normalisation que nous étudions principalement dans cette thèse est la normalisation en $N^{-1/2}$, que nous appelons normalisation diffusive. La question de la convergence de système de particules en normalisation diffusive a été peu étudiée dans la littérature, comparée à la normalisation en N^{-1} (que nous appelons normalisation linéaire). En normalisation linéaire, il est habituel d'observer une propagation du chaos : les particules du système limite sont indépendantes. Dans les cadres classiques, les résultats de ce type reposent sur des calculs sur les processus ponctuels, et leurs preuves montrent des convergences L^1. En normalisation diffusive, nous démontrons des convergences en loi. Pour démontrer ces convergence en loi, nous pouvons regrouper les techniques utilisées en deux catégories : celles liées à la théorie des processus de Markov, et celles liées à

un nouveau type de problème martingale. Contrairement à la situation de la normalisation linéaire, les particules du système limite ne sont pas, en général, indépendantes. En effet, en normalisation diffusive, le terme d'interaction limite est stochastique. Ce terme crée un bruit commun dans le système limite. Nous observons que les particules limites sont indépendantes conditionnellement à ce bruit commun. Nous appelons cette propriété la propagation du chaos conditionnelle.

Dans les Chapitres 2 et 3, la dynamique de chaque système à N particules est caractérisée par un processus de Markov de dimension finie indépendante de N. Nous démontrons la convergence de cette suite de processus et obtenons une vitesse de convergence explicite pour leurs semi-groupes. Pour cela, nous utilisons la notion de générateur infinitésimal et une formule de type Trotter-Kato. En notant, P^N (resp. $\bar{P}$) le semi-groupe du processus du système à N particules (resp. processus limite) et A^N (resp. $\bar{A}$) son générateur, nous utilisons et démontrons la formule suivante dans notre cadre

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds.$$

Dans les Chapitres 4 et 5, le processus qui caractérise le système à N particules est de dimension N. Ceci rend l'étude directe de ces processus plus compliquée, car l'espace de travail dépend de N. Pour étudier ces modèles, nous utilisons le fait que les systèmes de particules que nous manipulons sont échangeables (i.e. la loi du système est invariante par permutations à support finie). Pour montrer la convergence en loi de systèmes échangeables finis vers un système échangeable infini, il suffit de montrer que les mesures empiriques des systèmes finis convergent vers ce qu'on appelle la mesure directrice du système infini. Pour montrer cette convergence, nous utilisons un nouveau type de problème martingale. Cette nouveauté vient du fait que les solutions du problème sont des lois de mesures aléatoires et non des lois de processus. Notre approche repose sur deux propriétés fondamentales : n'importe quelle limite en loi de la suite des mesures empiriques est solution du problème, et la loi de la mesure directrice du système infini est l'unique solution du problème.

La différence principale entre les Chapitres 4 et 5 est que dans le Chapitre 5, le modèle est énoncé dans un cadre McKean-Vlasov plus général et surtout qu'il y a une dépendance spatiale dans les interactions des systèmes. Dans le Chapitre 4, le terme d'interaction (qui est le même que celui des chapitres précédents) donne naissance à un bruit commun sous la forme d'un mouvement brownien. Dans le Chapitre 5, la dépendance spatiale des interactions donne naissance à un bruit commun sous la forme d'un bruit blanc.

Dans le Chapitre 6, nous nous intéressons à la normalisation linéaire dans le cadre des équations et des limites de type McKean-Vlasov. Nous montrons que ces équations sont bien posées (au sens fort) et la propagation du chaos. Ces questions sont classiques quand on considère des coefficients lipschitziens. Dans ce chapitre, nous considérons des coefficients localement lipschitziens. Contrairement au cadre Lipschitz, où les preuves reposent principalement sur le lemme de Grönwall, dans notre cadre, nous utilisons à la place le lemme d'Osgood avec un argument de troncature pour gérer la constante de Lipschitz locale. Ces deux arguments suffisent pour montrer l'unicité trajectorielle des équations différentielles stochastiques et la propagation du chaos, mais des difficultés techniques apparaissent pour montrer l'existence d'une solution forte des équations, ce que nous faisons via un schéma d'itération de Picard. Dans un premier temps, ce schéma d'itération ne nous permet que de construire une solution faible. Nous construisons une solution forte en utilisant une généralisation des résultats de Yamada et Watanabe.

Dans la première partie de cette thèse, les systèmes de particules que nous étudions sont des systèmes de processus de Hawkes. Ce sont des systèmes de processus ponctuels, dont les intensités sont solutions d'équations différentielles stochastiques de type convolution dirigées par ces processus ponctuels. Dans la seconde partie, nous étudions certaines questions introduites dans le Chapitre 4 dans un cadre plus général : les équations de type McKean-Vlasov. En effet, bien que les systèmes de particules finis introduits dans le Chapitre 4 ne soient pas liés a priori aux équations McKean-Vlasov, les équations différentielles stochastiques qui caractérisent le système limite sont du type McKean-Vlasov.

Un résumé plus détaillé de chaque chapitre se trouve à la Section 1.7. Commençons par rappeler certaines notions utiles sur les processus ponctuels et l'échangeabilité.

1.2 Processus ponctuel et mesure ponctuelle aléatoire

Dans cette section, nous rappelons quelques propriétés connues sur les processus ponctuels, qui nous seront utiles dans la suite. Les lecteurs peuvent trouver des études complètes de certains aspects des processus ponctuels dans les livres Daley and Vere-Jones (2003) et Ikeda and Watanabe (1989).

Définition 1.2.1. *Un processus ponctuel est un processus stochastique càdlàg, défini sur $\mathbb{R}_+$, constant par morceaux, dont les amplitudes de sauts sont toujours égales à un.*

Dans la littérature, la définition ci-dessus correspond à un processus ponctuel simple. Comme nous n'étudions que ce type de processus ponctuel, nous ne faisons pas la distinction.

Remarque 1.2.2. *Si $Z = (Z_t)_{t\geq 0}$ est un processus ponctuel, alors il existe une famille de variables aléatoires $(T_n)_{n\in\mathbb{N}^*}$ à valeurs dans $\mathbb{R}_+ \cup \{+\infty\}$ telle que :*

- *pour tout $n \in \mathbb{N}^*$, $T_n < T_{n+1}$,*
- *pour tout $t \geq 0$,*
$$Z_t = \sum_{n\in\mathbb{N}^*} n\mathbb{1}_{\{T_n\leq t<T_{n+1}\}}.$$

Ainsi, on peut toujours identifier un processus ponctuel avec un ensemble aléatoire de points sur $\mathbb{R}_+$, ce que l'on verra aussi comme une mesure ponctuelle aléatoire sur $\mathbb{R}_+$. Définissons formellement cette dernière notion.

La notion de mesure aléatoire n'est pas spécifique à $\mathbb{R}_+$, et dans la suite nous aurons besoin de cette notion sur des espaces plus généraux. Dans la suite, E désignera toujours un espace polonais muni de sa tribu borélienne.

Définition 1.2.3. *On note $\mathcal{M}(E)$ l'ensemble des mesures sur E muni de la tribu qui rend mesurable les applications $m \in \mathcal{M}_c(E) \mapsto m(B) \in \mathbb{R}_+ \cup \{+\infty\}$ (B borélien de E). On note $\mathcal{M}_c(E)$ le sous-ensemble des mesures de comptage sur E (i.e. les mesures qui s'écrivent comme une somme au plus dénombrable de masses de Dirac)*

Une mesure (resp. mesure ponctuelle) aléatoire sur E est une variable aléatoire à valeurs dans $\mathcal{M}(E)$ (resp. $\mathcal{M}_c(E)$).

Dans la suite, nous identifierons parfois processus ponctuel et mesure ponctuelle aléatoire définie sur $\mathbb{R}_+$. En effet, nous avons déjà vu comment définir un ensemble aléatoire de points à partir d'un processus ponctuel Z (voir Remarque 1.2.2), il suffit alors de considérer la mesure qui compte ces points. Réciproquement, si m est une mesure ponctuelle aléatoire, alors $Z_t := m([0, t])$ est un processus ponctuel.

Nous allons maintenant définir une classe importante de mesure ponctuelle aléatoire : les mesures de Poisson.

Définition 1.2.4. *Une mesure ponctuelle π sur E est une mesure de Poisson si :*

- *pour tout B borélien de E, $\pi(B)$ suit une loi de Poisson,*
- *pour tout $B_1, \dots, B_n$ boréliens disjoints, $(\pi(B_1), \dots, \pi(B_n))$ est une famille indépendante.*

Dans ce cas, on appelle intensité de π la mesure (déterministe) m définie par $m(B) := \mathbb{E}[\pi(B)]$.

Comme l'intensité d'une mesure de Poisson caractérise sa loi, nous introduirons toujours les mesures de Poisson par leur intensité.

Nous allons maintenant définir l'intensité stochastique d'un processus ponctuel, qui est une notion centrale pour définir et manipuler les processus de Hawkes.

Définition 1.2.5. *On dit qu'un processus progressivement mesurable $\lambda : t \in \mathbb{R}_+ \mapsto \lambda_t \in \mathbb{R}_+$ est l'intensité stochastique d'un processus ponctuel Z si, pour tout processus positif et prévisible $(C_t)_{t\geq 0}$,*

$$\mathbb{E}\left[\int_{\mathbb{R}_+} C_s dZ_s\right] = \mathbb{E}\left[\int_0^{+\infty} C_s \lambda_s ds\right].$$

Les mesures de Poisson sont fondamentales pour définir des processus ponctuels à partir de leur intensité stochastique. En effet, dans la suite, nous utiliserons souvent implicitement l'exemple et le lemme suivants.

Exemple 1.2.6. *Une mesure de Poisson sur $\mathbb{R}_+$ d'intensité $\lambda(t)dt$ (où λ est une fonction positive mesurable (déterministe)) admet comme intensité stochastique $\lambda_t = \lambda(t)$.*

Lemme 1.2.7. *Soit π une mesure de Poisson sur $\mathbb{R}_+ \times E$ d'intensité $dt \cdot d\nu(v)$, où ν est une mesure sur E $\sigma-$finie. Alors, pour tout processus positif et prévisible $(C_{t,v})_{(t,v)\in\mathbb{R}_+\times E}$, on a*

$$\mathbb{E}\left[\int_{\mathbb{R}_+\times E} C_{t,v} d\pi(t,v)\right] = \mathbb{E}\left[\int_0^{+\infty}\int_E C_{t,v} d\nu(v) dt\right].$$

Une conséquence immédiate de ce lemme est la suivante : si $(\lambda_t)_{t\geq 0}$ est un processus prévisible positif, et π une mesure de Poisson sur $\mathbb{R}_+ \times \mathbb{R}_+$ d'intensité $dt \cdot dz$, alors $Z_t := \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\leq\lambda_s\}} d\pi(s,z)$ est un processus ponctuel d'intensité stochastique $(\lambda_t)_{t\geq 0}$.

Cette construction de processus ponctuels est fondamentale, car on peut toujours écrire un processus ponctuel admettant une intensité stochastique prévisible sous cette forme (voir Lemme 4 de Brémaud and Massoulié (1996)).

Lemme 1.2.8. *Soit Z un processus ponctuel admettant une intensité stochastique prévisible $(\lambda_t)_{t\geq 0}$. Alors, il existe une mesure de Poisson π sur $\mathbb{R}_+ \times \mathbb{R}_+$ d'intensité $dt \cdot dz$ telle que, pour tout $t \geq 0$,*

$$Z_t = \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\leq\lambda_s\}} d\pi(s,z). \tag{1.1}$$

Dans la suite, nous aurons besoin d'introduire des processus ponctuels sous une forme similaire à (1.1), car nous utiliserons un paramètre de bruit, donné par la mesure de Poisson.

La notion d'intensité stochastique repose sur celle de prévisibilité, elle dépend donc de la filtration sur laquelle on se place. Dans toute la suite, quand nous étudierons un modèle, nous utiliserons toujours, implicitement, la filtration générée par tous les processus du modèle. La cohérence de ce choix de filtration avec la notion d'intensité stochastique et de prévisibilité est garantie par le lemme suivant (voir Proposition III.2 de Chevallier (2013)).

Lemme 1.2.9. *Soit $Z = (Z_t)_{t\geq 0}$ un processus ponctuel admettant $(\lambda_t)_{t\geq 0}$ comme $(\mathcal{F}_t)_t$ −intensité. Soit $(\mathcal{G}_t)_{t\geq 0}$ une filtration telle que, pour tout $t \geq 0$, $\mathcal{G}_t$ est indépendante de $\mathcal{F}_t$. Alors Z admet $(\lambda_t)_{t\geq 0}$ comme $(\mathcal{F}_t \vee \mathcal{G}_t)_t$ −intensité.*

Terminons cette section avec un lemme important qui est souvent utilisé de manière plus ou moins implicite.

Lemme 1.2.10. *Soit $(Z_t)_t$ un processus ponctuel qui admet une intensité stochastique $(\lambda_t)_t$. Alors le processus*

$$Z_t - \int_0^t \lambda_s ds$$

est une martingale locale appelée processus ponctuel compensé.

1.3 Échangeabilité

Dans les Chapitres 4 et 5, nous utilisons beaucoup les propriétés des systèmes dits échangeables. Dans la suite, nous appelons mesure de probabilité aléatoire (ou loi aléatoire) sur un espace mesurable E, une mesure aléatoire m (au sens de la Définition 1.2.3) tel que $m(E) = 1$ presque sûrement.

Définition 1.3.1. *Soient E un espace mesurable et $(X_i)_{i\in I}$ un système de v.a. sur E (fini ou infini). On dit que $(X_i)_{i\in I}$ est échangeable si pour toute permutation σ de I à support fini,*

$$\mathcal{L}((X_i)_{i\in I}) = \mathcal{L}((X_{\sigma(i)})_{i\in I}).$$

Nous pouvons déjà remarquer qu'un système i.i.d. est l'exemple le plus simple de système échangeable.

Cette notion est étroitement liée à la notion de mélange.

Définition 1.3.2. *Soient E un espace mesurable, $(X_i)_{i\in I}$ un système de v.a. sur E, et m une mesure de probabilité aléatoire sur E. On dit que le système $(X_i)_{i\in I}$ est un mélange dirigé par m si, conditionnellement à m, les variables X_i $(i \in I)$ sont i.i.d. de loi m. Cette mesure aléatoire m est unique si elle existe (voir Lemme (2.15) de Aldous (1983)), et elle est appelée la mesure directrice du système.*

Si $X = (X_i)_{i\in I}$ est un mélange dirigé par une mesure m de loi Q, alors on peut formellement écrire

$$\mathcal{L}(X)(\cdot) = \int_{\mathcal{P}(E)} \eta^{\otimes I}(\cdot) dQ(\eta).$$

Le lien entre la notion de mélange et la notion d'échangeabilité est donné par le théorème de de Finetti (Théorème (3.1) de Aldous (1983)) :

Théorème 1.3.3. *Soient E un espace polonais et $(X_i)_{i\in I}$ un système infini de v.a. sur E. Alors le système $(X_i)_{i\in I}$ est échangeable si et seulement s'il s'écrit comme un mélange dirigé par une certaine mesure aléatoire.*

Décrivons maintenant deux méthodes pour déterminer en pratique la mesure directrice m d'un système infini échangeable $(X_i)_{i\geq 1}$. La première est une application conditionnelle du théorème de Glivenko-Cantelli. En effet, en rappelant que conditionnellement à m, les variables X_i $(i \geq 1)$ sont i.i.d. de loi m, on sait que m doit être la limite presque sûre de la suite des mesures empiriques

$$m^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{X_i}.$$

La deuxième méthode s'appuie sur le Lemme (2.12) de Aldous (1983). D'après ce lemme, s'il existe une tribu $\mathcal{F}$ telle que, conditionnellement à $\mathcal{F}$, les variables X_i $(i \geq 1)$ sont i.i.d., alors le système $(X_i)_{i\geq 1}$ est un mélange dirigé par $\mathcal{L}(X_1|\mathcal{F})$. Autrement dit, si on arrive à identifier le bruit commun d'un système échangeable, alors la mesure directrice s'écrit comme la loi conditionnelle de n'importe laquelle des coordonnées sachant ce bruit commun.

Terminons cette section avec un résultat central dans l'étude de la convergence de grande échelle de systèmes échangeables (il s'agit de la Proposition (7.20) de Aldous (1983)).

Théorème 1.3.4. *Soit E un espace polonais. Soient $(X_k^N)_{1\leq k\leq N}$ et $(X_k)_{k\geq 1}$ $(N \in \mathbb{N}^*)$ des systèmes échangeables sur E, m^N la mesure empirique de X^N $(N \in \mathbb{N}^*)$ et m la mesure directrice de X. Les assertions suivantes sont équivalentes :*

- *$(X_k^N)_{1\leq k\leq N}$ converge en loi vers $(X_k)_{k\geq 1}$ sur $E^{\mathbb{N}^*}$ muni de la topologie produit.*
- *m^N converge en loi vers m sur $\mathcal{P}(E)$ muni de la topologie de la convergence faible.*

Nous invitons les lecteurs qui cherchent un cours complet sur la notion d'échangeabilité à lire le cours d'Aldous donné à Saint-Flour Aldous (1983).

Nous venons de faire des rappels sur deux notions générales : les processus ponctuels et l'échangeabilité. Passons maintenant à une notion plus spécifique à cette thèse : les processus de Hawkes.

1.4 Processus de Hawkes

Dans la littérature, il existe des variantes dans la définition des processus de Hawkes. Informellement, nous pouvons définir un processus de Hawkes comme un système de processus ponctuels dont les intensités stochastiques sont solution d'équations de type convolution dirigées par les processus ponctuels du système.

Dans cette thèse, nous appelons système de processus de Hawkes, tout système de processus ponctuels $(Z^i)_{i\in I}$ dont les intensités λ^i $(i \in I)$ vérifient une équation similaire à (1.2) ci-dessous. Nous ne donnons pas de définition générale de tels systèmes, car nous étudions différents modèles qui ne vérifieraient pas une même définition.

$$\lambda_t^i = f^i\left(\sum_{j\in I}\int_{[0,t[} h^{ji}(t-s)dZ_s^j\right). \tag{1.2}$$

Dans la suite, quand nous manipulons des processus de Hawkes, nous utilisons les formules du même type que (1.2). Signalons qu'il peut aussi être utile d'écrire cette formule différemment : en notant, pour chaque $i \in I$, T_k^i $(k \geq 1)$ les instants de sauts du processus ponctuel Z^i, il est possible de définir λ^i comme

$$\lambda_t^i = f^i \left(\sum_{j \in I} \sum_{k \geq 1} \mathbb{1}_{\{T_k^j < t\}} h^{ji}(t - T_k^j) \right).$$

L'avantage de cette formule est que l'on voit bien que chaque point des processus ponctuels apportent une contribution aux intensités stochastiques. En particulier, quand on étudie des processus de Hawkes linéaires (i.e. quand les fonctions f^i $(i \in I)$ sont affines et que les fonctions h^{ji} $(i, j \in I)$ sont positives), on peut interpréter ces processus commes des processus de branchements : si $f^i(x) = ax + b$ (avec $a, b > 0$), quand le processus ponctuel Z^i crée un point, soit ce point est créé par le taux de base b (un tel point est un "migrant"), soit il est créé par la contribution d'un autre point (un tel point est un "descendant"). Nous n'étudierons pas cet aspect des processus de Hawkes dans cette thèse.

1.5 Applications des processus de Hawkes

Les processus de Hawkes ont été introduits par Hawkes (1971) pour modéliser les tremblements de terre au Japon. Récemment, les processus de Hawkes ont connu un regain de popularité, car ils ont permis de modéliser de manière pertinente des phénomènes dans des domaines variés. Nous pouvons citer des applications en séisméologie (Helmstetter and Sornette (2002), Y. Kagan (2009), Ogata (1999)), en finance (Aït-Sahalia et al. (2015), Lu and Abergel (2018), Bauwens and Hautsch (2009), Hewlett (2006)), en génomique (Reynaud-Bouret and Schbath (2010)), en réseaux sociaux (Zhou et al. (2013)) et en neurosciences (Grün et al. (2010), Pillow et al. (2008), Reynaud-Bouret et al. (2014)).

L'idée derrière les modélisations dans ces différents domaines est toujours la même : on modélise un réseau de particules (ou une seule particule) qui effectuent chacune une action avec un certain taux. Quand une particule effectue une action, cette action modifie les taux auquels les autres particules effectuent les leurs. Autrement dit, les particules s'excitent et/ou s'inhibent mutuellement. Détaillons deux de ces applications.

1.5.1 Modélisation d'un réseau de neurones

Considérons un réseau de N neurones numérotés de 1 à N. Le fonctionnement d'un réseau de neurones repose sur le fait que les neurones s'envoient des décharges les uns aux autres. On considère souvent que l'activité d'un tel réseau est caractérisé par les instants auxquels ces décharges ont lieu.

Pour chaque $1 \leq i \leq N$, notons Z^i le processus ponctuel qui compte les instants de décharge du neurone i (i.e. $Z^i([0,t])$ est le nombre de décharges que le neurone i a envoyées avant le temps t). Dans ce modèle, chaque neurone i envoie ses décharges au taux aléatoire $\lambda_t^i = f^i(X_{t-}^i)$ défini par (1.2), où

$$X_t^i := \sum_{j=1}^N \int_{[0,t]} h^{ji}(t-s) dZ_s^j$$

représente le potentiel de membrane du neurone i à l'instant t.

La formule ci-dessus s'interprète de la manière suivante : quand un neurone j envoie une décharge à un instant s, cette décharge modifie les valeurs des potentiels X_t^i $(t \geq s)$ pour tous les neurones i pour lesquels h^{ji} n'est pas la fonction nulle. Plus précisément, à l'instant s de la décharge, chaque X^i fait un saut d'amplitude $h^{ji}(0)$, puis l'effet de la décharge évolue au cours du temps, et cette évolution est décrite par la fonction h^{ji}.

Autrement dit, quand un neurone envoie une décharge à un autre neurone, le potentiel de membrane de cet autre neurone est modifié, et donc son taux de décharge aussi. C'est ce qui correspond à une excitation ou à une inhibition, selon si le taux augmente ou diminue. Dans le cas où les fonctions f^i sont croissantes, dire que la fonction h^{ji} est positive (resp. négative) signifie que les décharges du neurone j excitent (resp. inhibent) le neurone i. On peut aussi imaginer des fonctions h^{ji} ni positives, ni négatives. Dans ce cas, l'effet des décharges du neurone j sur le neurone i (i.e. excitatrice ou inhibitrice) dépend du moment où elles ont eues lieu. Ce type de dynamique est impossible en pratique en neurosciences si $i \neq j$: étant donné une relation entre deux neurones, cette relation est soit excitatrice soit inhibitrice. Toutefois, il peut être intéressant de considérer une fonction h^{ii} très négative au voisinage de zéro puis égale à zéro pour modéliser la période réfractaire du neurone après ses décharges.

En pratique, les fonctions f^i et h^{ji} $(1 \leq i, j \leq N)$ ont quelques propriétés remarquables. En effet, il est naturel de considérer des fonctions h^{ji} qui tendent vite vers zéros en l'infini (par exemple, des fonctions exponentielles négatives) car les potentiels de membrane des neurones retournent à leurs valeurs de repos (que l'on suppose égales à zéro) relativement vite. En ce qui concerne les fonctions f^i, on s'intéresse à des fonctions qui "alternent" entre deux états. Expliquons ce qu'on entend par là sur des exemples :

$$\begin{cases} f(x) = 0 & \text{si } x < x_0 \\ f(x) = 1 & \text{si } x \geq x_0 \end{cases}, \quad \begin{cases} f(x) = 0 & \text{si } x < 0 \\ f(x) = (x/x_0)^{10} & \text{si } x \geq 0 \end{cases}, \quad f(x) = \frac{\pi}{2} + \arctan(x - x_0).$$

La propriété commune aux exemples de fonctions f ci-dessus, est une transition de phase rapide entre deux états. En effet, pour $x < x_0$, la valeur de $f(x)$ est proche de 0, et pour $x > x_0$, $f(x)$ est proche de sa borne supérieure. On peut interpréter cette propriété du point de vue de la modélisation en neurosciences de la manière suivante : si le potentiel X_t^i du neurone i est inférieur à x_0, son taux de décharge est presque nul, donc la probabilité que le neurone i n'envoie aucune décharge est élevée. Le neurone i est alors considéré comme étant inactif. Inversement, si $X_t^i > x_0$, le taux de décharge est presque optimal, et donc le neurone i est dans l'état actif. C'est cette dualité entre l'état actif et l'état inactif qui est intéressante du point de vue des neurosciences.

Cette modélisation est bien entendu très simplifiée par rapport à la réalité. Au Chapitre 4, nous verrons un modèle proche de celui-là, dans lequel nous prendrons en compte une autre propriété des neurones : quand un neurone envoie une décharge, son potentiel de membrane retourne presque immédiatement à sa valeur de repos.

1.5.2 Modélisation en génomique

Dans cette modélisation en génomique, la notion d'excitation et d'inhibittion intervient dans le processus de régulation des gènes. Avant d'expliquer où interviennent les processus de Hawkes, détaillons ce processus de régulation. Il faut d'abord savoir ce qu'est un gène et ce qu'est l'ADN. L'ADN est une molécule formée de deux séquences complémentaires de bases nucléiques. Il existe quatre bases nucléiques : adénine (A), thymine (T), cytosine (C) et guanine (G). Pour simplifier l'explication, nous ferons comme s'il n'y avait qu'une seule séquence.

Les gènes sont des fragments d'ADN (ce sont donc des séquences de bases nucléiques) qui ont un site promoteur situé sur l'ADN (ce site est aussi une séquence de bases nucléiques). Un gène peut produire des protéines (on dit alors qu'il est actif) qui vont ensuite exciter ou inhiber la production de protéines d'autres gènes. Pour qu'un gène soit actif, il faut qu'une ARN polymérase s'accroche au site promoteur du gène pour initier les phases de transcription et de traduction (ce qui résulte en la production de protéines).

Il faut savoir que les protéines sont aussi capables de s'accrocher à l'ADN, comme le fait l'ARN polymérase. Chaque protéine aura une certaine affinité avec chaque séquence de bases nucléiques. Il y a donc, pour chaque protéine, des sites sur l'ADN où la protéine peut se fixer. Ce sont des sites opérateurs. Si une protéine a un site opérateur proche du site promoteur d'un gène, alors cette protéine peut influencer l'activation de ce gène de deux manières : une protéine peut être un inhibiteur ou un excitateur d'un gène.

L'inhibition est un processus simple : il suffit d'imaginer une configuration dans laquelle un site opérateur et un site promoteur se chevauchent, de telle sorte qu'il soit impossible que la protéine inhibitrice et l'ARN polymérase associées à ces sites soient liées en même temps. Autrement dit, la protéine associée à ce site opérateur peut empêcher l'ARN polymérase de se lier au site promoteur du gène, et ainsi empêcher le gène de s'activer. Ceci est illustré sur la figure 1.1 où deux sites ont une base nucléique en commun.

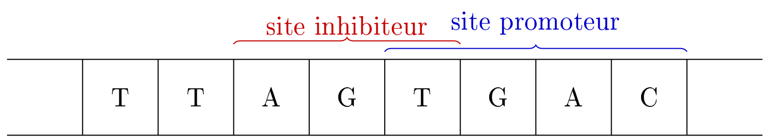

FIGURE 1.1 : Exemple de fragment d'ADN avec un site inhibiteur d'un site promoteur

Pour comprendre l'excitation, il faut savoir que les gènes peuvent se lier entre eux, en plus de se lier à l'ADN. Ainsi, si un site excitateur est très proche d'un site promoteur, alors quand la protéine et l'ARN polymérase vont se lier sur les deux sites, elles vont aussi se lier entre elles. Par conséquent, la protéine associée au site excitateur peut aider le gène à produire des protéines, puisque si elle se lie au site opérateur, alors l'ARN polymérase aura une meilleure affinité avec le site promoteur car, en plus des liaisons avec l'ADN, elle aura des liaisons avec la protéine. Le site excitateur doit donc être suffisamment proche du site promoteur pour que les protéines puissent se lier à l'ARN polymérase.

On vient de décrire l'excitation et l'inhibition du site promoteur, mais il est aussi possible de d'exciter ou d'inhiber des sites opérateurs de la même façon, ce qui peut conduire à des régulations complexes.

Les fragments d'ADN qui interviennent dans ces processus d'activation des gènes s'appellent des Transcription Regulatory Elements (TRE). Par exemple, les séquences de bases nucléiques qui correspondent à des sites promoteurs, excitateurs ou inhibiteurs sont des TRE. Le même TRE peut apparaitre à différents endroits de l'ADN.

Nous pouvons maintenant expliquer comment modéliser ces phénomènes avec les processus de Hawkes. Contrairement aux autres modélisations, les variables t et s qui apparaissent dans la formule (1.2) ne représentent pas des variables de temps, mais des variables d'espace. Chaque élément $t \in \mathbb{R}_+$ est une position sur la séquence d'ADN. Supposons que l'on s'intéresse à un nombre fini de TRE différents que l'on numérote de 1 à N. Alors nous considérons un système de processus de Hawkes $(Z^i)_{1 \le i \le N}$ de dimension N. Pour chaque $1 \le i \le N$, les atomes de Z^i sont les positions de l'ADN où le TRE numéro i est actif (i.e. pour un site promoteur, avec une ARN polymérase, et pour un site opérateurs, avec une protéine).

Nous pouvons donc interpréter la formule (1.2) dans ce contexte de la manière suivante : dire que l'un des processus ponctuels Z^i $(1 \le i \le N)$ charge un point à un instant $t \in \mathbb{R}$, signifie qu'il y a un TRE i actif à la position t. Ce TRE peut donc influencer les probabilités que d'autres TRE arrivent après t. Donnons un exemple concret d'application. Pour faire simple, nous considérerons trois TRE : TRE1 qui correspond à un site promoteur, TRE2 (resp. TRE3) un site excitateur (resp. inhibiteur) de TRE1 tels que TRE2 et TRE3 ont une sous-séquence de nuclétoides en commun (voir figure 1.2). Fixons des fonctions f^i croissantes qui tendent vers 0 en $-\infty$. Dans cet exemple, il serait cohérent de choisir la fonction nulle pour h^{1i}, h^{32} et h^{ii} $(1 \le i \le 3)$, des fonctions "très négatives" pour h^{23} et h^{31} (voir figure 1.2), et une fonction positive pour h^{21}.

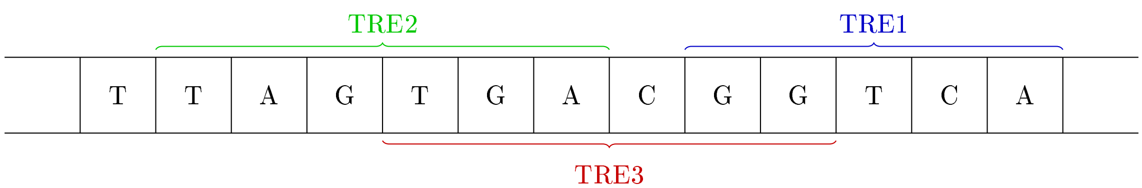

FIGURE 1.2 : Exemple d'application des processus de Hawkes à la génomique

Dans ce contexte, il est naturel de considérer des fonctions $h^{ji} : \mathbb{R}_+ \to \mathbb{R}$ à support compact. En effet, un TRE n'aura d'effet qu'à une distance relativement courte.

L'un des désavantages de cette modélisation est que l'ADN est orienté dans un seul sens. Sur l'exemple précédent (qui correspond à la figure 1.2), l'ADN est orientée de gauche à droite. Cela vient de la formule (1.2) où chaque atome t des processus ponctuels modifie les intensités après t, mais pas avant. Par exemple, sur la figure 1.2, on peut modéliser le fait que TRE2 inhibe TRE3, et que TRE3 inhibe TRE1 (de gauche à droite), mais pas que TRE3 inhibe TRE2 et que TRE1 inhibe TRE3 (de droite à gauche).

1.6 Limite de grande échelle de processus de Hawkes

La question à laquelle nous allons nous intéresser dans la première partie cette thèse est la convergence des systèmes de processus de Hawkes quand la dimension du système tend vers l'infini. Plus précisément, pour chaque $N \in \mathbb{N}^*$, nous considérons un système de processus de Hawkes $\left(Z^{N,1}, \dots, Z^{N,N}\right)$ tel que chaque $Z^{N,i}$ admet une intensité stochastique de la forme

$$\lambda_t^N = f\left(\sum_{j=1}^{N} \int_{[0,t[} h^N(t-s)dZ_s^{N,j}\right). \tag{1.3}$$

Pour que l'expression dans (1.3) converge, il faut normaliser la somme pour la forcer à converger. La normalisation la plus naturelle pour forcer cette somme à converger est la normalisation linéaire en N^{-1}, pour se ramener à une sorte de loi des grands nombres. Ce modèle a été étudié par Delattre et al. (2016), et la convergence du système (1.3) a été démontrée dans leur Théorème 8. L'autre normalisation qui nous intéressera plus dans cette thèse est la normalisation diffusive en $N^{-1/2}$.

La question de la convergence de systèmes de N particules quand N tend vers l'infini intervient naturellement dans le cadre de la modélisation des réseaux de neurones (voir Section 1.5.1). En effet, bien que la taille d'un réseau de neurones est finie, elle est quand même très grande. Il est donc logique d'approcher un système de taille $N \gg 1$ par un système de taille infinie. L'intérêt d'une telle approximation repose sur le fait qu'un système de taille infinie peut être plus simple à étudier. Ceci vient du fait que simuler N processus ponctuels quand N est grand peut s'avérer algorithmiquement coûteux. Mais dans le modèle limite, ces processus ponctuels vont soit se moyenniser (dans le cadre de la modélisation linéaire) soit se "transformer" en mouvement brownien. Dans les deux cas, il existe des techniques pour modéliser ces termes limites.

Expliquons comment comprendre la forme possible des systèmes de processus ponctuels limites. Pour chaque entier naturel N, considérons le système de processus ponctuels $(Z^{N,1}, ..., Z^{N,N})$ défini par (1.3). Supposons que ce système converge quand N tend vers l'infini vers un système de processus que nous notons $(\bar{Z}^i)_{i\geq 1}$. Expliquons comment fonctionne cette convergence dans les cadres des deux normalisations mentionnées plus tôt (i.e. la normalisation linéaire en N^{-1} et la normalisation diffusive en $N^{-1/2}$), et quelle est la dynamique du système limite associé.

1.6.1 Normalisation linéaire

Le type de dynamique considérée dans le cadre de la normalisation linéaire est la suivante :

$$\lambda_t^N = f\left(\frac{1}{N}\sum_{j=1}^{N}\int_{[0,t[} h(t-s)dZ_s^{N,j}\right). \tag{1.4}$$

Si nous la comparons avec (1.3), nous avons juste écrit $h^N = N^{-1}h$ où h est une fonction qui ne dépend pas de N. De manière informelle, d'après les résultats de type "loi des grands nombres" (bien que les termes de la somme ne sont pas indépendants) l'intensité limite $\bar{\lambda}$ devrait exister et s'écrire comme

$$\bar{\lambda}_t = f\left(\mathbb{E}\left[\int_{[0,t[} h(t-s)d\bar{Z}_s^1\right]\right),$$

avec $\bar{Z}^1$ un processus ponctuel d'intensité $(\bar{\lambda}_t)_{t\geq 0}$. En particulier, $\bar{\lambda}$ est une fonction déterministe qui vérifie l'équation fonctionnelle suivante

$$\bar{\lambda}_t = f\left(\int_0^t h(t-s)\bar{\lambda}_s ds\right).$$

Notons que l'équation ci-dessus est l'équivalent de l'équation (8) de Delattre et al. (2016).

En rappelant que les interactions des systèmes à N particules $(Z^{N,i})_{1\leq i\leq N}$ viennent de l'intensité λ^N, nous pouvons remarquer que le système limite $(\bar{Z}^i)_{i\geq 1}$ est un système i.i.d. caractérisé par une intensité déterministe. C'est pour cette raison que l'on qualifie cette convergence de propagation du chaos.

1.6.2 Normalisation diffusive

En normalisation diffusive, il faut prendre plus de précautions pour définir les modèles correctement. En effet, si on reconnait une application de pseudo-loi des grands nombres dans le cadre linéaire, le résultat correspondant à la normalisation diffusive est le théorème central limite. Pour utiliser ce résultat, il ne suffit pas de normaliser la somme de (1.3) en $N^{-1/2}$, il faut aussi centrer les termes de la somme. Une première idée pour définir un tel modèle serait de se donner une collection de variables aléatoires i.i.d. $U^i(t)$ $(i \in \mathbb{N}^*, t \in \mathbb{R}_+)$, et de définir l'intensité λ^N par

$$\lambda_t^N = f\left(\frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[} U^j(s)h(t-s)dZ_s^{N,j}\right). \tag{1.5}$$

Il y a toutefois un problème dans la formulation (1.5) : l'intégrande n'est pas prévisible, puisque, pour chaque $j \geq 1$ fixé, le processus $s \mapsto U^j(s)$ ne peut pas être prévisible. Cela se comprend bien intuitivement car dire qu'un processus X est prévisible revient à dire que pour tout instant t, la valeur de X_t peut être déduite de la trajectoire de X sur $[0, t[$. Or, par construction, la variable $U^j(t)$ est indépendante de toutes les variables $U^j(s)$ avec $s < t$. C'est pour cela que le processus U^j n'est pas prévisible.

Remarque 1.6.1. *Le fait que l'intégrande de (1.5) n'est pas prévisible n'empêche pas de définir l'intégrale : à ω fixé, les processus ponctuels $Z^{N,j}$ sont des vraies mesures ponctuelles, donc les intégrales ont bien un sens presque sûrement. Le problème de la non-prévisibilité, c'est qu'il est compliqué, voir impossible de faire des calculs sur ce genre d'intégrale, car le Lemme 1.2.7 ne s'applique pas. Ce qui implique notamment que l'intégrale n'a pas les bonnes propriétés des intégrales stochastiques, c'est-à-dire que l'intégrale compensée n'est pas, a priori, une martingale locale.*

L'idée pour bien définir l'intensité λ^N consiste à écrire directement les processus ponctuels $Z^{N,j}$ comme des amincissements de mesures de Poisson (i.e. sous la forme (1.1)). L'intérêt de cette écriture est de pouvoir introduire des variables centrées en tant que variables de ces mesures de Poisson plutôt qu'en tant que variables $U^j(s)$ introduites à part. Plus précisément, nous considérons des intensités λ^N définies par

$$\lambda_t^N = f\left(\frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[\times\mathbb{R}_+\times\mathbb{R}} uh(t-s)\mathbb{1}_{\{z\leq\lambda_s^N\}}d\pi^j(s,z,u)\right), \tag{1.6}$$

où les π^j $(j \geq 1)$ sont des mesures de Poissons i.i.d. d'intensité $dt \cdot dz \cdot d\mu(u)$, avec μ une loi sur $\mathbb{R}$ centrée. Le paramètre t représente le temps, le paramètre z permet de faire l'amincissement pour retrouver la bonne intensité de saut λ^N, et le paramètre u correspond aux variables $U^j(s)$ de la formule (1.5). Informellement, les deux formules (1.5) et (1.6) définissent la même dynamique. Dans la suite nous ne manipulerons que les expressions du second type, car, comme nous l'avons expliqué plus tôt, la première formule est mal posée.

Expliquons maintenant comment doit se comporter l'intensité limite. Contrairement à la normalisation linéaire où l'intensité limite est déterministe (par une "loi des grands nombres"), ici l'intensité limite est aléatoire. D'après un "théorème central limite", l'intensité limite $\bar{\lambda}$ devrait être dirigé par un processus gaussien, et ce processus devrait être à accroissements indépendants : ce processus devrait donc être un mouvement brownien W. Ainsi, l'intégrale de (1.6) devrait tendre vers une intégrale stochastique dirigée par W dont l'intégrande est la racine carrée de l'intégrande de la limite des variations quadratiques de l'intégrale de (1.6). Donc, en remarquant que

$$\left\langle \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[\times\mathbb{R}_+\times\mathbb{R}} uh(t-s)\mathbb{1}_{\{z\leq\lambda_s^N\}}d\pi^j(s,z,u)\right\rangle_t = \sigma^2\int_0^t h(t-s)^2\lambda_s^N ds,$$

où σ^2 est la variance de la loi μ, l'intensité limite $\bar{\lambda}$ devrait vérifier

$$\bar{\lambda}_t = f\left(\sigma \int_0^t h(t-s)\sqrt{\bar{\lambda}_s} dW_s\right).$$

Ce genre de processus, solution d'équation de convolution dirigée par un mouvement brownien s'appelle un processus de Volterra. L'étude des processus de Volterra est une question délicate que nous ne traiterons pas. Dans la suite de la thèse, nous étudierons des processus de Hawkes avec des noyaux de convolutions particuliers qui nous permettront de nous ramener à des équations différentielles stochastiques classiques. Les lecteurs qui s'intéressent aux noyaux de convolutions généraux peuvent se reporter à la Section 7.2.1 pour trouver des pistes de réflexion.

1.7 Organisation

Dans la première partie de cette thèse, nous étudierons des limites de grande échelle de processus de Hawkes en normalisation diffusive. Le premier modèle que nous verrons sera un système de processus de Hawkes de dimension N où tous les processus ponctuels d'un même système auront la même intensité stochastique. Ce premier modèle sera donc relativement classique, car il consiste à étudier la convergence d'une suite de processus réels de dimension un. Le deuxième modèle sera une généralisation du premier à un cadre multi-populations. Les techniques de preuves seront les mêmes. Dans le troisième modèle, nous ajouterons des éléments aux modèles qui sont naturels du point de vue de la modélisation des réseaux de neurones. Ces éléments nous éloigneront du cadre de travail des équations différentielles stochastiques classiques, en faisant naturellement intervenir des équations de type McKean-Vlasov conditionnel (i.e. des équations différentielles stochastiques où les coefficients dépendent de la loi conditionnelle de la solution) dans le système limite.

La deuxième partie de la thèse généralisera naturellement la première. En effet, elle sera consacrée à l'étude de la limite de grande échelle de systèmes de N particules de type McKean-Vlasov (i.e. des systèmes d'équations différentielles stochastiques dont les coefficients dépendent de la mesure empirique du système). Nous étudierons deux modèles : le premier sera en normalisation diffusive $N^{-1/2}$, comme ceux de la première partie, et le second en normalisation linéaire N^{-1}. Les équations de McKean-Vlasov sont naturelles dans le cadre de travail des systèmes de particules en champ moyen : ce phénomène peut être interprété comme une loi des grands nombres. En effet, dans les exemples où la dynamique du système à N particules est dirigée par une équation différentielle stochastique, les interactions en champ moyen peuvent s'exprimer comme une dépendance des coefficients par rapport à la mesure empirique du système. Et quand N tend l'infini, cette mesure empirique converge vers la loi de n'importe quelle particule du système limite. Par conséquent, les coefficients de l'équation différentielle stochastique du système limite dépendent naturellement de la loi de la solution de cette équation. On peut trouver des exemples illustrant ce phénomène en modélisation de systèmes de neurones dans De Masi et al. (2015) et Fournier and Löcherbach (2016), en modélisation de portfolio dans Fischer and Livieri (2016), et en jeux en champ moyen dans Carmona et al. (2016).

Détaillons rapidement le contenu de chacun des chapitres de ces deux parties.

1.7.1 Chapitres 2 et 3 : Processus de Hawkes avec noyaux exponentiels et d'Erlang

Dans ces deux chapitres, nous manipulons des systèmes de processus de Hawkes de la forme suivante :

$$\left\{\begin{array}{ll} Z_t^{N,i} = & \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\{z\leq\lambda_s^N\}} d\pi^i(s,z,u) \quad (1\leq i\leq N), \\ \lambda_t^N = & f\left(\frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[\times\mathbb{R}_+\times\mathbb{R}} u\cdot h(t-s)\mathbb{1}_{\{z\leq\lambda_s^N\}} d\pi^j(s,z,u)\right), \end{array}\right. \tag{1.7}$$

avec $\left(\pi^i\right)_{i\in\mathbb{N}^*}$ une famille i.i.d. de mesures de Poisson sur $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}$ d'intensité $dt\cdot dz\cdot d\mu(u)$, et μ une mesure de probabilité sur $\mathbb{R}$ centrée.

La convergence des processus ponctuels repose principalement sur la convergence de leur intensité stochastique. Dans notre cas, cette intensité est de la forme $\lambda_t^N = f\left(X_{t-}^N\right)$, avec

$$X_t^N := \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\cdot h(t-s)\mathbb{1}_{\left\{z\leq f\left(X_{s-}^N\right)\right\}} d\pi^j(s,z,u).$$

La plupart des résultats de ces chapitres concernent le processus $\left(X_t^N\right)_{t\geq 0}$. Dans chaque modèle que nous étudions, ces résultats viendront du fait que X^N est solution d'une équation différentielle stochastique dirigée par les mesures de Poisson $\left(\pi^i\right)_{1\leq i\leq N}$. La raison pour laquelle dans nos modèles X^N est solution d'une équation différentielle stochastique, est que l'on considère des noyaux de convolution h particuliers. En effet, dans le Chapitre 2, nous prenons une fonction h exponentielle $h(t) := e^{-\alpha t}$, ce qui implique que X^N est solution de

$$dX_t^N = -\alpha X_t^N + \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\mathbb{1}_{\left\{z\leq f(X_{s-}^N)\right\}} d\pi^j(s,z,u). \tag{1.8}$$

Dans le Chapitre 3, la fonction h est plus générale car elle s'écrit comme produit d'une fonction exponentielle et d'un polynôme. Dans ce cas, X^N n'est pas directement solution d'une équation différentielle stochastique, mais nous pouvons écrire X^N comme une coordonnée d'un processus multi-dimensionnel qui est solution d'une équation différentielle stochastique. Notons que la dimension de ce processus ne dépend pas de N, et que c'est pour cela que les techniques de preuve du Chapitre 2 s'adaptent aisément au Chapitre 3.

Le but du Chapitre 2 est de montrer quatre résultats : la convergence en loi du processus X^N vers un processus limite $\bar{X}$, expliciter la vitesse de convergence associée à l'erreur faible, la convergence du noyau markovien de X^N vers la mesure invariante de $\bar{X}$ et la convergence des processus ponctuels $Z^{N,i}$.

D'après le raisonnement de la Section 1.6.2, $\bar{X}$ devrait a priori être solution de

$$d\bar{X}_t = -\alpha\bar{X}_t dt + \sigma\sqrt{f(\bar{X}_t)}dW_t,$$

avec W un mouvement brownien. Ce mouvement brownien W correspond à la limite de la contribution des mesures de Poisson de (1.8). Cette limite peut être interprétée comme un théorème central limite.

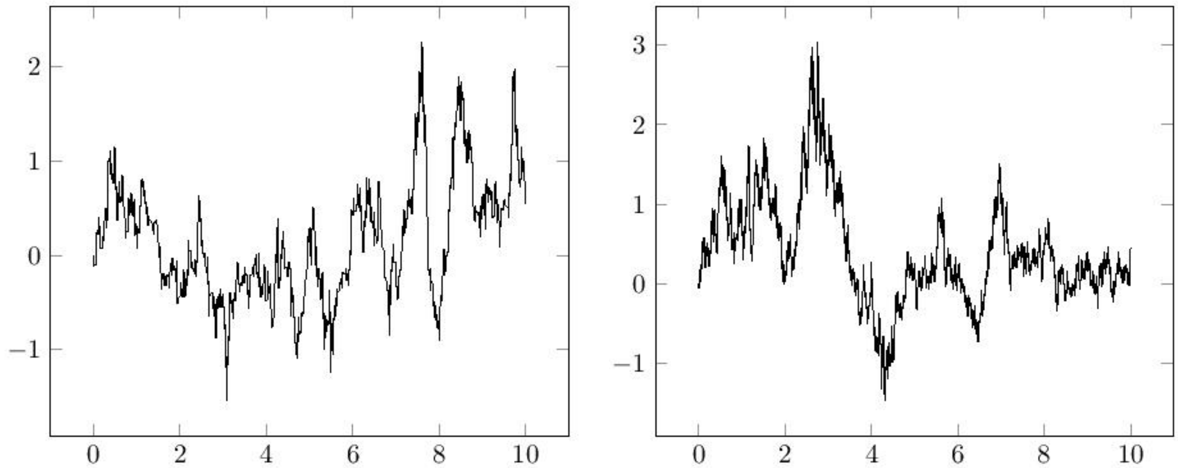

FIGURE 1.3 : Simulations de trajectoires de $(X_t^N)_{0\leq t\leq 10}$ avec $X_0^N = 0$, $\alpha = 1$, $\mu = \mathcal{N}(0,1)$, $f(x) = 1 + x^2$, $N = 100$ (image de gauche) et $N = 500$ (image de droite).

Nous montrons la convergence de X^N vers $\bar{X}$ en utilisant des résultats généraux sur les semi-martingales : nous montrons que le triplet caractéristique de X^N converge vers celui de $\bar{X}$. Ensuite, nous obtenons une vitesse de convergence du semi-groupe P^N de X^N vers le semi-groupe $\bar{P}$ de $\bar{X}$. Pour cela, nous utilisons la vitesse de convergence de leurs générateurs, notés respectivement A^N et $\bar{A}$, ainsi que la formule suivante, que nous démontrons dans notre cadre

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds.$$

Puis nous montrons que P_t^N converge faiblement vers la mesure invariante de $\bar{X}$ quand N et t tendent conjointement vers l'infini. Pour montrer ce résultat, nous utilisons, la vitesse de convergence explicite de P^N vers $\bar{P}$.

Finalement, nous montrons la convergence des processus ponctuels $Z^{N,i}$ en proposant deux preuves. L'une est une application du Théorème $IX.4.15$ of Jacod and Shiryaev (2003) qui exploite le fait que $(X^N, Z^{N,i})$ est une semi-martingale. L'autre est une application de notre Théorème B.0.1 qui permet de montrer la convergence de processus ponctuels via la convergence de leurs intensités stochastiques.

1.7.2 Chapitre 4 : Processus de Hawkes à mémoire variable

Le modèle de ce chapitre, proche de celui du Chapitre 2, est plus adapté à la modélisation des réseaux de neurones. En effet, nous modélisons le phénomène biologique de la repolarisation du potentiel de membrane des neurones : lorsqu'un neurone envoie une décharge, son potentiel de membrane retourne rapidement vers sa valeur de repos, que nous supposerons égale à zéro.

Au lieu d'étudier un processus X^N de dimension un solution de (1.8), nous étudions un système

$(X^{N,i})_{1\le i\le N}$ solution de

$$
\begin{aligned}
X_t^{N,i} &= X_0^{N,i} - \alpha \int_0^t X_s^{N,i} ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} X_{s-}^{N,i} \mathbf{1}_{\{z\le f(X_{s-}^{N,i})\}} \pi^i(ds,dz,du) \\
&\quad + \frac{1}{\sqrt{N}} \sum_{j\ne i} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbf{1}_{\{z\le f(X_{s-}^{N,j})\}} \pi^j(ds,dz,du).
\end{aligned} \tag{1.9}
$$

Nous interprétons $X_t^{N,i}$ comme le potentiel de membrane du $i-$ième neurone d'un réseau de N neurones à l'instant t.

C'est le troisième terme du membre droit de l'équation ci-dessus qui correspond à la repolarisation. Les sauts de ce terme correspondent à des sauts de réinitialisation, car lorque le neurone i émet une décharge, ce terme de saut force son potentiel $X^{N,i}$ à sauter vers zéro immédiatement.

Dans ces équations, il y a un système de processus de Hawkes sous-jacent $(Z^{N,i})_{1\le i\le N}$ défini par

$$
Z_t^{N,i} = \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{s-}^{N,i})\right\}} d\pi^i(s,z,u).
$$

En effet, le système $(X^{N,i})_{1\le i\le N}$ possède la dynamique de type Hawkes suivante

$$
X_t^{N,i} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{]L_t^i,t]\times\mathbb{R}_+\times\mathbb{R}} e^{-\alpha(t-s)} u \mathbb{1}_{\left\{z\le f(X_{s-}^{N,j})\right\}} d\pi^j(s,z,u) + e^{-\alpha t} X_0^{N,i} \mathbf{1}_{L_t^i=0},
$$

où $L_t^i = \sup\{0 \le s \le t : \Delta Z_s^{N,i} = 1\}$ est le dernier instant de décharge du neurone i avant le temps t, avec la convention $\sup\emptyset := 0$. Autrement dit, l'intégrale sur le passé, qui part de zéro avec la dynamique classique des processus de Hawkes, est remplacée par une intégrale qui part du dernier instant de saut. On dit que ce sont des processus à mémoire variable.

Le but du chapitre est d'étudier la convergence en loi du système $(X^{N,i})_{1\le i\le N}$ quand N tend vers l'infini. Le système limite $(\bar{X}^i)_{i\ge 1}$ est de la forme suivante :

$$
\begin{aligned}
\bar{X}_t^i &= \bar{X}_0^i - \alpha \int_0^t \bar{X}_s^i ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-}^i \mathbf{1}_{\{z\le f(\bar{X}_{s-}^i)\}} \pi^i(ds,dz,du) \\
&\quad + \sigma \int_0^t \sqrt{\mathbb{E}\left[f\left(\bar{X}_s^i\right) \middle| \mathcal{W}\right]} dW_s,
\end{aligned} \tag{1.10}
$$

où W est encore un mouvement brownien standard de dimension un, et $\mathcal{W}$ la tribu engendrée par ce brownien.

Notons que ce mouvement brownien est commun à toutes les particules du système limite, et donc ce système n'est pas i.i.d., mais seulement conditionnellement i.i.d. étant donné W. C'est pour cela que la mesure directrice du système est la loi conditionnelle sachant ce mouvement brownien. Cette mesure directrice apparaît dans l'équation (1.10) sous forme d'espérance conditionnelle (cette espérance est la fonction f intégrée par rapport à la mesure directrice), car la variation quadratique du terme de saut de (1.9) s'écrit comme la fonction f intégrée par rapport à la mesure empirique du système

$$
\left\langle \frac{1}{\sqrt{N}} \sum_{j\ne i} \int_{[0,\cdot]\times\mathbb{R}_+\times\mathbb{R}} u \mathbf{1}_{\{z\le f(X_{s-}^{N,j})\}} \pi^j(ds,dz,du) \right\rangle_t = \frac{1}{N} \sum_{j\ne i} f(X_t^{N,j}) = \mu_t^N(f) - \frac{1}{N} f(X_t^{N,i}),
$$

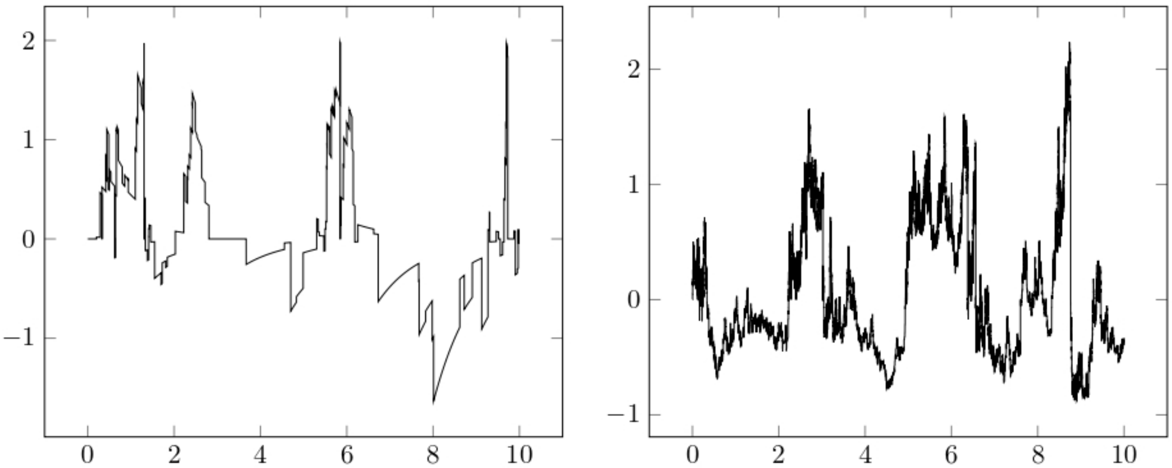

FIGURE 1.4 : Simulations de trajectoires de $(X_t^{N,1})_{0\leq t\leq 10}$ avec $X_0^{N,i} = 0$ $(1 \leq i \leq N)$, $\alpha = 1$, $\nu = \mathcal{N}(0,1)$, $f(x) = 2.2 + 1.4\arctan(10x - 2)$, $N = 10$ (image de gauche) et $N = 1000$ (image de droite).

avec $\mu_t^N := N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$. Voir la Section 4.1 pour plus détails sur l'intuition de la forme de l'équation différentielle stochastique (1.10).

Les équations ci-dessus n'étant pas classiques, nous montrons d'abord que ce système d'équations est bien posé, en introduisant une distance appropriée en nous inspirant des preuves de Fournier and Löcherbach (2016) et de Graham (1992). En effet, comme dans Fournier and Löcherbach (2016), le compensateur du terme de saut fait intervenir la fonction $x \mapsto xf(x)$ qui n'est pas lipschitzienne, et comme dans Graham (1992), nous travaillons à la fois avec un terme brownien et un terme de saut. Notons que des difficultés techniques apparaissent dans notre cadre par rapport à ces deux références à cause du fait que l'espérance de l'équation (1.10) est conditionnelle et que cette espérance est l'intégrande d'une intégrale brownienne.

Puis, pour montrer la convergence du système X^N vers le système $\bar{X}$, nous ne pouvons pas appliquer les techniques du chapitre précédent, à cause de difficultés techniques. C'est pour cela que nous démontrons la convergence de la mesure empirique de X^N vers la mesure directrice de $\bar{X}$. Nous pouvons remarquer que ces deux convergences sont équivalentes car nous manipulons des systèmes échangeables. Pour montrer la convergence en loi des mesures empiriques, nous introduisons un nouveau type de problème martingale dont la loi de la mesure directrice du système limite est la seule solution.

1.7.3 Chapitre 5 : Limite diffusive de systèmes McKean-Vlasov

Dans le chapitre 4, nous étudions l'équation différentielle stochastique suivante

$$d\bar{X}_t = -\alpha\bar{X}_t dt + \sigma\sqrt{\mu_t(f)}dW_t - \bar{X}_{t-}dZ_t,$$

avec $\mu_t(f) = \mathbb{E}\left[f(\bar{X}_t)\middle| W\right]$ et Z un processus ponctuel d'intensité $f(\bar{X}_{t-})$. Cette équation n'est pas classique. Elle peut être appelée "équation de McKean-Vlasov conditionnelle". Dans le cadre classique des équations de McKean-Vlasov, les coefficients dépendent à la fois de la solution de l'équation et aussi de la loi de cette solution. Ici, nous parlons d'équation de McKean-Vlasov conditionnelle parce que μ_t est la loi conditionnelle de la solution.

Le but du chapitre 5 est de généraliser les résultats de la partie précédente à ce cadre. Nous étudions des limites de grande échelle de systèmes de particules de type McKean-Vlasov en régime diffusif. Plus précisément, nous étudions le système suivant.

$$\begin{aligned} dX_t^{N,i} &= b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)d\beta_t^i \\ &+ \frac{1}{\sqrt{N}} \sum_{k=1,k\neq i}^{N} \int_{\mathbb{R}_+\times E} \Psi(X_{t-}^{N,k}, X_{t-}^{N,i}, \mu_{t-}^N, u^k, u^i) \mathbb{1}_{\left\{z\leq f(X_{t-}^{N,k}, \mu_{t-}^N)\right\}} d\pi^k(t,z,u), 1\leq i\leq N, \end{aligned} \quad (1.11)$$

où $\mu_t^N = N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$, $E = \mathbb{R}^{\mathbb{N}^*}$, avec β^i $(i\geq 1)$ des mouvements browniens i.i.d. standards de dimension un et π^k $(k\geq 1)$ des mesures de Poisson i.i.d. d'intensité $ds\cdot dz\cdot \nu(du)$, avec ν une mesure produit sur $E = \mathbb{R}^{\mathbb{N}^*}$. Pour manipuler ce type d'équation, nous supposons que les coefficients sont lipschitziens où l'espace des variables d'espace est $\mathbb{R}$ muni de sa topologie habituelle et l'espace des variables mesures est muni de la distance de Wasserstein d'ordre un.

Remarque 1.7.1. *Dans ce modèle, nous ne considérons plus le terme de saut de réinitialisation du Chapitre 4 pour simplifier le modèle.*

Remarquons que le système est caractérisé par des équations McKean-Vlasov conditionnelles (ce conditionnement apparaît à cause de la normalisation diffusive).

Pour prouver la convergence des systèmes ci-dessus, nous utilisons des techniques similaires à celles du Chapitre 4 : l'argument principal de la preuve est un problème martingale. Cependant il y a plusieurs difficultés supplémentaires dans ce modèle comparé au précédent. La différence principale est la forme du système limite où des bruits blancs apparaissent à la place du mouvement brownien. La raison pour laquelle les mouvements browniens ne sont pas suffisants dans ce modèle c'est la dépendance spatiale de la fonction d'amplitude des sauts Ψ (i.e. la dépendance en $X^{N,k}$ et $X^{N,i}$ dans 1.11). En effet, si la fonction Ψ ne dépendait que des deux dernières variables u^k et u^i, le système d'équations limites serait encore dirigé par des mouvement browniens, comme c'est expliqué dans l'Exemple 5.0.3 du Chapitre 5.

Il y a une autre difficulté technique dans ce modèle comparé à celui du Chapitre 4 : la convergence faible d'une suite de mesures $(\mu_t^N)_N$ implique la convergence de $(\mu_t^N(f))_N$ (pour n'importe quelle fonction f continue bornée), mais elle n'implique pas en général la convergence plus forte de $(\mu_t^N)_N$ au sens de la distance de Wasserstein. Ce dernier point est problématique car la régularité des coefficients de nos équations est exprimée par rapport à la distance de Wasserstein et non par rapport à la distance de Prokhorov.

1.7.4 Chapitre 6 : Existence, unicité et limite linéaire de sytèmes McKean-Vlasov avec des coefficients localement lipschitziens

Le but de ce chapitre est d'étudier des équations et des systèmes de type McKean-Vlasov. Contrairement au Chapitre 5, nous travaillons en régime linéaire : la force des interaction d'un système à

N particules est de l'ordre de N^{-1}. L'intérêt des résultats de ce chapitre est de travailler avec des coefficients non-lipschitziens.

Nous avons déjà rencontré ce type de coefficients dans l'équation (1.9). En effet, le compensateur du terme de saut fait intervenir la fonction $x \mapsto xf(x)$, qui n'est a priori pas lipschitzienne (à notre connaissance, il n'existe pas d'hypothèse "naturelle" sur la fonction f qui rende cette fonction lipschitzienne). Dans le Chapitre 4, nous contournons cette difficulté en introduisant une distance appropriée et une hypothèse particulière. Dans le Chapitre 6, nous travaillons avec une hypothèse générique sur la régularité des coefficients des équations différentielles stochastiques : nous supposons que ces coefficients sont localement lipschitziens et nous nous donnons une condition sur la croissance de la constante de Lipschitz locale par rapport aux variables des fonctions concernées. Nous ne montrons ces résultats qu'en normalisation linéaire et que pour des "vraies" équations de McKean-Vlasov (i.e. non conditionnelle), car il y a plusieurs difficultés techniques supplémentaires liées aux mesures aléatoires.

Les questions d'existence et d'unicité des solutions d'équations McKean-Vlasov et la question de propagation du chaos de systèmes McKean-Vlasov sont classiques dans la littérature. Dans le cas où l'on s'intéresse à des équations sans terme de saut (i.e. uniquement avec un terme de dérive et un terme brownien) il existe beaucoup de résultats avec peu d'hypothèses sur la régularité des coefficients : Gärtner (1988) et Lacker (2018) sur les questions d'existence, d'unicité et de propagation du chaos, et Mishura and Veretennikov (2020) et Chaudru de Raynal (2020) sur les questions d'existence et d'unicité sous différentes hypothèses. Remarquons que dans le cadre où les coefficients des équations différentielles stochastiques ne sont pas lipschitziens, l'existence et l'unicité des solutions ne sont montrées que dans un sens faible.

Dans le Chapitre 6, nous étudions des équations McKean-Vlasov avec un terme de saut. Les questions mentionnées avant ont aussi été étudiées dans ce cadre par Graham (1992) et Andreis et al. (2018) dans le cas où les coefficients sont lipschitziens. La nouveauté de nos résultats repose sur le fait que nous travaillons avec des coefficients localement lipschitziens.

Nous montrons d'abord que, dans ce cadre, l'équation de McKean-Vlasov suivante est bien posée au sens fort :

$$dX_t = b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \int_{\mathbb{R}_+ \times E} \Phi(X_{t-}, \mu_{t-}, u) \mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}} d\pi(t, z, u),$$

où μ_t est la loi de X_t, W un mouvement brownien, π une mesure de Poisson sur E un espace mesurable (voir la Section 6.1 pour plus de détails sur les notations). Pour montrer ce résultat avec des coefficients localement lipschitziens, nous adaptons les preuves du cas où les coefficients sont globalement lipschitziens. Nous utilisons un argument de troncature pour gérer la dépendance des constantes de Lipschitz locales par rapport aux variables. Contrairement au cas globalement Lipschitz, le lemme de Grönwall ne permet pas de conclure immédiatement. Dans le cas localement Lipschitz, il faut utiliser une généralisation de ce lemme : le lemme d'Osgood. L'unicité des solutions vient rapidement à partir de ce lemme, mais il y a d'autres difficultés pour montrer l'existence des solutions. Nous contruisons une solution faible en utilisant un schéma d'itération de Picard, mais le fait que les coefficients sont seulement localement lipschitziens ne nous permet pas de montrer que ce schéma converge dans un sens L^1. À la place, nous prouvons qu'il existe une sous-suite qui converge en loi vers une certaine limite qui est solution de l'équation. Certaines difficultés techniques apparaissent pour deux raisons. La première est que le schéma de Picard ne converge pas directement, mais admet seulement une sous-suite convergeante. Ce qui implique qu'il faut contrôler la variation entre deux étapes successives du schéma. La deuxième est que nous ne montrons qu'une convergence en loi, ce n'est donc pas évident que la limite est bien solution de l'équation. Nous le montrons en exploitant ses caractéristiques de semi-martingale.

Le deuxième résultat de ce chapitre est la propagation du chaos des sytèmes McKean-Vlasov dans ce cadre localement Lipschitz. Plus précisément, il s'agit de la convergence de ce système de particules :

$$\begin{aligned} dX_t^{N,i} =& b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+\times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i)\mathbb{1}_{\left\{z\leq f(X_{t-}^{N,i}, \mu_{t-}^N)\right\}} d\pi^i(t,z,v) \\ &+ \frac{1}{N}\sum_{j=1}^N \int_{\mathbb{R}_+\times F^{\mathbb{N}^*}} \Theta(X_{t-}^{N,j}, X_{t-}^{N,i}, \mu_{t-}^N, v^j, v^i)\mathbb{1}_{\left\{z\leq f(X_{t-}^{N,j}, \mu_{t-}^N)\right\}} d\pi^j(t,z,v), \end{aligned}$$

avec $\mu_t^N := N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$, vers le système infini

$$\begin{aligned} d\bar{X}_t^i =& b(\bar{X}_t^i, \bar{\mu}_t)dt + \sigma(\bar{X}_t^i, \bar{\mu}_t)dW_t^i + \int_{\mathbb{R}_+\times F^{\mathbb{N}^*}} \Psi(\bar{X}_{t-}^i, \bar{\mu}_{t-}, v^i)\mathbb{1}_{\left\{z\leq f(\bar{X}_{t-}^i, \bar{\mu}_{t-})\right\}} d\pi^i(t,z,v) \\ &+ \int_{\mathbb{R}}\int_{F^{\mathbb{N}^*}} \Theta(x, \bar{X}_t^i, \bar{\mu}_t, v^1, v^2) f(x,\bar{\mu}_t) d\nu(v) d\bar{\mu}_t(x), \end{aligned}$$

où $\bar{\mu}_t := \mathcal{L}(\bar{X}_t)$, quand N tend vers l'infini (voir la Section 6.2 pour plus de détails sur les notations). La preuve de ce résultat repose sur des arguments similaires à ceux utilisés pour montrer l'unicité des solutions de l'équations McKean-Vlasov : un argument de troncature et le lemme d'Osgood.

1.8 Notations

Dans la thèse, nous utilisons les notations suivantes :

- Si E est un espace métrique, nous notons $\mathcal{P}(E)$ l'espace des lois sur E muni de la topologie de la convergence faible (i.e. muni de la distance de Prokhorov).

- Nous notons $\mathcal{P}_p(\mathbb{R})$ l'espace des lois sur $\mathbb{R}$ avec un moment d'ordre p fini. Sauf mention contraire, cet espace est muni de la distance de Wasserstein d'ordre p, notée W_p et définie par : quelque soit $m_1, m_2 \in \mathcal{P}_p(\mathbb{R})$,

  $$W_p(m_1,m_2) = \inf_{X_1\sim m_1, X_2\sim m_2} \mathbb{E}\left[|X_1 - X_2|^p\right]^{1/p}.$$

  Le théorème 6.9 et la définition 6.8 de Villani (2008) donnent plusieurs caractérisations de la convergence de cette distance. Remarquons qu'il existe une autre caractérisation de la distance W_1, appelée "dualité de Kantorovich-Rubinstein" (voir Remarque 6.5 de Villani (2008)) :

  $$W_1(m_1,m_2) = \sup_{f\in Lip_1} \int_{\mathbb{R}} f(x)dm_1(x) - \int_{\mathbb{R}} f(x)dm_2(x),$$

  avec Lip_1 l'espace des fonctions lipschitziennes réelles dont la constante de Lipschitz est inférieure à un.

- Si X est une variable aléatoire, nous notons $\mathcal{L}(X)$ sa loi.

- Si X et X_n ($n \in \mathbb{N}$) sont des variables aléatoires, nous écrivons $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ pour dire "$(X_n)_n$ converge en loi vers X".

- Si $g : \mathbb{R}^p \to \mathbb{R}$ est une fonction n fois dérivable, nous notons $||g||_{n,\infty} = \sum_{|\alpha|\leq n} ||\partial_\alpha g||_\infty$.

- Si $g : \mathbb{R} \to \mathbb{R}$ est une fonction mesurable et λ une mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que g est intégrable par rapport à λ, nous notons $\lambda(g)$ pour $\int_\mathbb{R} g d\lambda$.

- Nous notons $C_b^n(\mathbb{R}^p)$ l'ensemble des fonctions $g : \mathbb{R}^p \to \mathbb{R}$ de classe C^n telles que $||g||_{n,\infty} < +\infty$, et nous écrivons $C_b(\mathbb{R}^p)$ au lieu de $C_b^0(\mathbb{R}^p)$. Finalement, $C^n(\mathbb{R}^p)$ est l'ensemble des fonctions de classes C^n non-nécessairement bornées.

- Si g est une fonction réelle et I est un intervalle, nous notons $||g||_{\infty,I} = \sup_{x\in I}|g(x)|$.

- Nous notons $C_c^n(\mathbb{R}^p)$ le sous-espace de $C^n(\mathbb{R}^p)$ restreint aux fonctions à support compact.

- Dans la suite, $||\cdot||_p$ désigne la norme p sur $\mathbb{R}^d$ ($d \in \mathbb{N}^*$).

- Soient $T > 0$ et (G, d) un espace polonais, alors $D([0, T], G)$ (resp. $D(\mathbb{R}_+, G)$) désigne l'espace des fonctions càdlàg à valeurs dans G définies sur $[0, T]$ (resp. $\mathbb{R}_+$) muni de la topologie de Skorokhod. C'est un espace polonais. Rappelons que la convergence d'une suite $(x_n)_n$ de $D([0, T], G)$ vers une certaine fonction x pour cette topologie est équivalente à l'éxistence de fonctions continues et strictement croissantes λ_n ($n \in \mathbb{N}$) telles que $\lambda_n(0) = 0$, $\lambda_n(T) = T$ et chacune des suites
$$\sup_{0\leq t\leq T} |\lambda_n(t) - t| \text{ et } \sup_{0\leq t\leq T} d(x(\lambda_n(t)), x_n(t))$$
tendent vers zéro quand n tend vers l'infini. Dans la suite, nous appellons une telle suite $(\lambda_n)_n$ une suite de changement de temps (en anglais, "a sequence of time-changes").

- Nous notons $\mathcal{M}(E)$ l'espace des mesures localement finies sur l'espace mesurable E, muni de la topologie de la convergence vague. Et $\mathcal{N}(E)$ est le sous-espace contenant les mesures ponctuelles simples. Dans la suite, nous omettons le E dans les notations. Nous nous reportons à la section $A2.6$ de Daley and Vere-Jones (2003) pour une étude de cet espace.

- Nous notons C les constantes arbitraires qui apparaissent dans les calculs. Ainsi la valeur de C peut changer d'une ligne à l'autre dans une même équation. De plus, si la constante C dépend d'un autre paramètre θ (qui n'est pas un paramètre du modèle), nous écrivons C_θ.

1.9 Notation (in English)

- If E is a metric space, we note $\mathcal{P}(E)$ the space of distributions on E endowed with the topology of the weak convergence (i.e. endowed with Prohorov metric).

- We note $\mathcal{P}_p(\mathbb{R})$ the space of distributions on $\mathbb{R}$ with a finite $p-$order moment. This space is always endowed with Wasserstein metric of order p, denoted by W_p and defined as follows: for all $m_1, m_2 \in \mathcal{P}_p(\mathbb{R})$,
$$W_p(m_1, m_2) = \inf_{X_1\sim m_1, X_2\sim m_2} \mathbb{E}\left[|X_1 - X_2|^p\right]^{1/p}.$$

Theorem 6.9 and Definition 6.8 of Villani (2008) give some characterizations of the convergence for this metric. Note that there exists another characterization of the metric W_1, called "Kantorovich-Rubinstein's duality" (see Remark 6.5 of Villani (2008)):

$$W_1(m_1, m_2) = \sup_{f \in Lip_1} \int_{\mathbb{R}} f(x) dm_1(x) - \int_{\mathbb{R}} f(x) dm_2(x),$$

where Lip_1 is the space of real valued Lipschitz functions with Lipschitz constant nongreater than one.

- If X is a random variable (r.v.), we note $\mathcal{L}(X)$ its distribution.
- If X and X_n ($n \in \mathbb{N}$) are r.v., we write $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$ for "$(X_n)_n$ converges in distribution to X".
- If $g : \mathbb{R}^p \to \mathbb{R}$ is a function that is n times differentiable, we note $||g||_{n,\infty} = \sum_{|\alpha|\leq n} ||\partial_\alpha g||_\infty$.
- If $g : \mathbb{R} \to \mathbb{R}$ is a measurable function and λ a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that g is integrable w.r.t. λ, we note $\lambda(g)$ for $\int_{\mathbb{R}} g d\lambda$.
- We note $C^n_b(\mathbb{R}^p)$ the set of functions $g : \mathbb{R}^p \to \mathbb{R}$ that are n times continuously differentiable such that $||g||_{n,\infty} < +\infty$, and we write $C_b(\mathbb{R}^p)$ instead of $C^0_b(\mathbb{R}^p)$. Finally, $C^n(\mathbb{R}^p)$ is the set of functions that are n times continuously differentiable.
- If g is a real function and I an interval, we note $||g||_{\infty,I} = \sup_{x\in I}|g(x)|$.
- We note $C^n_c(\mathbb{R}^p)$ the subset of $C^n(\mathbb{R}^p)$ restricted to the functions that are compactly supported.
- In the following, $|| \cdot ||_p$ denotes the $p-$norm on $\mathbb{R}^d$ ($d \in \mathbb{N}^*$).
- Let $T > 0$ and (G, d) a Polish space, then $D([0, T], G)$ (resp. $D(\mathbb{R}_+, G)$) denotes the space of càdlàg $G-$valued functions defined on $[0, T]$ (resp. $\mathbb{R}_+$) endowed with Skorohod topology. This is a Polish space. Recall that the convergence of a sequence $(x_n)_n$ of $D([0, T], G)$ to some function x for this topology is equivalent to the existence of continuous and increasing functions λ_n ($n \in \mathbb{N}$) such that $\lambda_n(0) = 0$, $\lambda_n(T) = T$ and both sequences

$$\sup_{0\leq t\leq T} |\lambda_n(t) - t| \text{ and } \sup_{0\leq t\leq T} d(x(\lambda_n(t)), x_n(t))$$

  vanish as n goes to infinity. In the following, we call such a sequence $(\lambda_n)_n$ a sequence of time-changes.
- We note $\mathcal{M}(E)$ the space of locally finite measures on the measurable space E endowed with the topology of the vague convergence. We note $\mathcal{N}(E)$ the subspace that contains only the simple point measures. In the following, we omit the E in te notation. We refer to section $A2.6$ of Daley and Vere-Jones (2003) for a survey on this space.
- We note C any arbitrary constant in the computations. So the value of C can change in an equation. In addition, if the constant C depends on a parameter θ (which is not a model parameter), we write C_θ.

Part I

Propagation of chaos for mean field models of Hawkes processes in a diffusive regime

Introduction of the part

In this part, we prove the convergence of sequences of systems of Hawkes processes in different models, in a mean field framework. Namely, for each $N \in \mathbb{N}^*$, we consider a system of point processes $\left(Z^{N,1}, \dots, Z^{N,N}\right)$ such that each $Z^{N,i}$ has intensity of the form

$$\lambda_t^N = f\left(\sum_{j=1}^N \int_{]-\infty,t[} h^N(t-s)dZ_s^{N,j}\right). \tag{1.12}$$

In order to make the expression in (1.12) converge, the most natural way would be to scale the sum in N^{-1}, to reduce the problem to a kind of law of large numbers. This point of view has e.g. been adopted by Delattre et al. (2016). Here, instead of reducing the problem to a law of large numbers, we reduce it to a central limit theorem. This means that we scale the sum of (1.12) in $N^{-1/2}$, and that we center the terms of the sum. To this end, we consider intensities with stochastic jump heights of the form

$$\lambda_t^N = f\left(\frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[} h(t-s)U_j(s)dZ_s^{N,j}\right), \tag{1.13}$$

where $(U_j(s))_{j\in\mathbb{N}^*, s\geq 0}$ is an i.i.d. family of centered random variables.

The problem with the formulation of (1.13) is that the processes $(U_j(s))_{s\geq 0}$ cannot be predictable, and so the integrand cannot be manipulated. So, we formulate differently the stochastic jump heights, replacing the variables $U_j(s)$ by noise parameters of Poisson measures. In the following, we use systems of Hawkes processes in the following form:

$$\begin{cases} Z_t^{N,i} = & \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\{z\leq\lambda_s^N\}} d\pi^i(s,z,u) \quad (1\leq i\leq N), \\ \lambda_t^N = & f\left(\frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t[\times\mathbb{R}_+\times\mathbb{R}} u\cdot h(t-s)\mathbb{1}_{\{z\leq\lambda_s^N\}} d\pi^j(s,z,u)\right), \end{cases} \tag{1.14}$$

where $\left(\pi^i\right)_{i\in\mathbb{N}^*}$ is an i.i.d. family of Poisson measures on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}$ having intensity $dt\cdot dz\cdot d\mu(u)$, and μ is a centred probability measure on $\mathbb{R}$.

The convergence of the point processes relies mainly on the convergence of their stochastic intensity. In our case, this intensity is of the form $\lambda_t^N = f\left(X_{t-}^N\right)$, with

$$X_t^N := \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\cdot h(t-s)\mathbb{1}_{\left\{z\leq f\left(X_{s-}^N\right)\right\}} d\pi^j(s,z,u).$$

In the following, the majority of the results concern the process $\left(X_t^N\right)_{t\geq 0}$. In every model that we study, these results will be based on the fact that X^N is solution of a stochastic differential equation driven by the Poisson measures $\left(\pi^i\right)_{1\leq i\leq N}$.

In particular, we prove the convergence of X^N in distribution in Skorohod space. Contrarily to the situation considered in Delattre et al. (2016), the limit of X^N is stochastic.

In Chapter 2, for each N the process X^N is a $1-$dimensional process. In Chapter 3, this process X^N is multi-dimensional, but the dimension does not depend on N. For these two chapters, the proofs rely on analytical technics based on the convergence of the infinitesimal generators of the processes, and a convergence speed is given. In Chapter 4, X^N is a $N-$dimensional process. Its convergence is proven using a new kind of martingale problem.

Chapter 2

Exponential kernel

This chapter is based on Erny et al. (2019).

We start studying the model given by (1.14) with exponential kernel, that is $h(t) := e^{-\alpha t}$ for some $\alpha > 0$. The explicit model is given by the following system of Hawkes processes :

$$\begin{cases} Z_t^{N,i} = & \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\leq f\left(X_{s-}^N\right)\right\}} d\pi^i(s,z,u) \quad (1\leq i\leq N), \\ X_t^N = & X_0^N e^{-\alpha t} + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{[0,t[\times\mathbb{R}_+\times\mathbb{R}} u\cdot e^{-\alpha(t-s)} \mathbb{1}_{\left\{z\leq f\left(X_{s-}^N\right)\right\}} d\pi^j(s,z,u), \\ X_0^N \sim & \nu_0^N, \end{cases} \tag{2.1}$$

where ν_0^N is a probability measure on $\mathbb{R}$ and the π^i $(i\in\mathbb{N}^*)$ are i.i.d. Poisson measures on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}$ having intensity $dt\cdot dz\cdot d\mu(u)$, and μ is a centred probability measure on $\mathbb{R}$.

We point out that the form of the function $h(t) = e^{-\alpha t}$ is extremely important, since it guarantees X^N to be a Markov process. Actually, we can even prove that any strong solution of (2.1) is solution of (2.2).

Lemma 2.0.1. *X^N satisfies*

$$dX_t^N = -\alpha X_t^N dt + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{(z,u)\in\mathbb{R}_+\times\mathbb{R}} u \mathbb{1}_{\left\{z\leq f\left(X_{t-}^N\right)\right\}} d\pi^j(s,z,u). \tag{2.2}$$

Proof. Applying the formula of integration by parts to

$$X_t^N = e^{-\alpha t}\times\left(X_0^N + \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\cdot e^{\alpha s}\mathbb{1}_{\left\{z\leq f\left(X_{s-}^N\right)\right\}} d\pi^j(s,z,u)\right),$$

we obtain exactly the result. □

It is possible to prove the strong well-posedness of (2.1) using a reasoning similar as in the proof of Theorem 6 in Delattre et al. (2016). For the sake of simplicity, in this chapter, we consider that the process X^N is directly defined by (2.2).

The aim of this chapter is to study, under appropriate conditions on f, the convergence in distribution in Skorohod space of X^N to a diffusion $\bar{X}$ that is solution of the stochastic differential equation:

$$\begin{cases} d\bar{X}_t = & -\alpha\bar{X}_t dt + \sigma\sqrt{f\left(\bar{X}_t\right)}dB_t, \\ \bar{X}_0 \sim & \bar{\nu}_0 \end{cases} \tag{2.3}$$

where σ^2 is the variance of μ, $(B_t)_{t\geq 0}$ is a one-dimensional standard Brownian motion, and $\bar{\nu}_0$ is a probability measure on $\mathbb{R}$. This result is formally stated at Theorem 2.2.1.

Apart from proving the aforementioned convergence, we prove a rate of convergence at Theorem 2.2.3, and deduce from it the convergence of the Markovian kernel of X^N to the invariant measure of $\bar{X}$ (see Theorem 2.3.1). Besides, we also prove in Theorem 2.4.1 the convergence of the system of Hawkes processes $\left(Z^{N,1},\dots,Z^{N,N}\right)$ in distribution in $D\left(\mathbb{R}_+,\mathbb{R}\right)^{\mathbb{N}^*}$ endowed with the product topology.

In Section 2.1, we formally state the assumptions we use throughout the chapter. Section 2.2 is devoted to prove the convergence of X^N in Skorohod space, as well as the rate of convergence of its semigroup and of its Markovian kernel. In Section 2.3, we prove the convergence of the Markovian kernel of X^N to the invariant measure of $\bar{X}$. Finally we prove the convergence of the systems of Hawkes processes, that is the convergence of the point processes in Section 2.4.

2.1 Assumptions on the model

Firstly, we need an assumption that guarantees the process $\bar{X}$ to exist.

Assumption 2.1. *$\sqrt{f}$ is a positive and Lipschitz continuous function, having Lipschitz constant L.*

Under Assumption 2.1, it is classical that the equation (2.3) admits a unique non-exploding strong solution (see remark IV.2.1, Theorems IV.2.3, IV.2.4 and IV.3.1 of Ikeda and Watanabe (1989)).

Assumption 2.2.

- $\int_{\mathbb{R}} x^4 d\bar{\nu}_0(x) < \infty$ *and* $\sup\limits_{N\in\mathbb{N}^*}\int_{\mathbb{R}} x^4 d\nu_0^N(x) < \infty$.
- *μ is a centered probability measure having a fourth moment, we note σ^2 its variance.*

Assumption 2.2 allows us to control the moments up to order four of the processes $\left(X_t^N\right)_t$ and $\left(\bar{X}_t\right)_t$ (see Lemma 2.2.5) and to prove the convergence of the generators of the processes $\left(X_t^N\right)_t$ (see Proposition 2.2.6).

Assumption 2.3. *We assume that f is C^4 and for each $1\leq k\leq 4$, $\left(\sqrt{f}\right)^{(k)}$ is bounded by some constant m_k.*

Assumption 2.3 guarantees that the stochastic flow associated to (2.3) has regularity properties with respect to the initial condition $\bar{X}_0 = x$. This will be the main tool to obtain uniform in time estimates of the limit semigroup, see Proposition 2.2.7.

Example 2.1.1. *The functions $f(x) = 1+x^2$, $f(x)=\sqrt{1+x^2}$ and $f(x) = (\pi/2+\arctan x)^2$ satisfy Assumptions 2.1 and 2.3.*

Assumption 2.4. *X_0^N converges in distribution to $\bar{X}_0$.*

Obviously, Assumption 2.4 is a necessary condition for the convergence in distribution of X^N to $\bar{X}$.

Let us prove that our model is well-defined.

Proposition 2.1.2. *If Assumptions 2.1 and 2.2 hold, the equation (2.2) admits a unique non-exploding strong solution.*

Proof. It is well known that if f is bounded, there is a unique strong solution of (2.2) (see Theorem IV.9.1 of Ikeda and Watanabe (1989)). In the general case we reason in a similar way as in the proof of Proposition 2 in Fournier and Löcherbach (2016). Consider the solution $(X_t^{N,K})_{t\in\mathbb{R}_+}$ of the equation (2.2) where f is replaced by $f_K : x \in \mathbb{R} \mapsto f(x) \wedge \sup_{|y|\leq K} f(y)$ for some $K \in \mathbb{N}^*$. Introduce moreover the stopping time

$$\tau_K^N = \inf\left\{t \geq 0 \ : \ \left|X_t^{N,K}\right| \geq K\right\}.$$

Since for all $t \in \left[0, \tau_K^N \wedge \tau_{K+1}^N\right]$, $X_t^{N,K} = X_t^{N,K+1}$, we know that $\tau_K^N(\omega) \leq \tau_{K+1}^N(\omega)$ for all ω. Then we can define τ^N as the non-decreasing limit of τ_K^N. With a classical reasoning relying on Ito's formula and Grönwall's lemma (see the proof of Lemma 2.2.5.(iv)), we can prove that

$$\sup_{0\leq s\leq t} \mathbb{E}\left[\left(X_{s\wedge\tau_K^N}^{N,K}\right)^2\right] \leq C_t\left(1 + x^2\right), \tag{2.4}$$

where $C_t > 0$ does not depend on K. As a consequence, we know that almost surely, $\tau^N = +\infty$. Indeed,

$$\mathbb{P}\left(\tau_K^N \leq t\right) \leq \mathbb{P}\left(|X_{t\wedge\tau_K^N}^{N,k}| \geq K\right) \leq K^{-2}C_t(1+x^2) \underset{N\to\infty}{\longrightarrow} 0,$$

then

$$\mathbb{P}\left(\tau^N \leq t\right) \leq \mathbb{P}\left(\bigcap_{K\in\mathbb{N}^*} \{\tau_K^N \leq t\}\right) = \lim_{K\to\infty} \mathbb{P}\left(\tau_K^N \leq t\right) = 0.$$

Consequently

$$\mathbb{P}\left(\tau^N < \infty\right) \leq \mathbb{P}\left(\bigcup_{t\in\mathbb{N}^*} \{\tau^N \leq t\}\right) = 0.$$

So we can simply define X_t^N as the limit of $X_t^{N,K}$, as K goes to infinity. Now we show that X^N satisfies equation (2.2). Consider some $\omega \in \Omega$ and $t > 0$, and choose K such that $\tau_K^N(\omega) > t$. Then we know that for all $s \in [0,t]$, $X_s^N(\omega) = X_s^{N,K}(\omega)$ and $f(X_{s-}^N(\omega)) = f_K(X_{s-}^{N,K}(\omega))$. Moreover, as $X^{N,K}(\omega)$ satisfies the equation (2.2) with f replaced by f_K, we know that $X^N(\omega)$ verifies the equation (2.2) on $[0,t]$. This holds for all $t > 0$. As a consequence, we know that X^N satisfies the equation (2.2). This proves the existence of strong solution. The uniqueness is a consequence of the uniqueness of strong solutions of (2.2), if we replace f by f_K in (2.2), and of the fact that any strong solution $(Y_t^N)_t$ equals necessarily $(X_t^{N,K})_t$ on $[0, \tau_K^N]$. □

2.2 Convergence of $\left(X^N\right)_N$ in distribution in Skorohod topology

The goal of this section is to prove two results: the convergence of X^N to $\bar{X}$ and the rate of this convergence.

2.2.1 Convergence of X^N

Let us begin with the most classical proof.

Theorem 2.2.1. *Under Assumptions 2.1, 2.2 and 2.4, X^N converges to $\bar{X}$ in distribution in $D(\mathbb{R}_+, \mathbb{R})$.*

Proof. This proof relies on the fact that X^N ($N \in \mathbb{N}^*$) and $\bar{X}$ are semimartingales. We use a semimartingale convergence theorem stating that, the convergence of the characteristics of semimartingales implies the convergence of the semimartingales: Theorem $IX.4.15$ of Jacod and Shiryaev (2003).

Using the notation therein, with the truncation function $h = Id$, we have $K^N(x, dy) := Nf(x)\mu(\sqrt{N}dy)$, $b'^N(x) = -\alpha x$, and $\tilde{c}'^N(x) = \int K^N(x, dy)y^2 = \sigma^2 f(x)$. One can note that b'^N and $\tilde{c}'^N$ do not depend on N, and that, for any function $g \in C_b(\mathbb{R})$ null around zero, $\int_{\mathbb{R}} g(y)K^N(\cdot, dy)$ vanishes as N goes to infinity.

Then, Theorem $IX.4.15$ of Jacod and Shiryaev (2003) implies the result. □

2.2.2 Rate of convergence of X^N

In this section we provide a rate of convergence for Theorem 2.2.1. More precisely, we prove the convergence speed of the semigroups of the processes. This convergence speed is expressed in terms of the following parameters

$$\beta := \max\left(\frac{1}{2}\sigma^2 L^2 - \alpha, 2\sigma^2 L^2 - 2\alpha, \frac{7}{2}\sigma^2 L^2 - 3\alpha\right) \tag{2.5}$$

and, for any $T > 0$ and any fixed $\varepsilon > 0$,

$$K_T := (1 + 1/\varepsilon)\int_0^T (1 + s^2)e^{\beta s}\left(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)(T-s)}\right) ds. \tag{2.6}$$

Remark 2.2.2. *If $\alpha > 7/6\,\sigma^2 L^2$, then $\beta < 0$, and one can choose $\varepsilon > 0$ such that $\sigma^2 L^2 - 2\alpha + \varepsilon < 0$, implying that $\sup_{T>0} K_T < \infty$.*

Theorem 2.2.3. *Under Assumptions 2.1, 2.2 and 2.3, for all $T \geq 0$, for each $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,*

$$\sup_{0 \leq t \leq T} \left|P_t^N g(x) - \bar{P}_t g(x)\right| \leq C(1 + x^2)K_T ||g||_{3,\infty} \frac{1}{\sqrt{N}}.$$

In particular, if $\alpha > \frac{7}{6}\sigma^2 L^2$, then

$$\sup_{t \geq 0} \left|P_t^N g(x) - \bar{P}_t g(x)\right| \leq C(1 + x^2)||g||_{3,\infty} \frac{1}{\sqrt{N}}.$$

We prove Theorem 2.2.3 studying the convergence of the infinitesimal generator of X^N to that of $\bar{X}$.

Throughout this chapter, we consider extended generators similar to those used in Meyn and Tweedie (1993) and in Davis (1993), because the classical notion of generator does not suit to our framework. We introduce and study such a kind of generator at Appendix A.

Indeed, in the general theory of semigroups, one defines the generators on some Banach space. In the frame of semigroups related to Markov processes, one generally considers $(C_b(\mathbb{R}), ||\bullet||_\infty)$. In this context, the generator A of a semigroup $(P_t)_t$ is defined on the set of functions $\mathcal{D}(A) = \left\{ g \in C_b(\mathbb{R}) : \exists h \in C_b(\mathbb{R}), \left|\left| \frac{1}{t}(P_t g - g) - h \right|\right|_\infty \underset{t \to 0}{\longrightarrow} 0 \right\}$. Then one denotes the previous function h as Ag. If A is the generator of a diffusion, we can only guarantee that $\mathcal{D}(A)$ contains the functions that have a compact support, but to prove Proposition A.0.3, we need to apply the generators of the processes $\left(X_t^N\right)_t$ and $\left(\bar{X}_t\right)_t$ to functions of the type $\bar{P}_s g$, and we cannot guarantee that $\bar{P}_s g$ has compact support even if we assume g to be in $C_c^\infty(\mathbb{R})$.

As this definition slightly differs from one reference to another, we define explicitly the extended generator in Definition A.0.1 and we prove the results on extended generators that we need in Appendix A. We note A^N the extended generator of X^N and $\bar{A}$ that of $\bar{X}$. Before proving the convergence of these generators, we state a lemma which characterizes the generators for some test functions. This lemma is a straightforward consequence of Ito's formula and Lemma 2.2.5.

Lemma 2.2.4. *$C_b^2(\mathbb{R}) \subseteq \mathcal{D}'\left(\bar{A}\right)$, and for all $g \in C_b^2(\mathbb{R})$ and $x \in \mathbb{R}$, we have*

$$\bar{A}g(x) = -\alpha x g'(x) + \frac{1}{2}\sigma^2 f(x) g''(x).$$

Moreover, $C_b^1(\mathbb{R}) \subseteq \mathcal{D}'\left(A^N\right)$, and for all $g \in C_b^1(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$A^N g(x) = -\alpha x g'(x) + N f(x) \int_{\mathbb{R}} \left[g\left(x + \frac{u}{\sqrt{N}}\right) - g(x) \right] d\mu(u).$$

Let us prove useful a priori bounds on the moments of X^N and $\bar{X}$.

Lemma 2.2.5. *Under Assumptions 2.1 and 2.2, the following holds.*

(i) For all $\varepsilon > 0,\ t > 0$ and $x \in \mathbb{R}$, $\mathbb{E}_x\left[(X_t^N)^2\right] \leq C(1 + 1/\varepsilon)(1 + x^2)(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t})$, for some $C > 0$ independent of N, t, x and ε.

(ii) For all $\varepsilon > 0,\ t > 0$ and $x \in \mathbb{R}$, $\mathbb{E}_x\left[(\bar{X}_t)^2\right] \leq C(1 + 1/\varepsilon)(1 + x^2)(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t})$, for some $C > 0$ independent of t, x and ε.

(iii) For all $N \in \mathbb{N}^*, T > 0$, $\mathbb{E}\left[(\sup_{0 \leq t \leq T} |X_t^N|)^2\right] < +\infty$ and $\mathbb{E}\left[(\sup_{0 \leq t \leq T} |\bar{X}_t|)^2\right] < +\infty$.

(iv) For all $T > 0, N \in \mathbb{N}^*$, $\sup_{0 \leq t \leq T} \mathbb{E}_x\left[(X_t^N)^4\right] \leq C_T(1 + x^4)$ and $\sup_{0 \leq t \leq T} \mathbb{E}_x\left[(\bar{X}_t)^4\right] \leq C_T(1 + x^4)$.

Proof. We begin with the proof of (i). Let $\Phi(x) = x^2$ and A^N be the extended generator of $(X_t^N)_{t \geq 0}$. One can note that, applying Fatou's lemma to the inequality (2.4), one obtains for all $t \geq 0, \sup_{s \leq t} \mathbb{E}\left[(X_s^N)^2\right]$ is finite. As a consequence $\Phi \in \mathcal{D}'(A^N)$ (in the sense of Definition A.0.1). And, recalling that μ is centered and that $\sigma^2 := \int_{\mathbb{R}} u^2 d\mu(u)$, we have for all $x \in \mathbb{R}$,

$$A^N \Phi(x) = -2\alpha\Phi(x) + \sigma^2 f(x) \leq -2\alpha\Phi(x) + \sigma^2 \left(L|x| + \sqrt{f(0)}\right)^2$$

$$\leq (\sigma^2 L^2 - 2\alpha)\Phi(x) + 2\sigma^2 L|x|\sqrt{f(0)} + \sigma^2 f(0).$$

Let $\varepsilon > 0$ be fixed, and $\eta_\varepsilon = 2\sigma^2 L\sqrt{f(0)}/\varepsilon$. Using that, for every $x \in \mathbb{R}, |x| \leq x^2/\eta_\varepsilon + \eta_\varepsilon$, we have

$$A^N\Phi(x) \leq c_\varepsilon \Phi(x) + d_\varepsilon, \tag{2.7}$$

with $c_\varepsilon = \sigma^2 L^2 - 2\alpha + \varepsilon$ and $d_\varepsilon = O(1/\varepsilon)$. Let us assume that $c_\varepsilon \neq 0$, possibly by reducing $\varepsilon > 0$. Considering $Y_t^N := e^{-c_\varepsilon t}\Phi(X_t^N)$, by Ito's formula,

$$\begin{aligned} dY_t^N =& - c_\varepsilon e^{-c_\varepsilon t}\Phi(X_t^N)dt + e^{-c_\varepsilon t} d\Phi(X_t^N) \\ =& - c_\varepsilon e^{-c_\varepsilon t}\Phi(X_t^N)dt + e^{-c_\varepsilon t} A^N\Phi(X_t^N)dt + e^{-c_\varepsilon t} dM_t, \end{aligned}$$

where $(M_t)_{t\geq 0}$ is a $\mathbb{P}_x-$martingale. Using (2.7), we obtain

$$Y_t^N \leq Y_0^N + d_\varepsilon \int_0^t e^{-c_\varepsilon s} ds + \int_0^t e^{-c_\varepsilon s} dM_s = Y_0^N + \frac{d_\varepsilon}{c_\varepsilon}\left(1 - e^{-c_\varepsilon t}\right),$$

implying

$$\mathbb{E}_x\left[Y_t^N\right] \leq \mathbb{E}_x\left[Y_0^N\right] + \frac{d_\varepsilon}{c_\varepsilon}\left(e^{-c_\varepsilon t} + 1\right).$$

One deduces

$$\mathbb{E}_x\left[\left(X_t^N\right)^2\right] \leq x^2 e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t} + \frac{C}{\varepsilon}\left(e^{(\sigma^2 L^2 - 2\alpha+\varepsilon)t} + 1\right), \tag{2.8}$$

for some constant $C > 0$ independent of t, ε, N.
The proof of (ii) is analogous and therefore omitted.
Now we prove (iii). From

$$X_t^N = X_0^N - \alpha\int_0^t X_s^N ds + \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{]0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbb{1}_{\left\{z\leq f(X_{s-}^N)\right\}} d\pi_j(s,z,u),$$

we deduce

$$\begin{aligned} \left(\sup_{0\leq s\leq t}\left|X_t^N\right|\right)^2 \leq& 3\left(X_0^N\right)^2 + 3\alpha^2 t\int_0^t (X_s^N)^2 ds \\ &+ 3\sum_{j=1}^N\left(\sup_{0\leq s\leq t}\left|\int_{]0,s]\times\mathbb{R}_+\times\mathbb{R}} u\mathbb{1}_{\left\{z\leq f(X_{r-}^N)\right\}} d\pi_j(r,z,u)\right|\right)^2. \end{aligned} \tag{2.9}$$

Applying Burkholder-Davis-Gundy inequality to the last term above in (2.9), we can bound its expectation by

$$\begin{aligned} 3N\mathbb{E}\left[\int_{]0,t]\times\mathbb{R}_+\times\mathbb{R}} u^2\mathbb{1}_{\left\{z\leq f(X_{s-}^N)\right\}} d\pi_j(s,z,u)\right] \leq& 3N\sigma^2\int_0^t \mathbb{E}\left[f(X_{s-}^N)\right] ds \\ \leq& 3N\sigma^2 C\int_0^t \left(1 + \mathbb{E}\left[(X_s^N)^2\right]\right) ds. \end{aligned} \tag{2.10}$$

Now, bounding the expectation of (2.9) by (2.10), and using point (i) of the lemma we conclude the proof of (iii).

Finally, (iv) can be proved in classical way, applying Itô's formula and Grönwall's lemma. Let us explain how to prove this property for the process X^N. The proof for $\bar{X}$ is similar. According to Itô's formula, for every $t \geq 0$,

$$\begin{aligned}(X_t^N)^4 = (X_0^N)^4 - 4\alpha \int_0^t (X_s^N)^4 ds \\ + \sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left[(X_{s-}^N + \frac{u}{\sqrt{N}})^4 - (X_{s-}^N)^4 \right] \mathbb{1}_{\left\{z \leq f(X_{s-}^N)\right\}} d\pi_j(s,z,u) \\ \leq (X_0^N)^4 + \sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left[(X_{s-}^N + \frac{u}{\sqrt{N}})^4 - (X_{s-}^N)^4 \right] \mathbb{1}_{\left\{z \leq f(X_{s-}^N)\right\}} d\pi_j(s,z,u).\end{aligned}$$

Let us recall that u is centered and has a finite fourth moment, and that f is subquadratic. Introducing the stopping times $\tau_K^N := \inf\{t > 0 : |X_t^N| > K\}$ for $K > 0$, and $u_K^N(t) := \mathbb{E}\left[(X_{t\wedge\tau_K^N}^N)^4\right]$, it follows from the above that for all $t \geq 0$,

$$u_K^N(t) \leq C + Ct + C\int_0^t u_K^N(s)ds,$$

where C is a constant independent of t, N and K. This implies that for all $t \geq 0$,

$$\sup_{N\in\mathbb{N}^*} \sup_{K>0} \sup_{0\leq s\leq t} u_K^N(s) < \infty.$$

Consequently, the stopping times τ_K^N tend to infinity as K goes to infinity, and Fatou's lemma allows to conclude. □

Before establishing the rate of convergence for the semigroups, we do it for their generators.

Proposition 2.2.6. *If Assumptions 2.1 and 2.2 hold, then for all $g \in C_b^3(\mathbb{R})$,*

$$\left|\bar{A}g(x) - A^N g(x)\right| \leq |f(x)|.\|g'''\|_\infty \frac{1}{6\sqrt{N}} \int_{\mathbb{R}} |u|^3 d\mu(u).$$

Proof. For $g \in C_b^3(\mathbb{R})$, if we note U a random variable having distribution μ, we have

$$\begin{aligned}\left|A^N g(x) - \bar{A}g(x)\right| \leq & |f(x)| \left| N\mathbb{E}\left[g\left(x + \frac{U}{\sqrt{N}}\right) - g(x)\right] - \frac{1}{2}\sigma^2 g''(x) \right| \\ = & |f(x)| N \left| \mathbb{E}\left[g\left(x + \frac{U}{\sqrt{N}}\right) - g(x) - \frac{U}{\sqrt{N}} g'(x) - \frac{U^2}{2N} g''(x)\right]\right| \\ \leq & |f(x)| N\mathbb{E}\left[\left| g\left(x + \frac{U}{\sqrt{N}}\right) - g(x) - \frac{U}{\sqrt{N}} g'(x) - \frac{U^2}{2N} g''(x)\right|\right].\end{aligned}$$

Using Taylor-Lagrange's inequality, we obtain the result. □

We prove in Appendix A, in Proposition A.0.3, that

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds. \tag{2.11}$$

This formula allows to obtain a speed of convergence of P_t^N to $\bar{P}_t$.

Obviously, to be able to apply the above formula, we need to ensure the regularity of $x \mapsto \bar{P}_s g(x)$, together with a control of the associated norms $||\bar{P}_s g||_{k,\infty}$, for suitable k. This is done in the next proposition.

Proposition 2.2.7. *If Assumptions 2.1, 2.2 and 2.3 hold, then for all $t \geq 0$ and for all $g \in C_b^3(\mathbb{R})$, the function $x \mapsto \bar{P}_t g(x)$ belongs to $C_b^3(\mathbb{R})$ and satisfies*

$$\left|\left|\left(\bar{P}_t g\right)'''\right|\right|_\infty \leq C||g||_{3,\infty}(1+t^2)e^{\beta t}, \tag{2.12}$$

with $\beta = \max(\frac{1}{2}\sigma^2 L^2 - \alpha, 2\sigma^2 L^2 - 2\alpha, \frac{7}{2}\sigma^2 L^2 - 3\alpha)$. Moreover, for all $T > 0$,

$$\sup_{0\leq t\leq T}||\bar{P}_t g||_{3,\infty} \leq Q_T||g||_{3,\infty} \tag{2.13}$$

for some $Q_T > 0$, and for all $i \in \{0,1,2\}$ and $x \in \mathbb{R}$, $s \mapsto (\bar{P}_s g)^{(i)}(x) = \frac{\partial^i}{\partial x^i}(\bar{P}_s g(x))$ is continuous.

Proof. To begin with, we use Theorem 1.4.1 of Kunita (1986) to prove that the flow associated to the stochastic differential equation (2.3) admits a modification which is C^3 with respect to the initial condition x (see also Theorem 4.6.5 of Kunita (1990)). Indeed the local characteristics of the flow are given by

$$b(x,t) = -\alpha x \quad \text{and} \quad a(x,y,t) = \sigma^2\sqrt{f(x)f(y)},$$

and, under Assumptions 2.1 and 2.3, they satisfy the conditions of Theorem 1.4.1 of Kunita (1986):

- $\exists C, \forall x,y,t, |b(x,t)| \leq C(1+|x|)$ and $|a(x,y,t)| \leq C(1+|x|)(1+|y|)$.
- $\exists C, \forall x,y,t, |b(x,t) - b(y,t)| \leq C|x-y|$ and $|a(x,x,t) + a(y,y,t) - 2a(x,y,t)| \leq C|x-y|^2$.
- $\forall 1 \leq k \leq 4, 1 \leq l \leq 4-k$, $\frac{\partial^k}{\partial x^k} b(x,t)$ and $\frac{\partial^{k+l}}{\partial x^k \partial y^l} a(x,y,t)$ are bounded.

In the following, we consider the process $(\bar{X}_t^{(x)})_t$, solution of the equation (2.3) and satisfying $\bar{X}_0^{(x)} = x$. Then we can consider a modification of the flow $\bar{X}_t^{(x)}$ which is C^3 with the respect to the initial condition $x = \bar{X}_0^{(x)}$. It is then sufficient to control the moment of the derivatives of $\bar{X}_t^{(x)}$ with respect to x, since with those controls we will have

$$\begin{aligned}
\bar{P}_t g(x) =& \mathbb{E}\left[g\left(\bar{X}_t^{(x)}\right)\right], \quad \left(\bar{P}_t g\right)'(x) = \mathbb{E}\left[\frac{\partial \bar{X}_t^{(x)}}{\partial x} g'\left(\bar{X}_t^{(x)}\right)\right], \\
\left(\bar{P}_t g\right)''(x) =& \mathbb{E}\left[\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} g'\left(\bar{X}_t^{(x)}\right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x}\right)^2 g''\left(\bar{X}_t^{(x)}\right)\right], \\
\left(\bar{P}_t g\right)'''(x) =& \mathbb{E}\left[\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} g'\left(\bar{X}_t^{(x)}\right) + 3\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \cdot \frac{\partial \bar{X}_t^{(x)}}{\partial x} g''\left(\bar{X}_t^{(x)}\right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x}\right)^3 g'''\left(\bar{X}_t^{(x)}\right)\right].
\end{aligned} \tag{2.14}$$

We start with the representation

$$\bar{X}_t^{(x)} = xe^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} \sqrt{f\left(\bar{X}_s^{(x)}\right)} dB_s.$$

This implies

$$\frac{\partial \bar{X}_t^{(x)}}{\partial x} = e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} \frac{\partial \bar{X}_s^{(x)}}{\partial x} \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) dB_s. \tag{2.15}$$

Writing $U_t = e^{\alpha t} \frac{\partial \bar{X}_t^{(x)}}{\partial x}$ and

$$M_t = \int_0^t \sigma \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) dB_s, \tag{2.16}$$

we obtain $U_t = 1 + \int_0^t U_s dM_s$, whence

$$U_t = \exp\left(M_t - \frac{1}{2} < M >_t\right).$$

Notice that this implies $U_t > 0$ almost surely, whence $\frac{\partial \bar{X}_t^{(x)}}{\partial x} > 0$ almost surely. Hence

$$U_t^p = e^{pM_t - \frac{p}{2}<M>_t} = \exp\left(pM_t - \frac{1}{2}p^2 < M >_t\right) e^{\frac{1}{2}p(p-1)<M>_t} = \mathcal{E}(M)_t e^{\frac{1}{2}p(p-1)<M>_t}.$$

Since $\left(\sqrt{f}\right)'$ is bounded, M_t is a martingale, thus $\mathcal{E}(M)$ is an exponential martingale with expectation 1, implying that

$$I\!\!E U_t^p \leq e^{\frac{1}{2}p(p-1)\sigma^2 m_1^2 t}, \tag{2.17}$$

where m_1 is the bound of $(\sqrt{f})'$ introduced in Assumption 2.3. In particular we have

$$\mathbb{E}\left[\left(\frac{\partial \bar{X}_t^{(x)}}{\partial x}\right)^2\right] \leq e^{(\sigma^2 m_1^2 - 2\alpha)t} \text{ and } \mathbb{E}\left[\left|\frac{\partial \bar{X}_t^{(x)}}{\partial x}\right|^3\right] \leq e^{(3\sigma^2 m_1^2 - 3\alpha)t}. \tag{2.18}$$

Differentiating (2.15) with respect to x, we obtain

$$\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} = \sigma \int_0^t e^{-\alpha(t-s)} \left[\frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x}\right)^2 \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right)\right] dB_s. \tag{2.19}$$

We introduce $V_t = \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} e^{\alpha t}$ and deduce from this that

$$V_t = \sigma \int_0^t \left[V_s \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) + e^{-\alpha s} U_s^2 \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right)\right] dB_s,$$

which can be rewritten as

$$dV_t = V_t dM_t + Y_t dB_t, V_0 = 0, Y_t = \sigma e^{-\alpha t} U_t^2 \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_t^{(x)}\right),$$

with M_t as in (2.16). Applying Ito's formula to $Z_t := V_t/U_t$ (recall that $U_t > 0$), we obtain

$$dZ_t = \frac{Y_t}{U_t} dB_t - \frac{Y_t}{U_t} d < M, B >_t,$$

such that, by the precise form of Y_t and since $Z_0 = 0$,

$$Z_t = \sigma \int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) dB_s - \sigma^2 \int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) ds.$$

Using Jensen's inequality, (2.17) and Burkholder-Davis-Gundy inequality, for all $t \geq 0$,

$$\begin{aligned}\mathbb{E}\left[Z_t^4\right] \leq C \Bigg(\mathbb{E}\left[\left(\int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) dB_s\right)^4\right] \\ +\mathbb{E}\left[\left(\int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) ds\right)^4\right]\Bigg) \\ \leq C\left(t+t^3\right) \int_0^t e^{(6\sigma^2 m_1^2 - 4\alpha)s} ds \\ \leq C\left(t+t^3\right)\left(1+t+e^{(6\sigma^2 m_1^2-4\alpha)t}\right). \quad (2.20)\end{aligned}$$

We deduce that

$$\mathbb{E}\left[V_t^2\right] \leq \mathbb{E}\left[Z_t^4\right]^{1/2} \mathbb{E}\left[U_t^4\right]^{1/2} \leq C(t^{1/2}+t^2) e^{3\sigma^2 m_1^2 t},$$

whence

$$\mathbb{E}\left[\left(\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2}\right)^2\right] \leq C(t^{1/2}+t^2) e^{(3\sigma^2 m_1^2 - 2\alpha)t}. \quad (2.21)$$

Finally, differentiating (2.19), we get

$$\begin{aligned}\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} = \sigma \int_0^t e^{-\alpha(t-s)} \Bigg[\frac{\partial^3 \bar{X}_s^{(x)}}{\partial x^3} \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) + 3 \frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} \frac{\partial \bar{X}_s^{(x)}}{\partial x} \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) \\ + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x}\right)^3 \left(\sqrt{f}\right)^{(3)} \left(\bar{X}_s^{(x)}\right)\Bigg] dB_s.\end{aligned}$$

Introducing $W_t = e^{\alpha t} \frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3}$, we obtain

$$W_t = \sigma \int_0^t \left[W_s \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) + 3 e^{-\alpha s} U_s V_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) + e^{-2\alpha s} U_s^3 \left(\sqrt{f}\right)^{(3)} \left(\bar{X}_s^{(x)}\right)\right] dB_s.$$

Once again we can rewrite this as

$$dW_t = W_t dM_t + Y_t' dB_t, W_0 = 0,$$

where

$$Y_t' = \sigma \left(3 e^{-\alpha t} U_t V_t \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_t^{(x)}\right) + e^{-2\alpha t} U_t^3 \left(\sqrt{f}\right)^{(3)} \left(\bar{X}_t^{(x)}\right)\right),$$

whence, introducing $Z'_t = \frac{W_t}{U_t}$,

$$Z'_t = \int_0^t \frac{Y'_s}{U_s} dB_s - \int_0^t \frac{Y'_s}{U_s} d < M, B >_s .$$

As previously, we obtain,

$$\begin{aligned}
\mathbb{E}\left[(Z'_t)^2\right] \leq & C(1+t)\int_0^t \mathbb{E}\left[\left(\frac{Y'_s}{U_s}\right)^2\right] ds \\
\leq & C(1+t)\int_0^t \left(e^{-2\alpha s}\mathbb{E}\left[V_s^2\right] + e^{-4\alpha s}\mathbb{E}\left[U_s^4\right]\right) ds \\
\leq & C(1+t)\int_0^t \left((s^{1/2}+s^2)e^{(3\sigma^2 m_1^2-2\alpha)s} + e^{(6\sigma^2 m_1^2-4\alpha)s}\right) ds \\
\leq & C(1+t^3)\int_0^t \left(e^{(3\sigma^2 m_1^2-2\alpha)s} + e^{(6\sigma^2 m_1^2-4\alpha)s}\right) ds \\
\leq & C(1+t^3)\int_0^t \left(1 + e^{(6\sigma^2 m_1^2-4\alpha)s}\right) ds \quad \leq C(1+t^4)\left(1+e^{(6\sigma^2 m_1^2-4\alpha)t}\right). \qquad (2.22)
\end{aligned}$$

As a consequence,

$$\begin{aligned}
\mathbb{E}\left[|W_t|\right] \leq & \mathbb{E}\left[(Z'_t)^2\right]^{1/2}\mathbb{E}\left[U_t^2\right]^{1/2} \leq C(1+t^2)\left(1+e^{(3\sigma^2 m_1^2-2\alpha)t}\right)e^{\frac{1}{2}\sigma^2 m_1^2 t} \\
\leq & C(1+t^2)\left(e^{\frac{1}{2}\sigma^2 m_1^2 t} + e^{(\frac{7}{2}\sigma^2 m_1^2-2\alpha)t}\right),
\end{aligned}$$

implying

$$\mathbb{E}\left[\left|\frac{\partial^3 \bar{X}_t^{(x)}}{\partial^3 x}\right|\right] \leq C(1+t^2)\left(e^{(\frac{1}{2}\sigma^2 m_1^2-\alpha)t} + e^{(\frac{7}{2}\sigma^2 m_1^2-3\alpha)t}\right). \qquad (2.23)$$

Finally, using Cauchy-Schwarz inequality, and inserting (2.18), (2.21) and (2.23) in (2.14),

$$\left|\left|(\bar{P}_t g)'''\right|\right|_\infty \leq C||g||_{3,\infty}(1+t^2)\left(e^{(\frac{1}{2}\sigma^2 m_1^2-\alpha)t} + e^{2(\sigma^2 m_1^2-\alpha)t} + e^{(\frac{7}{2}\sigma^2 m_1^2-3\alpha)t}\right),$$

which proves the first assertion of the proposition. The proof of the second assertion, equation (2.13), follows similarly. Finally to prove the third assertion, we first study the regularity of the first derivative. Notice that $t \mapsto \frac{\partial \bar{X}_t^{(x)}}{\partial x}$ is almost surely continuous by equation (2.15). Now take any sequence $t_n \to t$. By (2.18), the family of random variables $\left\{\frac{\partial \bar{X}_{t_n}^{(x)}}{\partial x} g'(\bar{X}_{t_n}^{(x)}), n \geq 1\right\}$ is uniformly integrable. As a consequence, the second formula in (2.14) implies that $(\bar{P}_{t_n} g)'(x) \to (\bar{P}_t g)'(x)$ as $n \to \infty$, whence the desired continuity. The argument is similar for the second derivative, using (2.19) and (2.21). That concludes the proof. □

Now we can compute a rate of convergence of the semigroups.

Proof of Theorem 2.2.3. Step 1. The main part of the proof will be to check that the conditions stated in Proposition A.0.3 below allowing to obtain formula (2.11) are satisfied. This will be done in *Step 2* below. Indeed, once this is shown, the rest of the proof will be a straightforward consequence of Proposition 2.2.6, since, applying formula (2.11),

$$\begin{aligned}
\left|\bar{P}_t g(x) - P_t^N g(x)\right| =& \left|\int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds\right| \\
\leq& \int_0^t \mathbb{E}_x^N \left[\left|\bar{A}\left(\bar{P}_s g\right)\left(X_{t-s}^N\right) - A^N \left(\bar{P}_s g\right)\left(X_{t-s}^N\right)\right|\right] ds \\
\leq& C \frac{1}{\sqrt{N}} \int_0^t \left|\left|\left(\bar{P}_s g\right)'''\right|\right|_\infty \mathbb{E}_x \left[f\left(X_{t-s}^N\right)\right] ds \\
\leq& C \frac{1}{\sqrt{N}} ||g||_{3,\infty} \int_0^t \left((1+s^2) e^{\beta s} \left(1 + \mathbb{E}_x \left[\left(X_{t-s}^N\right)^2\right]\right)\right) ds \\
\leq& C \left(1 + \frac{1}{\varepsilon}\right) \frac{1}{\sqrt{N}} ||g||_{3,\infty} (1+x^2) \int_0^t (1+s^2) e^{\beta s} \left(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)(t-s)}\right) ds,
\end{aligned}$$

where we have used respectively Proposition 2.2.7 and Lemma 2.2.5.(i) to obtain the two last inequalities above, and ε is any positive constant.

Step 2. Now we show that X^N and $\bar{X}$ satisfy the hypotheses of Proposition A.0.3. To begin with we know that $\bar{X}$ and X^N satisfy hypotheses (i) and (ii), using Lemma 2.2.5.

We now show how hypothesis (iii) can be proved for the processes X^N and $\bar{X}$ solving the equations (2.2) and (2.3), based on Lemma 2.2.5. Indeed, by Lemma 2.2.5, Itô's isometry and Jensen's inequality, for all $0 \leq s \leq t$,

$$\begin{aligned}
\mathbb{E}_x^N \left[(X_t^N - X_s^N)^2\right] =& \mathbb{E}_x \left[\left(-\alpha \int_s^t X_r^N dr + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]s,t] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbb{1}_{\left\{z \leq f(X_{r-}^N)\right\}} d\pi_j(r,z,u)\right)^2\right] \\
\leq& 2\alpha^2 (t-s) \int_s^t \mathbb{E}_x^N \left[(X_r^N)^2\right] dr + 2\sigma^2 \int_s^t \mathbb{E}_x^N \left[f(X_r^N)\right] dr \\
\leq& 2\alpha^2 C_t (1+x^2)(t-s)^2 + 2\sigma^2 C_t (1+x^2)(t-s).
\end{aligned}$$

This proves that X^N satisfies hypothesis (iii). A similar computation holds true for $\bar{X}$.

$\bar{P}$ satisfies hypothesis (iv) thanks to Proposition 2.2.7.

We now show that $\bar{P}$ satisfies hypothesis (v) using the calculations of the proof of Proposition 2.2.7. We first study the regularity of the first derivative. Notice that $t \mapsto \frac{\partial \bar{X}_t^{(x)}}{\partial x}$ is almost surely continuous by equations (2.15). Now take any sequence $t_n \to t$. By (2.18), the family of random variables $\left\{\frac{\partial \bar{X}_{t_n}^{(x)}}{\partial x} g'(\bar{X}_{t_n}^{(x)}), n \geq 1\right\}$ is uniformly integrable. As a consequence, the second formula in (2.14) implies that $(\bar{P}_{t_n} g)'(x) \to (\bar{P}_t g)'(x)$ as $n \to \infty$, whence the desired continuity. The argument is similar for the second derivative, using (2.19) and (2.21).

Using Lemma 2.2.4, we see directly that $\bar{A}$ and A^N satisfy hypotheses (vi) and $(viii)$. In addition (vii) is straightforward for $\bar{A}$, and it is a consequence of Lemma 2.2.8 below for A^N. The only remaining hypothesis (ix) is a straightforward consequence of Lemma 2.2.9 below. □

Lemma 2.2.8. *For all $g \in C_c^2(\mathbb{R})$ such that Supp $g \subseteq [-M, M]$, we have*

$$\left|\left|\left(A^N g\right)'\right|\right|_\infty \leq C \|g\|_{2,\infty} \left(1 + M^2\right),$$

for some constant $C > 0$.

Proof. We have

$$\begin{aligned}\left(A^N g\right)'(x) = & -\alpha g'(x) - \alpha x g''(x) - Nf'(x)g(x) - Nf(x)g'(x) \\ & + Nf'(x)\mathbb{E}\left[g\left(x + \frac{U}{\sqrt{N}}\right)\right] + Nf(x)\mathbb{E}\left[g'\left(x + \frac{U}{\sqrt{N}}\right)\right].\end{aligned}$$

Then it is clear that for all $x \in \mathbb{R}$, we have

$$\left|\left(A^N g\right)'(x)\right| \leq C_N ||g||_{2,\infty}\left(1+M^2\right) + \left|Nf'(x)\mathbb{E}\left[g\left(x + \frac{U}{\sqrt{N}}\right)\right]\right| + \left|Nf(x)\mathbb{E}\left[g'\left(x + \frac{U}{\sqrt{N}}\right)\right]\right|. \tag{2.24}$$

We bound the jump terms using the subquadraticty of f and f' (indeed with Assumptions 2.1 and 2.3, we know that f' is sublinear, and consequently subquadratic). We can write:

$$\begin{aligned}\mathbb{E}\left[\left|g'\left(x + \frac{U}{\sqrt{N}}\right)\right|\right] \leq & ||g'||_\infty \mathbb{E}\left[\mathbb{1}_{\{|x+U/\sqrt{N}| \leq M\}}\right] \\ = & ||g'||_\infty \mathbb{P}\left(\left\{x + \frac{U}{\sqrt{N}} \geq -M\right\} \cap \left\{x + \frac{U}{\sqrt{N}} \leq M\right\}\right) \\ \leq & ||g'||_\infty \mathbb{P}\left(\left\{x + \frac{|U|}{\sqrt{N}} \geq -M\right\} \cap \left\{x - \frac{|U|}{\sqrt{N}} \leq M\right\}\right) \\ = & ||g'||_\infty \mathbb{P}\left(\left\{|U| \geq -\sqrt{N}(M+x)\right\} \cap \left\{|U| \geq \sqrt{N}(x-M)\right\}\right).\end{aligned}$$

Then for $x > M+1$, using that $f(x) \leq C(1+x^2)$, and for a constant C that may change from line to line,

$$\begin{aligned}\left|f(x)\mathbb{E}\left[g'\left(x + \frac{U}{\sqrt{N}}\right)\right]\right| \leq & C||g'||_\infty \left(1+x^2\right)\mathbb{P}\left(|U| \geq \sqrt{N}(x-M)\right) \\ \leq & C\frac{1}{N}\mathbb{E}\left[U^2\right]||g'||_\infty \frac{1+x^2}{(x-M)^2} \\ \leq & C||g'||_\infty \left(1+M^2\right).\end{aligned}$$

The last inequality comes from the fact that the function $x \in [M+1, +\infty[\mapsto \frac{1+x^2}{(x-M)^2}$ is bounded by $1+(M+1)^2$. With the same reasoning, we know that for all $x < -M-1$, we have

$$\left|f(x)\mathbb{E}\left[g'\left(x + \frac{U}{\sqrt{N}}\right)\right]\right| \leq C||g'||_\infty \left(1+M^2\right).$$

This concludes the proof. □

Lemma 2.2.9. *Let $(g_k)_k$ be a sequence of $C_b^1(\mathbb{R})$ satisfying $\sup_k ||g_k'||_\infty < \infty$, and for all $x \in \mathbb{R}$, $g_k(x) \to 0$ as $k \to \infty$.*

Then for all bounded sequences of real numbers $(x_k)_k$, $g_k(x_k) \to 0$ as $k \to \infty$.

Proof. Let $(x_k)_k$ be a bounded sequence. In a first time, we suppose that $(x_k)_k$ converges to some $x \in \mathbb{R}$. Then we have $|g_k(x_k)| \leq ||g_k'||_\infty |x - x_k| + |g_k(x)|$ which converges to zero as k goes to infinity. In the general case, we show that for all subsequence of $(g_k(x_k))_k$, there exists a subsequence of the first one that converges to 0 (the second subsequence has to be chosen such that x_k converges). □

2.3 Convergence of the Markovian kernel

The main result of this section is the convergence of the Markovian kernel of X^N to the invariant measure of $\bar{X}$.

Theorem 2.3.1. *Under Assumptions 2.1 and 2.2, $\bar{X}$ is recurrent in the sense of Harris, having invariant probability measure $\pi(dx) = p(x)dx$ with density*

$$p(x) = C\frac{1}{f(x)} \exp\left(-\frac{2\alpha}{\sigma^2}\int_0^x \frac{y}{f(y)}dy\right).$$

Besides, if Assumption 2.3 holds, then for all $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

$$\left|P_t^N g(x) - \pi g\right| \leq C||g||_{3,\infty}(1+x^2)\left(\frac{K_t}{\sqrt{N}} + e^{-\gamma t}\right),$$

where C and γ are positive constants independent of N and t, and where K_t is defined in (2.6). In particular, $P_t^N(x,\cdot)$ converges weakly to π as $(N,t) \to (\infty,\infty)$, provided $K_t = o(\sqrt{N})$.

If we assume, in addition, that $\alpha > \frac{7}{6}\sigma^2 L^2$, then $P_t^N(x,\cdot)$ converges weakly to π as $(N,t) \to (\infty,\infty)$ without any condition on the joint convergence of (t,N), and we have, for any $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

$$\left|P_t^N g(x) - \pi g\right| \leq C||g||_{3,\infty}(1+x^2)\left(\frac{1}{\sqrt{N}} + e^{-\gamma t}\right).$$

We begin by proving some properties of the invariant measure of $\bar{P}_t$. In what follows we use the total variation distance between two probability measures ν_1 and ν_2 defined by

$$\|\nu_1 - \nu_2\|_{TV} = \frac{1}{2}\sup_{g:\|g\|_\infty \leq 1} |\nu_1(g) - \nu_2(g)|.$$

Proposition 2.3.2. *If Assumptions 2.1 and 2.2 hold, then the invariant measure π of $(\bar{P}_t)_t$ exists and is unique. Its density is given, up to multiplication with a constant, by*

$$p(x) = C\frac{1}{f(x)} \exp\left(-\frac{2\alpha}{\sigma^2}\int_0^x \frac{y}{f(y)}dy\right).$$

In addition, if Assumption 2.3 holds, then for every $0 < q < 1/2$, there exists some $\gamma > 0$ such that, for all $t \geq 0$,

$$||\bar{P}_t(x,\cdot) - \pi||_{TV} \leq C\left(1+x^2\right)^q e^{-\gamma t}.$$

Proof. In a first time, let us prove the positive Harris recurrence of $\bar{X}$ implying the existence and uniqueness of π. According to Example 3.10 of Khasminskii (2012) it is sufficient to show that $S(x) := \int_0^x s(y)dy$ goes to $+\infty$ (resp. $-\infty$) as x goes to $+\infty$ (resp. $-\infty$), where

$$s(x) := \exp\left(\frac{2\alpha}{\sigma^2}\int_0^y \frac{v}{f(v)}dv\right).$$

For $x > 0$, and using that f is subquadratic,

$$s(x) \geq \exp\left(C\int_0^y \frac{2v}{1+v^2}dv\right) = \exp\left(C\ln(1+y^2)\right) = (1+y^2)^C \geq 1,$$

implying that $S(x)$ goes to $+\infty$ as x goes to $+\infty$. With the same reasoning, we obtain that $S(x)$ goes to $-\infty$ as x goes to $-\infty$. Finally, the associated invariant density is given, up to a constant, by

$$p(x) = \frac{C}{f(x)s(x)}.$$

For the second part of the proof, take $V(x) = (1+x^2)^q$, for some $q < 1/2$, then

$$V'(x) = 2qx(1+x^2)^{q-1}, V''(x) = 2q(1+x^2)^{q-2}[2x^2(q-1) + (1+x^2)].$$

As $q < \frac{1}{2}$, $V''(x) < 0$ for x^2 sufficiently large, say, for $|x| \geq K$. In this case, for $|x| \geq K$,

$$\bar{A}V(x) \leq -2\alpha q x^2(1+x^2)^{q-1} \leq -2\alpha q \frac{x^2}{1+x^2} V(x) \leq -2q\alpha \frac{K^2}{1+K^2} V(x) = -cV(x).$$

So we obtain all in all for suitable constants c and d that, for any $x \in \mathbb{R}$,

$$\bar{A}V(x) \leq -cV(x) + d. \tag{2.25}$$

For any fixed $T > 0$, the sampled chain $(\bar{X}_{kT})_{k\geq 0}$ is Feller and $\pi-$irreducible. The support of π being $\mathbb{R}$, Theorem 3.4 of Meyn and Tweedie (1992) implies that every compact set is petite for the sampled chain. Then, as (2.25) implies the condition $(CD3)$ of Theorem 6.1 of Meyn and Tweedie (1993), we have the following bound: introducing for any probability measure μ the weighted norm

$$\|\mu\|_V := \sup_{g:|g|\leq 1+V} |\mu(g)|,$$

there exist $C, \gamma > 0$ such that

$$\|\bar{P}_t(x,\cdot) - \pi\|_V \leq C(1+V(x))e^{-\gamma t}.$$

This implies the result, since $||\cdot||_{TV} \leq ||\cdot||_V$. □

Now the proof of Theorem 2.3.1 is straightforward.

Proof of Theorem 2.3.1. The first part of the theorem has been proved in Proposition 2.3.2. For the second part, for any $g \in C_b^3(\mathbb{R})$,

$$\begin{aligned}\left|P_t^N g(x) - \pi g\right| &\leq \left|P_t^N g(x) - \bar{P}_t g(x)\right| + \left|\bar{P}_t g(x) - \pi g\right| \\ &\leq \frac{K_t}{\sqrt{N}}(1+x^2)||g||_{3,\infty} + ||g||_\infty ||\bar{P}_t(x,\cdot) - \pi||_{TV} \\ &\leq ||g||_{3,\infty} C\left(\frac{K_t}{\sqrt{N}}(1+x^2) + e^{-\gamma t}(1+x^2)^q\right),\end{aligned}$$

where we have used Theorem 2.2.3 and Proposition 2.3.2. Since $(1+x^2)^q \leq 1+x^2$, q being smaller than $1/2$, this implies the result. □

2.4 Convergence of the system of point processes

In this section, we show the convergence of the point processes $Z^{N,i}$ defined in (2.1) to some limit point processes $\bar{Z}^i$ having stochastic intensity $f(\bar{X}_t)$ at time t. To define the processes $\bar{Z}^i$ $(i \in \mathbb{N}^*)$, we fix a Brownian motion $(B_t)_{t\geq 0}$ on some probability space different from the one where the processes X^N $(N \in \mathbb{N}^*)$ and the Poisson random measures π_i $(i \in \mathbb{N}^*)$ are defined. Then we fix a family of i.i.d. Poisson random measures $\bar{\pi}_i$ $(i \in \mathbb{N}^*)$ on the same space as $(B_t)_{t\geq 0}$, independent of $(B_t)_{t\geq 0}$. This independence property is natural (see Proposition 2.4.4 below), and it allows us to consider the joint distributions $(\bar{X}, \bar{\pi}_1, \ldots, \bar{\pi}_k, \ldots)$, where $\bar{X}$ is defined as the solution of (2.3) driven by $(B_t)_{t\geq 0}$. The limit point processes $\bar{Z}^i$ are then defined by

$$\bar{Z}^i_t = \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\leq f(\bar{X}_s)\right\}} d\pi_i(s,z,u). \tag{2.26}$$

The main result of this section is

Theorem 2.4.1. *Under Assumptions 2.1, 2.2, 2.3 and 2.4, the sequence* $\left(Z^{N,1}, Z^{N,2}, \ldots, Z^{N,k}, \ldots\right)_N$ *converges to* $\left(\bar{Z}^1, \bar{Z}^2, \ldots, \bar{Z}^k, \ldots\right)$ *in distribution in* $D\left(\mathbb{R}_+, \mathbb{R}\right)^{\mathbb{N}^*}$.

Let us give a brief interpretation of the above result. Conditionally on $\bar{X}$, $\bar{Z}^1, \ldots, \bar{Z}^k$ are independent. Therefore, the above result can be interpreted as a conditional propagation of chaos property (compare to Carmona et al. (2016) dealing with the situation where all interacting components are subject to common noise). In our case, the common noise, that is, the Brownian motion B driving the dynamic of $\bar{X}$, emerges in the limit as a consequence of the central limit theorem.

Remark 2.4.2. *In Theorem 2.4.1, we implicitly define* $Z^{N,i} := 0$ *for each* $i \geq N+1$.

To prove the convergence of $Z^{N,i}$ to $\bar{Z}^i$ $(i \in \mathbb{N}^*)$, we start by proving the convergence of their stochastic intensities. This is a straightforward consequence of Theorem 2.2.1 and the following lemma.

Lemma 2.4.3. *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a continuous function. Then the function* $\Psi : x \in D(\mathbb{R}_+, \mathbb{R}) \to f \circ x \in D(\mathbb{R}_+, \mathbb{R})$ *is continuous.*

Proof. Let us consider a sequence $(x_n)_n$ of $D(\mathbb{R}_+, \mathbb{R})$ that converges to some x. We fix a $T > 0$ such that $(x_n)_n$ converges to x in $D([0,T], \mathbb{R})$. Then we can consider increasing functions λ_N defined on $[0,T]$ such that $\lambda_N(0) = 0$, $\lambda_N(T) = T$, $||Id - \lambda_N||_{\infty,[0,T]}$ vanishes and $||x_N - x \circ \lambda_N||_{\infty,[0,T]}$ vanishes as $N \to \infty$.

For N big enough, we know that $||x_N||_{\infty,[0,T]} \leq ||x||_{\infty,[0,T]} + 1$. Introducing the modulus of continuity w of f restricted to $\left[0, 2||x||_{\infty,[0,T]} + 1\right]$, $w : \left[0, 2||x||_{\infty,[0,T]} + 1\right] \to \mathbb{R}_+$, we have

$$||f \circ x_N - f \circ x \circ \lambda_N||_{\infty,[0,T]} \leq w\left(||x_N - x \circ \lambda_N||_{\infty,[0,T]}\right) \to 0$$

as $N \to \infty$. □

To prove the convergence of $Z^{N,i}$ to $\bar{Z}^i$, the convergence of their respective intensities (that is, the convergence of $f\left(X^N_t\right)_{t\geq 0}$ to $f\left(\bar{X}_t\right)_{t\geq 0}$) is not sufficient, since we also manipulate the Poisson random measure π_i. So we need to prove the convergence of the pair $\left(X^N, \pi_i\right)$ to $\left(\bar{X}, \bar{\pi}_i\right)$ in distribution in order to apply Theorem B.0.1. This theorem roughly states that, under appropriate conditions, the convergence in distribution in $D(\mathbb{R}_+, \mathbb{R})$ of the stochastic intensities of points processes implies the convergence the point processes in distribution in $D(\mathbb{R}_+, \mathbb{R})$.

Proposition 2.4.4. *Under Assumptions 2.1, 2.2, 2.3 and 2.4, for each $k \geq 1$, the sequence $\mathcal{L}\left(X^N, \pi_1, \ldots, \pi_k\right)$ converges weakly to $\mathcal{L}\left(\bar{X}\right) \otimes \mathcal{L}\left(\pi_1\right) \otimes \ldots \otimes \mathcal{L}\left(\pi_k\right)$.*

Proof. We just prove the proposition for $k = 1$ to simplify the proof, but the general case is almost the same.

Recall that $D(\mathbb{R}_+, \mathbb{R})$ is separable and complete (see Theorem 16.3 of Billingsley (1999)), and $\mathcal{M}$, the space of locally finite measures endowed with the topology of the vague convergence, is also separable and complete (Theorem $A2.6.III.(i)$ of Daley and Vere-Jones (2003)). Hence the product of the metric spaces $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M})$ is also separable and complete. Since the sequence $(X_N)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R})$ and π^1 is tight on $\mathcal{M}$, (see Theorem 1.3 of Billingsley (1999)), the couple (X_N, π^1) is tight on $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M})$.

Thus it suffices to show that any weakly convergent subsequence of $\mathcal{D}\left(X^N, \pi_1\right)$ converges to $\mathcal{D}(\bar{X}) \otimes \mathcal{D}(\pi^1)$ (see Corollary of Theorem 5.1 of Billingsley (1999)). To simplify the notations we assume that $\mathcal{D}(X^N, \pi^1)$ is already a weakly-converging subsequence, converging to some limit P.

Let $(Y, \pi) \in (D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M})$ such that $(Y, \pi) \sim P$. It is easy to see that

$$Y \sim \bar{X} \text{ and } \pi \sim \pi^1,$$

but we do not know yet if both are independent.

In the sequel we suppose that (Y, π) is defined on a filtered probability space $(\Omega', \mathcal{A}', (\mathcal{F}_t)'_{t \geq 0}, P')$, where

$$\mathcal{F}_t{}' = \bigcap_{T > t} \mathcal{F}_T^0, \mathcal{F}_t{}^0 = \sigma(Y_s, \pi([0, s] \times A), A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}),\ s \leq t).$$

Step 1. We show that π is a $(P', (\mathcal{F}_t{}^0)_{t \geq 0})-$Poisson random measure on $[0, +\infty[\times \mathbb{R}_+ \times \mathbb{R}$, with non-random compensator measure $dt \times \nu$ where $\nu = dz \times \mu(du)$.

For that sake, it is sufficient to show that for all $s < t$, disjoint sets $U_1, \ldots, U_k \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$, and $\lambda_1, \ldots, \lambda_k \geq 0$,

$$I\!\!E\left(\exp\left[-\sum_{i=1}^{k} \lambda_i \pi(]s, t] \times U_i)\right] | \mathcal{F}_s^0\right) = \exp\left[(t - s) \sum_{i=1}^{k} (e^{-\lambda_i} - 1) \nu(U_i)\right]. \tag{2.27}$$

To prove (2.27), it suffices to show that for all $s_1 < \ldots < s_n < s$, all bounded $\varphi_1, \ldots \varphi_n$, disjoint sets $U_1, \ldots, U_k \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$, and sets $V_1, \ldots, V_n \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$,

$$\begin{aligned} I\!\!E\left(\exp\left[-\sum_{i=1}^{k} \lambda_i \pi(]s, t] \times U_i)\right] \varphi_1(Z_{s_1}) \times \ldots \times \varphi_n(Z_{s_n})\right) = \\ \exp\left[(t - s) \sum_{i=1}^{k} (e^{-\lambda_i} - 1) \nu(U_i)\right] I\!\!E\left(\varphi_1(Z_{s_1}) \times \ldots \times \varphi_n(Z_{s_n})\right), \end{aligned} \tag{2.28}$$

where $Z_{s_i} = (Y_{s_i}, \pi([0, s_i[\times V_i)$

The previous equality holds if we replace Y by X^N and π by π_1, because $\pi^1(]s, t] \times U_i)$ and $Z^N_{s_1}, \ldots, Z^N_{s_n}$ are independent, where $Z^N_{s_i} = \left(X^N_{s_i}, \pi_1([0, s_i[\times V_i\right)$.

This implies that π is a $\left(P', (\mathcal{F}_t^0)_t\right)-$Poisson random measure. By right continuity of $s \mapsto \exp\left[(t - s) \sum_{i=1}^{m} (e^{-\lambda_i} - 1) \nu(U_i)\right]$, this implies that π is also a Poisson random measure with respect to $(P', (\mathcal{F}'_t)_t)$.

Step 2. Fix a test function $\varphi \in C_b^3$. Now we show that

$$\varphi(Y_t) - \varphi(Y_0) + \alpha \int_0^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_0^t \varphi''(Y_v) f(Y_v) dv \tag{2.29}$$

is a $\left(\mathcal{F}_t^0\right)_t$ −martingale. Fix $s_1 < s_2 < \ldots < s_n \leq s < t$ together with continuous and bounded test functions ψ_i and disjoint sets $U_1, \ldots, U_n \in \mathcal{B}(\mathbb{R}_+, \times \mathbb{R})$. Denote $Z_{s_i} = (Y_{s_i}, \pi([0, s_i[\times U_i)$. It suffices to show that

$$I\!\!E \left[\left(\varphi(Y_t) - \varphi(Y_s) + \alpha \int_s^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_s^t \varphi''(Y_v) f(Y_v) dv \right) \prod_{i=1}^n \psi_i(Z_{s_i}) \right] = 0. \tag{2.30}$$

To prove (2.30), we shall use that

$$W_t^N = \varphi\left(X_t^N\right) + \alpha \int_0^t \varphi'\left(X_v^N\right) X_v^N dv - N \int_0^t dv \int_{\mathbb{R}} d\mu(u) \left[\varphi\left(X_v^N + \frac{u}{\sqrt{N}} \right) - \varphi\left(X_v^N\right) \right] f\left(X_v^N\right)$$

is a $\left(\mathcal{F}_t^0\right)_t$ − martingale.

As a consequence, for all $N \geq 1$,

$$\mathbb{E}\left[\left(W_t^N - W_s^N\right) \prod_{k=1}^n \psi_k\left(Z_{s_k}^N\right) \right] = 0.$$

Using the integral form of the remainder in Taylor's formula applied to the expression in the jump term of W_t^N, we can write $W_t^N - W_s^N$ as

$$\varphi\left(X_t^N\right) - \varphi\left(X_s^N\right) + \alpha \int_s^t \varphi'\left(X_v^N\right) X_v^N dv - \frac{\sigma^2}{2} \int_s^t \varphi''\left(X_v^N\right) f\left(X_v^N\right) dv + \frac{1}{\sqrt{N}} \Phi,$$

where Φ is a random variable whose expectation is bounded uniformly in N. Thus,

$$\mathbb{E}\left[\left(W_t^N - W_s^N\right) \prod_{k=1}^n \psi_k\left(Z_{s_k}^N\right) \right] = \mathbb{E}\left[F_{s,t}\left(X^N, \pi_1\right)\right] + \frac{1}{\sqrt{N}} \mathbb{E}\left[\Phi \prod_{k=1}^n \psi_k\left(Z_{s_k}^N\right) \right],$$

where

$$F_{s,t}(x,m) = \left(\varphi(x_t) - \varphi(x_s) + \alpha \int_s^t \varphi'(x_v) x_v dv - \frac{\sigma^2}{2} \int_s^t \varphi''(x_v) f(x_v) dv \right) \prod_{k=1}^n \psi_k\left(x_{s_k}, m([0, s_k[\times U_k)\right)$$

is a continuous function on $D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}$. If $F_{s,t}$ was bounded we could make N go to infinity in the previous expression (since $\left(X^N, \pi_1\right)$ converge in distribution to (Y, π)). So we have to truncate and consider $F_{s,t}^M(x,m) := F_{s,t}(x,m) \cdot \xi_M\left(\sup_{0 \leq r \leq t} |x_r| \right)$, where $\xi_M : \mathbb{R} \to [0,1]$ is C^∞ and verifies $\mathbb{1}_{\{|x| \leq M\}} \leq \xi_M(x) \leq \mathbb{1}_{\{|x| \leq M+1\}}$.

Recall that we want to show (2.30), that is, $\mathbb{E}\left[F_{s,t}\left(Y,\pi\right)\right]=0$. We start from

$$
\begin{aligned}
\left|\mathbb{E}\left[F_{s,t}\left(Y,\pi\right)\right]\right| &= \left|\mathbb{E}\left[F_{s,t}\left(Y,\pi\right)\right]-\mathbb{E}\left[\left(W_t^N-W_s^N\right)\prod_{k=1}^{N}\psi_k\left(Z_{s_k}^N\right)\right]\right| \\
&\leq \left|\mathbb{E}\left[F_{s,t}\left(Y,\pi\right)\left(1-\xi_M\left(\sup_{0\leq r\leq t}|Y_r|\right)\right)\right]\right| \qquad (2.31)\\
&\quad+\left|\mathbb{E}\left[F_{s,t}\left(Y,\pi\right)\xi_M\left(\sup_{0\leq r\leq t}|Y_r|\right)\right]-\mathbb{E}\left[F_{s,t}\left(X^N,\pi_1\right)\xi_M\left(\sup_{0\leq r\leq t}\left|X_r^N\right|\right)\right]\right| \qquad (2.32)\\
&\quad+\left|\mathbb{E}\left[F_{s,t}\left(X^N,\pi_1\right)\left(1-\xi_M\left(\sup_{0\leq r\leq t}\left|X_r^N\right|\right)\right)\right]\right|. \qquad (2.33)
\end{aligned}
$$

Using the fact that $1-\xi_M(x)\leq \mathbb{1}_{\{|x|>M\}}$, Cauchy-Schwarz's inequality, Markov's inequality and Lemma 2.2.5, we can bound (2.31) and (2.33) by $C/\sqrt{M}$ for some $C>0$ that is independent of N.

Now, fix some $\varepsilon>0$ and consider a constant $M_\varepsilon>0$ such that (2.31) and (2.33) are smaller than ε. In a next step, we choose an integer N_ε big enough such that (2.32) is smaller than ε. As a consequence, $|\mathbb{E}\left[F_{s,t}(Y,\pi)\right]|\leq 3\varepsilon$ for all $\varepsilon>0$, whence $\mathbb{E}\left[F_{s,t}(Y,\pi)\right]=0$ which means that for all $\varphi\in C_b^3(\mathbb{R})$, the expression (2.29) is a $(\mathcal{F}_t^0)_t-$martingale.

In the following we need to prove that for all $\varphi\in C^3$ (not necessarily bounded), the expression (2.29) is a $(\mathcal{F}_t^0)_t-$local martingale. So we introduce the stopping times $\tau_K=\inf\{t>0:|Y_t|>K\}$, and for $\varphi\in C^3(\mathbb{R})$, we define $\varphi_K\in C_c^3(\mathbb{R})$ by $\varphi_K(x)=\varphi(x)\xi_K(x)$. Now if $F_{s,t}^{\varphi}$ denotes the function $F_{s,t}$ we used previously, by definition of F,τ_K and φ_K, we know that $\mathbb{E}\left[F^{\varphi}_{s\wedge\tau_K,t\wedge\tau_K}(Y,\pi)\right]=\mathbb{E}\left[F^{\varphi_K}_{s\wedge\tau_K,t\wedge\tau_K}(Y,\pi)\right]$ which equals 0, since the expression (2.29) with $\varphi_K\in C_b^3(\mathbb{R})$ is a martingale.

Hence we have shown that the expression in (2.29) is a $(\mathcal{F}_t^0)_t-$martingale if $\varphi\in C_b^3(\mathbb{R})$, and that it is a $(\mathcal{F}_t^0)_t-$local martingale if $\varphi\in C^3(\mathbb{R})$. By right-continuity of $s\mapsto Y_s$, this implies that the expression in (2.29) is martingale (resp. local martingale) with respect to $(\mathcal{F}_t')_t$ for $\varphi\in C_b^3(\mathbb{R})$ (resp. $\varphi\in C^3(\mathbb{R})$).

Step 3. Now we show that Y and π are independent. By Theorem II.2.42 of Jacod and Shiryaev (2003), *Step 2* implies that Y is a $(P',({\mathcal{F}_t}')_{t\geq 0})-$semi-martingale with characteristics $B_t=-\alpha\int_0^t Y_s ds$, $\nu(ds,dx)=0$, $C_t=\int_0^t\sigma^2 f(Y_s)ds$. Moreover, Theorem III.2.26 of Jacod and Shiryaev (2003) implies that there exists a Brownian motion B' defined on $(\Omega',\mathcal{A}',(\mathcal{G}_t)_{t\geq 0},P')$, such that Y is solution of

$$
Y_t=Y_0-\alpha\int_0^t Y_s ds+\sigma\int_0^t\sqrt{f(Y_s)}dB'_s.
$$

So B' is defined on the same space, but for the moment we do not know that this Brownian motion is indeed a Brownian with respect to the filtration we are interested in, that is, with respect to $({\mathcal{F}_t}')_{t\geq 0}$. To understand this last point we use the Lamperti transform. To do so, we need to introduce

$$
h(x):=\int_0^x\frac{1}{\sqrt{f(t)}}dt.
$$

Note that the function h above is well-defined since the function f is positive and continuous.

Using Ito's formula, one gets that $\tilde{Y}_t := h(Y_t)$ solves

$$d\tilde{Y}_t = -\alpha h'(Y_t)Y_t dt + \sigma h'(Y_t)\sqrt{f(Y_t)}dB'_t + \frac{\sigma^2}{2}h''(Y_t)f(Y_t)dt.$$

In other words,

$$\sigma B'_t = h(Y_t) - h(Y_0) + \alpha\int_0^t h'(Y_s)Y_s ds - \frac{\sigma^2}{2}\int_0^t h''(Y_s)f(Y_s)ds$$

is exactly of the form as in (2.29), for the test-function $\varphi = h$ that is C^3. Thus we know that $(B'_t)_t$ is a $(P', (\mathcal{F}_t{}')_{t\geq 0})-$local martingale.

By Theorem II.6.3 of Ikeda and Watanabe (1989) we can then conclude that B' and the Poisson random measure π - which are defined with respect to the same filtration, living on the same space - are independent, and thus also Y and π. □

Proof of Theorem 2.4.1. We know that, for any $k \geq 1$, $(f \circ X^N, \pi^1, ..., \pi^k)$ converges to $(f \circ \bar{X}, \pi^1, ..., \pi^k)$ in distribution in $D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^k$ where at the limit, $\bar{X}$ is independent of the Poisson measures (indeed Theorem 2.2.1 gives the convergence of X^N to $\bar{X}$, Lemma 2.4.3 allows to deduce the convergence of $f \circ X^N$ to $f \circ \bar{X}$, and Proposition 2.4.4 proves the independance between $\bar{X}$ and the Poisson measures). Then, Theorem B.0.1 implies the result. □

An alternative proof of Theorem 2.4.1 consists in applying Theorem IX.4.15 of Jacod and Shiryaev (2003) to the semimartingale $(X^N, Z^{N,1}, ..., Z^{N,k})$. This theorem gives in particular the convergence of $(Z^{N,1}, ..., Z^{N,k})$ to $(\bar{Z}^1, ..., \bar{Z}^k)$ in $D(\mathbb{R}_+, \mathbb{R}^k)$ for any $k \in \mathbb{N}^*$. Thus, it implies the convergence in $D(\mathbb{R}_+, \mathbb{R})^k$ for any $k \in \mathbb{N}^*$, whence the convergence in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$. This proof is more classical and easier to apply than the previous one. However we chosed to show the first proof because it can be used to case where the stochastic intensities are not semimartingales.

Chapter 3

Generalized Erlang kernel in a multipopulation frame

In this chapter, let us generalize the previous model in two directions: we consider a more general convolution kernel of the form of a polynomial function multiplied by an exponential factor, and we consider a multipopulation framework in a similar way as in Ditlevsen and Löcherbach (2017).

More precisely, we consider a finite number $d \in \mathbb{N}^*$ of populations, where each population is a meanfield system, and the interactions between the particles are interaction between the populations: for each $1 \leq k \leq d$, let $I(k)$ be the set of the populations $1 \leq l \leq k$ that send spikes to the neurons of the population k. Let us note N_k $(1 \leq k \leq d)$ the number of particles in the population k. As before, let $N = \sum_{i=1}^d N_i$ the total number of particles. Let h_k $(1 \leq k \leq d)$ be the convolution kernel of the population k:

$$h_k(t) := e^{-\alpha_k t} \sum_{i=0}^{n_k} a_{k,i} \frac{t^i}{i!}, \tag{3.1}$$

where $a_{k,i} \in \mathbb{R}$ and $\alpha_k > 0$ $(1 \leq k \leq d, 0 \leq i \leq n_k)$. The convolution kernel (3.1) is a generalization of the so-called Erlang kernel. This kind of kernel has been studied in Duarte et al. (2018).

We consider the following system of Hawkes processes:

$$\begin{cases} Z_t^{N,k,i} = & \displaystyle\int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z \leq f_k(X_{s-}^{N,k})\right\}} d\pi^{k,i}(s,z,u), \quad 1 \leq k \leq d, 1 \leq i \leq N_k, \\ X_t^{N,k} = & e^{-\alpha_k t} a_{k,0} X_0^{N,k} \\ & \displaystyle + \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \cdot h_k(t-s) \mathbb{1}_{\left\{z \leq f_l(X_{s-}^{N,l})\right\}} d\pi^{l,i}(s,z,u), \quad 1 \leq k \leq d, \end{cases} \tag{3.2}$$

where $\pi^{k,i}$ $(1 \leq k \leq d, i \geq 1)$ are independent Poisson measures on $\mathbb{R}_+^2 \times \mathbb{R}$ with respective intensities $dt \cdot dz \cdot d\mu_k(u)$, with μ^k $(1 \leq k \leq d)$ are centred probability measures on $\mathbb{R}$.

This system seems more complicated than the previous one with exponential kernels, because here, the process $(X^{N,k})_{1\leq k \leq d}$ is not Markovian nor a semimartingale. But we show that we can reason in a similar way as in the previous chapter since $(X^{N,k})_{1 \leq k \leq d}$ is a coordinate of a Markovian semimartingale.

Let us introduce, for $1 \leq k \leq d, 0 \leq r \leq n_k$,

$$
\begin{aligned}
Y_t^{N,k,r} :=& e^{-\alpha_k t} \sum_{j=r}^{n_k} a_{k,j} \frac{t^{j-r}}{(j-r)!} Y_0^{N,k,j} \\
&+ \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{[0,t] \times \mathbb{R}_+ \times \mathbb{R}} u \cdot \sum_{j=r}^{n_k} a_{k,j} \frac{(t-s)^{j-r}}{(j-r)!} e^{-\alpha_k (t-s)} \mathbb{1}_{\left\{z \leq f_l(X_{s-}^{N,l})\right\}} d\pi^{l,i}(s,z,u).
\end{aligned}
$$

One can note that, for all $1 \leq k \leq d$, $Y^{N,k,0} = X^{N,k}$ (provided that $(Y_0^{N,k,j})_{0 \leq j \leq n_k} = (X_0^{N,k}, 0, ..., 0))$, and that

Lemma 3.0.1.

$$
\begin{aligned}
dY_t^{N,k,r} =& -\alpha_k Y_t^{N,k,r} dt + Y_t^{N,k,r+1} dt \qquad (3.3) \\
&+ a_{k,r} \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{\mathbb{R}_+ \times \mathbb{R}} u \cdot \mathbb{1}_{\left\{z \leq f_l(Y_{s-}^{N,l,0})\right\}} d\pi^{l,i}(t,z,u), \quad 0 \leq r < n_k, \\
dY_t^{N,k,n_k} =& -\alpha_k Y_t^{N,k,n_k} dt \qquad (3.4) \\
&+ a_{k,n_k} \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{\mathbb{R}_+ \times \mathbb{R}} u \cdot \mathbb{1}_{\left\{z \leq f_l(Y_{s-}^{N,l,0})\right\}} d\pi^{l,i}(t,z,u).
\end{aligned}
$$

Proof. Let us just prove the case $r < n_k$. Note

$$
\tilde{Y}_t^{N,k,r} := e^{\alpha_k t} Y_t^{N,k,r},
$$

and remark that

$$
\begin{aligned}
\tilde{Y}_t^{N,k,r} =& \sum_{j=r}^{n_k} a_{k,j} \frac{t^{j-r}}{(j-r)!} Y_0^{N,k,j} \\
&+ \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{[0,t] \times \mathbb{R}_+ \times \mathbb{R}} u \sum_{j=r}^{n_k} a_{k,j} \frac{(t-s)^{j-r}}{(j-r)!} e^{\alpha s} \mathbb{1}_{\left\{z \leq f_l(Y_{s-}^{N,l,0})\right\}} d\pi^{l,i}(s,z,u).
\end{aligned}
$$

As a consequence

$$
d\tilde{Y}_t^{N,k,r} = \tilde{Y}^{N,k,r+1} dt + a_{k,r} e^{\alpha_k t} \sum_{l \in I(k)} \frac{1}{\sqrt{N_l}} \sum_{i=1}^{N_l} \int_{\mathbb{R}_+ \times \mathbb{R}} u \cdot \mathbb{1}_{\left\{z \leq f_l(Y_{t-}^{N,l,0})\right\}} d\pi^{l,i}(t,z,u).
$$

On the other hand, the integration by parts formula gives

$$
d\tilde{Y}_t^{N,k,r} = \alpha_k e^{\alpha_k t} Y_t^{N,k,r} dt + e^{\alpha_k t} dY_t^{N,k,r}.
$$

□

The limit system is given by the following stochastic differential equations

$$\begin{aligned} d\bar{Y}_t^{k,r} =& -\alpha_k \bar{Y}_t^{k,r} dt + \bar{Y}^{k,r+1} dt + a_{k,r} \sum_{l\in I(k)} \sigma_l \sqrt{f_l(\bar{Y}_t^{l,0})} dB_t^l, \quad 0 \le r < n_k, \\ d\bar{Y}_t^{k,n_k} =& -\alpha_k \bar{Y}_t^{k,n_k} dt + a_{k,n_k} \sum_{l\in I(k)} \sigma_l \sqrt{f_l(\bar{Y}_t^{l,0})} dB_t^l, \end{aligned} \tag{3.5}$$

where $(B^l)_{1\le l\le d}$ is a $d-$dimensional Brownian motion and σ_l^2 is the variance of μ^l. Consistently to the previous notation, we note $\bar{X}^k := \bar{Y}^{k,0}$.

As these equations are similar to (2.2), we prove similar results as in Chapter 2 using very similar arguments. As a consequence, we juste write the main steps of the proofs. The only result that we omit in this chapter compared to the previous one is the convergence of the Markovian kernel of the $N-$particle system to the invariant measure of the limit process (i.e. the result of Section 2.3). This kind of question about invariant measures is known to be more complicated in a multi-dimensional framework, and we do not investigate it in this thesis.

3.1 Assumptions

In this model, we need to assume that each population satisfies the assumptions of Section 2.1. We also need another assumption stating that, in the limit system, each population survives.

Assumption 3.1. *For each $1 \le k \le N$, we assume that N_k/N converges as N goes to infinity to some $\lambda_k \in]0,1[$.*

Then, as in the previous model, we need an assumption on the functions f_k $(1 \le k \le d)$ to guarantee the well-posedness of the stochastic differential equations.

Assumption 3.2. *for all $1 \le k \le d$, $\sqrt{f_k}$ is a positive and Lipschitz continuous function, having Lipschitz constant L_k.*

We also need to control the moments of the processes $\left(Y_t^{N,k,r}\right)_t$ and $\left(\bar{Y}_t^{k,r}\right)_t$, and to assume that the measures μ_k $(1 \le k \le d)$ are centered because of the normalization in $N^{-1/2}$.

Assumption 3.3. *For each $1 \le k \le d$,*

- $\int_{\mathbb{R}} x^4 d\bar{\nu}_{k,0}(x) < \infty$ *and* $\sup_{N\in\mathbb{N}^*} \int_{\mathbb{R}} x^4 d\nu_{k,0}^N(x) < \infty$.
- *μ_k is a centered probability measure having a fourth moment, we note σ_k^2 its variance.*

Assumption 3.4. *We assume that the functions f_k $(1 \le k \le d)$ are C^4 and for each $1 \le \beta \le 4$, $\left(\sqrt{f}\right)^{(\beta)}$ is bounded.*

Assumption 3.4 guarantees that the stochastic flow associated to (3.5) has regularity properties with respect to the initial condition $\bar{X}_0 = x$.

Assumption 3.5. *$X_0^{N,k}$ converges in distribution to $\bar{X}_0^k$.*

An adaptation of the proof of Proposition 2.1.2 allows to prove

Proposition 3.1.1. *If Assumptions 3.2 and 3.3 hold, the equations (3.3) and (3.4) admit unique non-exploding strong solutions.*

3.2 Convergence of $(Y^{N,k,r})_N$ in distribution

The goal of this section is to prove two results: the convergence of X^N to $\bar{X}$ and the rate of this convergence.

3.2.1 The convergence of $Y^{N,k,r}$

As in the previous chapter, the convergence of $(Y^{N,k,r})_{\substack{1\leq k\leq d\\1\leq r\leq n_k}}$ as N goes to infinity can be proven using convergence martingale theorem (Theorem $IX.4.15$ of Jacod and Shiryaev (2003)).

Theorem 3.2.1. *Under Assumptions 3.1, 3.2, 3.3 and 3.5, $(Y^{N,k,r})_{\substack{1\leq k\leq d\\1\leq r\leq n_k}}$ converges to $(\bar{Y}^{k,r})_{\substack{1\leq k\leq d\\1\leq r\leq n_k}}$ in distribution in $D(\mathbb{R}_+,\mathbb{R}^{\prod_{k=1}^d n_k})$.*

Proof. Using the notation of Theorem $IX.4.15$ of Jacod and Shiryaev (2003), with the truncation function $h = Id$, we have

$$\begin{aligned}
b'^{N,k,r}(y) =& -\alpha_k y^{k,r} + y^{k,r+1}, \quad r < n_k,\\
b'^{N,k,n_k} =& -\alpha_k y^{k,r},\\
\int g(v)K^N(y,dv) =& \sum_{l=1}^{d} N_l f_l(y^{l,0}) \int_{\mathbb{R}} g\left(\frac{1}{\sqrt{N_l}} \sum_{k\in I^{-1}(l)} \sum_{r=0}^{n_k} a_{k,r} u\cdot e_{k,r}\right) d\mu_l(u),\\
\tilde{c}'^{N,(k,r),(k',r')} =& \int v^{k,r}v^{k',r'}K^N(y,dv) = \sum_{l\in I(k)\cap I(k')} N_l f_l(y^{l,0}) \int_{\mathbb{R}} \frac{1}{\sqrt{N_l}^2} a_{k,r}a_{k',r'}u^2 d\mu_l(u)\\
=& a_{k,r}a_{k',r'} \sum_{l\in I(k)\cap I(k')} f_l(y^{l,0})\sigma_l^2,
\end{aligned}$$

where $e_{k,r}$ denotes the unit vector related to the coordinate (k,r).

The function $b'^{N,k,r}$ and $\tilde{c}'^{N,(k,r),(k',r')}$ do not depend on N. And for any continuous and bounded function g null around zero, $\int g(y)K^N(\cdot,dy)$ vanishes as N goes to infinity. So Theorem $IX.4.15$ of Jacod and Shiryaev (2003) implies the result. ☐

Theorem 3.2.2. *Under Assumptions 3.1, 3.2, 3.3 and 3.5, $(X^{N,k})_{1\leq k\leq d}$ converges to $(\bar{X}^k)_{1\leq k\leq d}$ in distribution in $D(\mathbb{R}_+,\mathbb{R}^d)$.*

Proof. Recalling that $X^{N,k} = Y^{N,k,0}$ and $\bar{X}^k = \bar{Y}^{k,0}$, the results is a straightforward consequence of Theorem 3.2.1. ☐

3.2.2 Rate of convergence of $Y^{N,k,r}$

In this section we provide a rate of convergence for Theorem 3.2.1.

Theorem 3.2.3. *Under Assumptions 3.2, 3.3 and 3.4, for all $T\geq 0$, for each $g\in C_b^3(\mathbb{R}^{\prod_{k=1}^d n_k})$ and $y\in\mathbb{R}^{\prod_{k=1}^d n_k}$,*

$$\sup_{0\leq t\leq T} \left|P_t^N g(y) - \bar{P}_t g(y)\right| \leq C_t\left(1+||y||_2^2\right)||g||_{3,\infty} \sum_{l=1}^{d}\frac{1}{\sqrt{N_l}}.$$

To begin with, we can prove bounds for the processes $Y^{N,k,r}$ and $\bar{Y}^{k,r}$ in the exact same way as in Lemma 2.2.5. Hence, we omit the proof.

Lemma 3.2.4. *Under Assumptions 3.2 and 3.3, the following holds.*

(i) For all $t > 0$ and $y \in \mathbb{R}^{\prod_{k=1}^d n_k}$, $\mathbb{E}_y\left[\left|\left|Y_t^N\right|\right|_2^2\right] \leq (1 + ||y||_2^2)C_T$, for some $C_T > 0$ independent of N and y.

(ii) For all $t > 0$ and $y \in \mathbb{R}^{\prod_{k=1}^d n_k}$, $\mathbb{E}_y\left[\left|\left|\bar{Y}_t\right|\right|_2^2\right] \leq (1 + ||y||_2^2)C_T$, for some $C_T > 0$ independent of N and y.

(iii) For all $N \in \mathbb{N}^*, T > 0$, $\mathbb{E}\left[\sup_{0\leq t\leq T} ||Y_t^N||_2^2\right] < +\infty$ and $\mathbb{E}\left[\sup_{0\leq t\leq T} ||\bar{Y}_t||_2^2\right] < +\infty$.

(iv) For all $T > 0, N \in \mathbb{N}^*$, $\sup_{0\leq t\leq T} \mathbb{E}_y\left[||Y_t^N||_4^4\right] \leq C_T(1 + y^4)$ and $\sup_{0\leq t\leq T} \mathbb{E}_y\left[||\bar{Y}_t||_4^4\right] \leq C_T(1 + y^4)$.

We begin by establishing a rate of convergence for the (extended) generators of the processes. Let us note A^N the generator of Y^N, and $\bar{A}$ that of $\bar{Y}$.

Lemma 3.2.5. *$C_b^2(\mathbb{R}^{\prod_{k=1}^d n_k}) \subseteq \mathcal{D}'\left(\bar{A}\right)$, and for all $g \in C_b^2(\mathbb{R}^{\prod_{k=1}^d n_k})$ and $y \in \mathbb{R}^{\prod_{k=1}^d n_k}$, we have*

$$\bar{A}g(y) = -\sum_{k=1}^d \alpha_k \sum_{r=0}^{n_k} y^{k,r}\partial_{(k,r)}g(y) + \sum_{k=1}^d \sum_{r=0}^{n_k-1} y^{k,r+1}\partial_{(k,r)}g(y)$$
$$+ \sum_{(k,r),(k',r')} a_{k,r}a_{k',r'}\left(\sum_{l\in I(k)\cap I(k')} \sigma_l^2 f_l(y^{l,0})\right)\partial^2_{(k,r),(k',r')}g(y).$$

Moreover, $C_b^1(\mathbb{R}^{\prod_{k=1}^d n_k}) \subseteq \mathcal{D}'\left(A^N\right)$, and for all $g \in C_b^1(\mathbb{R}^{\prod_{k=1}^d n_k})$ and $y \in \mathbb{R}^{\prod_{k=1}^d n_k}$, we have

$$A^N g(y) = -\sum_{k=1}^d \alpha_k \sum_{r=0}^{n_k} y^{k,r}\partial_{(k,r)}g(y) + \sum_{k=1}^d \sum_{r=0}^{n_k-1} y^{k,r+1}\partial_{(k,r)}g(y)$$
$$+ \sum_{l=1}^d N_l f_l(y^{l,0}) \int_{\mathbb{R}} \left[g\left(y + \frac{1}{\sqrt{N_l}} \sum_{k\in I^{-1}(k)} \sum_{r=0}^{n_k} a_{k,r}u \cdot e_{k,r}\right) - g(y)\right] d\mu_l(u).$$

Now we can give the convergence speed of the generators, using once again Taylor-Lagrange's inequality.

Proposition 3.2.6. *If Assumptions 3.2 and 3.3 hold, then for all $g \in C_b^3(\mathbb{R}^{\prod_{k=1}^d n_k})$,*

$$\left|\bar{A}g(y) - A^N g(y)\right| \leq C\,||g||_{3,\infty} \sum_{l=1}^d \frac{1}{\sqrt{N_l}} f_l(y^{l,0}) \int_{\mathbb{R}} |u|^3 d\mu_l(u),$$

where the constant C depends on the functions h_k $(1 \leq k \leq d)$ defined in (3.1).

As in the previous model, we obtain a convergence speed for the semigroups from that of the generators using Proposition A.0.3, stating that

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds, \tag{3.6}$$

under suitable assumptions on the processes.

In order to apply Proposition A.0.3 in this model, the only result whose proof is not the same as in the previous model (i.e. the model of Chapter 2) is the regularity of the limit semigroup (i.e. Proposition 2.2.7).

Proposition 3.2.7. *If Assumptions 3.2, 3.3 and 3.4 hold, then for all $t \geq 0$ and for all $g \in C_b^3(\mathbb{R}^{\prod_{k=1}^d n_k})$, the function $y \mapsto \bar{P}_t g(y)$ belongs to $C_b^3(\mathbb{R}^{\prod_{k=1}^d n_k})$ and satisfies*

$$\sup_{0 \leq t \leq T} \left|\left| \left(\bar{P}_t g\right)''' \right|\right|_\infty \leq Q_T C ||g||_{3,\infty},$$

with Q_T some constant that depends on T.

Remark 3.2.8. *Contrarily to Proposition 2.2.7, in Proposition 3.2.7, we do not provide an explicit expression for Q_T, because the computations are far more tedious.*

Proof. The idea of the proof is very similar to the proof of Proposition 2.2.7. The main part of the proof consists in obtaining controls on the moments of the derivatives of the stochastic flow of Y^N, to control the derivatives of its semigroup. These controls can be obtained in classical way applying Ito's formula and Grönwall's Lemma. □

Now we can compute a rate of convergence of the semigroups.

Proof of Theorem 3.2.3. We can prove with the same arguments as in the previous chapter that the semigroups P^N and $\bar{P}$ satisfy the hypothesis of Proposition A.0.3. Then, thanks to (3.6), we have

$$\begin{aligned}
\left|\bar{P}_t g(y) - P_t^N g(y)\right| =& \left| \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(y) ds \right| \\
\leq& \int_0^t \mathbb{E}_y^N \left[\left| \bar{A} \left(\bar{P}_s g\right) \left(Y_{t-s}^N\right) - A^N \left(\bar{P}_s g\right) \left(Y_{t-s}^N\right) \right| \right] ds \\
\leq& C_t \left(1 + ||y||_2^2\right) \sum_{l=1}^d \frac{1}{\sqrt{N_l}},
\end{aligned}$$

which proves the result. □

Chapter 4

Model with reset jumps : Hawkes processes with variable length memory

This chapter is based on Erny et al. (2020).

In this chapter, we study a particle system close to the one studied in Chapter 2. This system is designed to model neural systems. More precisely, it models the biological fact that, right after a neuron emits a spike, its membrane potential is reset to zero, interpreted as resting potential.

More precisely, we consider, for each $N \geq 1$, a family of i.i.d. Poisson measures $(\pi^i(ds, dz, du))_{i=1,\ldots,N}$ on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ having intensity measure $ds \cdot dz \cdot \nu(du)$ where ν is a centred probability measure on $\mathbb{R}$, as well as an i.i.d. family $(X_0^{N,i})_{i=1,\ldots,N}$ of $\mathbb{R}$-valued random variables independent of the Poisson measures. The object of this chapter is to study the convergence of the Markov process $X_t^N = (X_t^{N,1}, \ldots, X_t^{N,N})$ taking values in $\mathbb{R}^N$ and solving, for $i = 1, \ldots, N$, for $t \geq 0$,

$$\begin{cases} X_t^{N,i} &= X_0^{N,i} - \alpha \int_0^t X_s^{N,i} ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} X_{s-}^{N,i} \mathbf{1}_{\{z \leq f(X_{s-}^{N,i})\}} \pi^i(ds, dz, du) \\ & \quad + \frac{1}{\sqrt{N}} \sum_{j \neq i} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{s-}^{N,j})\}} \pi^j(ds, dz, du), \\ X_0^{N,i} &\sim \nu_0. \end{cases} \tag{4.1}$$

In these equations, there is an underlying system of Hawkes processes $(Z^{N,i})_{1 \leq i \leq N}$ that can be defined as

$$Z_t^{N,i} = \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z \leq f(X_{s-}^{N,i})\right\}} d\pi^i(s, z, u).$$

Indeed, the system $(X^{N,i})_{1 \leq i \leq N}$ satisfies the following Hawkes-type dynamics

$$X_t^{N,i} = \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]L_t^i, t]\times\mathbb{R}_+\times\mathbb{R}} e^{-\alpha(t-s)} u \mathbb{1}_{\left\{z \leq f(X_{s-}^{N,j})\right\}} d\pi^j(s, z, u) + e^{-\alpha t} X_0^{N,i} \mathbf{1}_{L_t^i = 0},$$

where $L_t^i = \sup\{0 \le s \le t : \Delta Z_s^{N,i} = 1\}$ is the last spiking time of neuron i before time t, with the convention $\sup \emptyset := 0$. Thus the integral over the past, starting from 0 in classical Hawkes-type dynamics, is replaced by an integral starting at the last jump time before the present time. Such processes are termed being of *variable length memory*.

The goal of this chapter is to study the convergence of $(X^{N,i})_{1\le i\le N}$ as N goes to infinity, in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$ endowed with the product topology, where $D(\mathbb{R}_+,\mathbb{R})$ is endowed with Skorohod topology. The associated limit system (given in (4.2) below) is not classical and not obvious. The next section is dedicated to explain informally how the limit particle system should a priori look like. Let us note $(\bar{X}^i)_{i\ge 1}$ the limit system of $(X^{N,i})_{1\le i\le N}$ as N goes to infinity.

4.1 Heuristics for the limit system

Let us remark that in (4.1), the only term that does depend on N is the jump term of the second line which is approximately given by

$$M_t^N = \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbb{1}_{\left\{z\le f(X_{s-}^{N,j})\right\}} d\pi^j(s,z,u).$$

Then in the infinite neuron model, each process $\bar{X}^i$ should solve the equation (4.1), where the term of the second line is replaced by $M_t := \lim_{N\to\infty} M_t^N$. Because of the scaling in $N^{-1/2}$, the limit martingale M_t will be a stochastic integral with respect to some Brownian motion, and its variance the limit of

$$\langle M_t^N \rangle^2 = \sigma^2 \int_0^t \frac{1}{N}\sum_{j=1}^N f(X_s^{N,j}) ds,$$

where σ^2 is the variance of ν. Therefore, the limit martingale (if it exists) must be of the form

$$M_t = \sigma \int_0^t \sqrt{\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^N f\left(\bar{X}_s^j\right)}\, dW_s = \sigma \int_0^t \sqrt{\lim_{N\to\infty} \bar{\mu}_s^N(f)} dW_s,$$

where $\bar{\mu}_s^N$ is the empirical measure of the system $\left(\bar{X}_s^j\right)_{1\le j\le N}$ and W is a one-dimensional standard Brownian motion.

Then, admitting that the sequence of empirical measures $(\bar{\mu}^N)_N$ converges to some measure μ, the limit system $(\bar{X}^i)_{i\ge 1}$ should satisfy

$$\begin{cases} \bar{X}_t^i &= \bar{X}_0^i - \alpha \int_0^t \bar{X}_s^i ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-}^i \mathbf{1}_{\{z\le f(\bar{X}_{s-}^i)\}} \pi^i(ds,dz,du) \\ & \quad + \int_0^t \sigma\sqrt{\mu_s(f)} dW_s, \\ \bar{X}_0^i &\sim \nu_0. \end{cases}$$

As $\bar{\mu}^N$ $(N\in\mathbb{N}^*)$ are empirical measures of infinite exchangeable systems, their limit μ, if it exists, is necessarily the directing measure of the exchangeable system $(\bar{X}^i)_{i\ge 1}$ (see Proposition (7.20) of Aldous (1983)). Moreover, as conditionally on W, the variables $\bar{X}^i$ $(i\ge 1)$ are i.i.d., Lemma (2.12) of Aldous (1983) implies that $\mu = \mathcal{L}(\bar{X}^1|W) = \mathcal{L}(\bar{X}^i|W)$.

As a consequence, the limit system $(\bar{X}^i)_{i\geq 1}$ should satisfy

$$\begin{cases} \bar{X}_t^i &= \bar{X}_0^i - \alpha \int_0^t \bar{X}_s^i ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-}^i \mathbf{1}_{\{z\leq f(\bar{X}_{s-}^i)\}} \pi^i(ds,dz,du) \\ & \quad +\sigma \int_0^t \sqrt{\mathbb{E}\left[f\left(\bar{X}_s^i\right) \middle| \mathcal{W}\right]} dW_s, \\ \bar{X}_0^i &\sim \nu_0. \end{cases} \tag{4.2}$$

In the above equation, $(W_t)_{t\geq 0}$ is a standard one-dimensional Brownian motion which is independent of the Poisson random measures, and $\mathcal{W} = \sigma\{W_t, t \geq 0\}$. Moreover, the initial positions $\bar{X}_0^i, i \geq 1$, are i.i.d., independent of W and of the Poisson random measures, distributed according to ν_0 which is the same probability measure as in (4.1). The common jumps of the particles in the finite system, due to their scaling in $1/\sqrt{N}$ and the fact that they are centred, by the Central Limit Theorem, create this single Brownian motion W_t which is underlying each particle's motion and which induces the common noise factor for all particles in the limit.

The limit equation (4.2) is not clearly well-posed and requires more conditions on the rate function f. Let us briefly comment on the type of difficulties that one encounters when proving trajectorial uniqueness of (4.2). Roughly speaking, the jump terms demand to work in an L^1-framework, while the diffusive terms demand to work in an L^2-framework. Graham (1992) proposes a unified approach to deal both with jump and with diffusion terms in a non-linear framework, and we shall rely on his ideas in the sequel. The presence of the random volatility term which involves conditional expectation causes however additional technical difficulties. Finally, another difficulty comes from the fact that the jumps induce non-Lipschitz terms of the form $\bar{X}_s^i f(\bar{X}_s^i)$. For this reason a classical Wasserstein-1$-$coupling is not appropriate for the jump terms. Therefore we propose a different distance which is inspired by the one already used in Fournier and Löcherbach (2016).

To simplify the notation, we just prove the well-posedness of each coordinate of the system (4.2). That is

$$\begin{cases} \bar{X}_t &= \bar{X}_0 - \alpha \int_0^t \bar{X}_s ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-} \mathbf{1}_{\{z\leq f(\bar{X}_{s-})\}} \pi(ds,dz,du) \\ & \quad +\sigma \int_0^t \sqrt{\mathbb{E}\left[f\left(\bar{X}_s\right) \middle| \mathcal{W}\right]} dW_s, \\ \bar{X}_0 &\sim \nu_0. \end{cases} \tag{4.3}$$

4.2 Assumptions on the model

In order to guarantee existence and uniqueness of a strong solution of (4.1), we introduce the following hypothesis.

Assumption 4.1. *The function f is Lipschitz continuous.*

In addition, we also need the following condition to obtain a priori bounds on some moments of the process $\left(X^{N,i}\right)_{1\leq i\leq N}$.

Assumption 4.2. *We assume that $\int_{\mathbb{R}} x d\nu(x) = 0$, $\int_{\mathbb{R}} x^2 d\nu(x) < +\infty$, and $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$.*

Under Assumptions 4.1 and 4.2, existence and uniqueness of strong solutions of (4.1) follow from Theorem IV.9.1 of Ikeda and Watanabe (1989), exactly in the same way as in Proposition 2.1.2.

The limit system (4.2) being less classical than (4.1), we need less standard Assumptions to prove its well-posedness.

Assumption 4.3. *1. We suppose that* $\inf f > 0$.
2. There exists a function $a \in C^2(\mathbb{R}, \mathbb{R}_+)$, *strictly increasing and bounded, such that, for a suitable constant* C, *for all* $x, y \in \mathbb{R}$,

$$|a''(x) - a''(y)| + |a'(x) - a'(y)| + |xa'(x) - ya'(y)| + |f(x) - f(y)| \leq C|a(x) - a(y)|.$$

Note that Assumption 4.3 implies Assumption 4.1 as well as the boundedness of the rate function f. Let us give some examples of functions f that satisfy Assumption 4.3.

Example 4.2.1. *Any positive bounded and lowerbounded function* $f \in C_b^1(\mathbb{R}, \mathbb{R}_+)$ *with*

$$|f'(x)| \leq \frac{C}{(1 + |x|)^{1+\varepsilon}} \tag{4.4}$$

for all $x \in \mathbb{R}$, *where* C *and* ε *are some positive constants, satisfies Assumption 4.3 with*

$$a(x) = \int_{-\infty}^{x} \frac{dy}{(1 + \psi(y))^{1+\varepsilon}},$$

where ψ *is any smooth non-negative function satisfying* $\psi(y) = |y|$ *for* $|y| \geq 1$. *If* (4.4) *holds with* $\varepsilon = 1$, *then we may choose simply* $a(x) = \arctan(x) + \pi/2$.

Finally, fix some $-\infty < a < b < \infty$. *Then any function* $f \in C_b^1(\mathbb{R}, \mathbb{R}_+)$ *which is constant below* a *and above* b *satisfies Assumption* (4.3).

Let us note that this kind of function is interesting from a neuroscience point of view, if it is in addition non decreasing. Indeed, when the potential of a neuron is below a *(resp. above* b*), its spiking rate is minimal (resp. maximal), such that the neuron can be considered as inactive (resp. active).*

In order to prove the convergence of the finite particle system to the limit system, we need to assume that the measure ν has a finite third moment.

Assumption 4.4.

$$\int_{\mathbb{R}} |u|^3 d\nu(u) < \infty.$$

4.3 Properties of the model

4.3.1 A priori estimates for $X^{N,i}$

In this subsection, we prove useful a priori upper bounds on some moments of the solutions of the stochastic differential equations (4.1).

Lemma 4.3.1. *Assume that 4.2 holds and that* f *is bounded:*

(i) for all $t > 0$, $\sup_{N \in \mathbb{N}^*} \sup_{0 \leq s \leq t} \mathbb{E}\left[\left(X_s^{N,1}\right)^2\right] < +\infty$,

(ii) for all $t > 0$, $\sup_{N\in\mathbb{N}^*} \mathbb{E}\left[\sup_{0\le s\le t} \left|X_s^{N,1}\right|\right] < +\infty$,

Proof. Step 1: Let us prove (i).

$$\begin{aligned}\left(X_t^{N,1}\right)^2 &= \left(X_0^{N,1}\right)^2 - 2\alpha \int_0^t \left(X_s^{N,1}\right)^2 ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left(X_s^{N,1}\right)^2 \mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,1}\right)\right\}} d\pi^j(s,z,u) \\ &\quad + \sum_{j=2}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left[\left(X_{s-}^{N,1} + \frac{u}{\sqrt{N}}\right)^2 - \left(X_{s-}^{N,1}\right)^2\right] \mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,j}\right)\right\}} d\pi^j(s,z,u) \\ &\le \left(X_0^{N,1}\right)^2 + \sum_{j=2}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left[\left(X_{s-}^{N,1} + \frac{u}{\sqrt{N}}\right)^2 - \left(X_{s-}^{N,1}\right)^2\right] \mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,j}\right)\right\}} d\pi^j(s,z,u).\end{aligned}$$

As f is bounded,

$$\mathbb{E}\left[\left(X_t^{N,1}\right)^2\right] \le \mathbb{E}\left[\left(X_0^{N,1}\right)^2\right] + \frac{\sigma^2}{N}\sum_{j=2}^N \int_0^t \mathbb{E}\left[f\left(X_s^{N,j}\right)\right] ds \le \mathbb{E}\left[\left(X_0^{N,1}\right)^2\right] + \sigma^2 ||f||_\infty t.$$

Step 2: Now we prove (ii).

$$\left|X_t^{N,1}\right| \le \left|X_0^{N,1}\right| + \alpha\int_0^t \left|X_s^{N,1}\right| ds + \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left|X_{s-}^{N,1}\right| \mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,1}\right)\right\}} d\pi^1(s,z,u) + \frac{1}{\sqrt{N}}|M_t^N|,$$

where M_t^N is the martingale $M_t^N = \sum_{j=2}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,j}\right)\right\}} d\pi^j(s,z,u)$. Then

$$\begin{aligned}\sup_{0\le s\le t} \left|X_s^{N,1}\right| \le \left|X_0^{N,1}\right| + \alpha\int_0^t |X_s^{N,1}| ds + \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left|X_{s-}^{N,1}\right| \mathbb{1}_{\left\{z\le f\left(X_{s-}^{N,1}\right)\right\}} d\pi^1(s,z,u) \\ + \frac{1}{\sqrt{N}} \sup_{0\le s\le t} |M_s^N|.\end{aligned}$$

To conclude the proof, it is now sufficient to notice that

$$\frac{1}{\sqrt{N}}\mathbb{E}\left[\sup_{0\le s\le t} |M_s^N|\right] \le \mathbb{E}\left[\frac{1}{N}[M^N]_t\right]^{1/2}$$

is uniformly bounded in N, since f is bounded, and to use the point (i) of the lemma. □

Let us end this section by proving some a priori estimates for the limit system.

Lemma 4.3.2. *Suppose that* f *is bounded and that* $\int_{\mathbb{R}} x^2 d\nu_0(x) < \infty$. *Then any solution* $(\bar{X}_t^1)_{t\ge 0}$ *of* (4.3) *a priori satisfies* $\mathbb{E}\left[\sup_{s\le t}\left(\bar{X}_s^1\right)^2\right] < \infty$ *for all* $t \ge 0$.

Proof. We first prove the weaker result

$$\sup_{s\le t} \mathbb{E}\left[\left(\bar{X}_s^1\right)^2\right] < \infty. \tag{4.5}$$

By Itô's formula,

$$
\begin{aligned}
(\bar X^1_t)^2 = (\bar X^1_0)^2 - 2\alpha \int_0^t (\bar X^1_s)^2 ds + 2\sigma \int_0^t \bar X^1_s \sqrt{\mu_s(f)} dW_s \\
+ \sigma^2 \int_0^t \mu_s(f) ds - \int_{[0,t]\times\mathbb{R}_+} (\bar X^1_{s-})^2 \mathbb{1}_{\{z \le f(\bar X^1_{s-})\}} d\pi(s,z) \\
\le (\bar X^1_0)^2 + 2\sigma \int_0^t \bar X^1_s \sqrt{\mu_s(f)} dW_s + \sigma^2 \int_0^t \mu_s(f) ds. \quad (4.6)
\end{aligned}
$$

Introducing, for any $M>0$, $\tau_M := \inf\{t>0 \ : \ |\bar X^1_t| > M\}$ and $u_M(t) := \mathbb{E}\left[(\bar X^1_{t\wedge\tau_M})^2\right]$, we have, for all $t \ge 0$,

$$
u_M(t) \le \mathbb{E}\left[(\bar X^1_0)^2\right] + \sigma^2 ||f||_\infty t.
$$

Then Grönwall's lemma implies that for all $T>0$,

$$
\sup_{M>0} \sup_{0\le t\le T} u_M(t) < \infty. \tag{4.7}
$$

This implies that the stopping times τ_M tend to infinity as M goes to infinity. Then (4.5) is a consequence of (4.7) and Fatou's lemma. Then, using Burkholder-Davis-Gundy inequality to control the martingale part in (4.6), we have, for all $t \ge 0$,

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0\le s\le t} (\bar X^1_s)^2\right] \le & \mathbb{E}\left[(\bar X^1_0)^2\right] + \sigma^2 ||f||_\infty t + 2\sigma ||f||_\infty \mathbb{E}\left[\left(\int_0^t (\bar X^1_s)^2 ds\right)^{1/2}\right] \\
\le & \mathbb{E}\left[(\bar X^1_0)^2\right] + \sigma^2 ||f||_\infty t + 2\sigma ||f||_\infty \left(1 + \int_0^t \mathbb{E}\left[(\bar X^1_s)^2\right] ds\right) ds.
\end{aligned}
$$

Then the result follows from point (4.5). □

4.3.2 Well-posedness of the limit system $(\bar X^i)_{i\ge1}$

In this section, we prove that the limit system is well-posed. More precisely, we prove the trajectorial uniqueness for the equations of the system (4.2), and the existence of a (unique) strong solution.

Theorem 4.3.3. *Grant Assumption 4.3.*
1. Pathwise uniqueness holds for the nonlinear stochastic differential equation (4.3).
2. If additionally, $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$, then there exists a unique strong solution $(\bar X_t)_{t\ge0}$ of the nonlinear stochastic differential equation (4.3).

Proof. Proof of Item 1. Consider two solutions $(\hat X_t)_{t\ge0}$ and $(\check X_t)_{t\ge0}$, $(\mathcal{F}_t)_t$−adapted, defined on the same probability space and driven by the same Poisson random measure π and the same Brownian motion W, and with $\hat X_0 = \check X_0$. We consider $Z_t := a(\hat X_t) - a(\check X_t)$. Denote $\hat\mu_s(f) = I\!E[f(\hat X_s)|\mathcal{W}_s]$ and $\check\mu_s(f) = I\!E[f(\check X_s)|\mathcal{W}_s]$.
Using Ito's formula, we can write

$$
Z_t = -\alpha \int_0^t \left(\hat X_s a'(\hat X_s) - \check X_s a'(\check X_s)\right) ds + \frac12 \int_0^t (a''(\hat X_s)\hat\mu_s(f) - a''(\check X_s)\check\mu_s(f))\sigma^2 ds
$$

$$
\begin{aligned}
&+ \int_0^t (a'(\hat{X}_s)\sqrt{\hat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})\sigma dW_s \\
&- \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(\hat{X}_{s-}) - a(\check{X}_{s-})]\mathbf{1}_{\{z\le f(\hat{X}_{s-})\wedge f(\check{X}_{s-})\}}\pi(ds,dz,du) \\
&+ \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(0) - a(\hat{X}_{s-})]\mathbf{1}_{\{f(\check{X}_{s-})<z\le f(\hat{X}_{s-})\}}\pi(ds,dz,du) \\
&+ \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(\check{X}_{s-}) - a(0)]\mathbf{1}_{\{f(\hat{X}_{s-})<z\le f(\check{X}_{s-})\}}\pi(ds,dz,du) =: A_t + M_t + \Delta_t,
\end{aligned}
$$

where A_t denotes the bounded variation part of the evolution, M_t the martingale part and Δ_t the sum of the three jump terms. Notice that

$$
M_t = \int_0^t (a'(\hat{X}_s)\sqrt{\hat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})\sigma dW_s
$$

is a square integrable martingale since f and a' are bounded.

We wish to obtain a control on $|Z_t^*| := \sup_{s\le t}|Z_s|$. We first take care of the jumps of $|Z_t|$. Notice first that, since f and a are bounded,

$$
\begin{aligned}
\Delta(x,y) := (f(x)\wedge f(y))|a(x)-a(y)| + |f(x)-f(y)|\, \Big| |a(0)-a(y)| + |a(0)-a(x)| \Big| \\
\le C|a(x)-a(y)|,
\end{aligned}
$$

implying that

$$
I\!E \sup_{s\le t}|\Delta_s| \le C I\!E \int_0^t |a(\hat{X}_s) - a(\check{X}_s)|ds \le Ct\, I\!E|Z_t^*|.
$$

Moreover, for a constant C depending on σ^2, $\|f\|_\infty, \|a\|_\infty, \|a'\|_\infty, \|a''\|_\infty$ and α,

$$
\begin{aligned}
I\!E \sup_{s\le t}|A_s| \le C\int_0^t I\!E|a'(\hat{X}_s)\hat{X}_s - a'(\check{X}_s)\check{X}_s|ds \\
+ C\left[\int_0^t |a''(\hat{X}_s) - a''(\check{X}_s)|ds + \int_0^t |\hat{\mu}_s(f) - \check{\mu}_s(f)|ds\right].
\end{aligned}
$$

We know that $|a'(\hat{X}_s)\hat{X}_s - a'(\check{X}_s)\check{X}_s| + |a''(\hat{X}_s) - a''(\check{X}_s)| \le C|a(\hat{X}_s) - a(\check{X}_s)| = C|Z_s|$. Therefore,

$$
I\!E \sup_{s\le t}|A_s| \le C I\!E\left[\int_0^t |Z_s|ds + \int_0^t |\hat{\mu}_s(f) - \check{\mu}_s(f)|ds\right].
$$

Moreover,

$$
|\hat{\mu}_s(f) - \check{\mu}_s(f)| = \left| I\!E\left(f(\hat{X}_s) - f(\check{X}_s)|\mathcal{W}\right)\right| \le I\!E\left(|f(\hat{X}_s) - f(\check{X}_s)||\mathcal{W}\right) \le I\!E(|Z_s||\mathcal{W}),
$$

and thus,

$$
I\!E\int_0^t |\hat{\mu}_s(f) - \check{\mu}_s(f)|ds \le I\!E\int_0^t |Z_s|ds \le tI\!E|Z_t^*|.
$$

Putting all these upper bounds together we conclude that for a constant C not depending on t,

$$\mathbb{E} \sup_{s \le t} |A_s| \le Ct\mathbb{E}|Z_t^*|.$$

Finally, we treat the martingale part using the Burkholder-Davis-Gundy inequality, and we obtain

$$\mathbb{E} \sup_{s \le t} |M_s| \le C\mathbb{E}\left[\left(\int_0^t (a'(\hat{X}_s)\sqrt{\hat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})^2 ds \right)^{1/2} \right].$$

But

$$\begin{aligned}(a'(\hat{X}_s)\sqrt{\hat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})^2 &\le C\left[((a'(\hat{X}_s) - a'(\check{X}_s))^2 + (\sqrt{\hat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)})^2 \right] \\ &\le C|Z_t^*|^2 + C(\sqrt{\hat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)})^2, \quad (4.8)\end{aligned}$$

where we have used that $|a'(x) - a'(y)| \le C|a(x) - a(y)|$ and that f and a' are bounded.

Finally, since $\inf f > 0$,

$$|\sqrt{\hat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)}|^2 \le C|\hat{\mu}_s(f) - \check{\mu}_s(f)|^2 \le C\left(\mathbb{E}(|Z_s^*||\mathcal{W}_s)\right)^2.$$

We use that $|Z_s^*| \le |Z_t^*|$, implying that $\mathbb{E}(|Z_s^*||\mathcal{W}) \le \mathbb{E}(|Z_t^*||\mathcal{W})$. Therefore we obtain the upper bound

$$|\sqrt{\hat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)}|^2 \le C\left(\mathbb{E}(|Z_t^*||\mathcal{W})\right)^2$$

for all $s \le t$, which implies the control of

$$\mathbb{E} \sup_{s \le t} |M_s| \le C\sqrt{t}\mathbb{E}|Z_t^*|.$$

The above upper bounds imply that, for a constant C not depending on t nor on the initial condition,

$$\mathbb{E}|Z_t^*| \le C(t + \sqrt{t})\mathbb{E}|Z_t^*|,$$

and therefore, for t_1 sufficiently small, $\mathbb{E}|Z_{t_1}^*| = 0$. We can repeat this argument on intervals $[t_1, 2t_1]$, with initial condition $\hat{X}_{t_1}$, and iterate it up to any finite T because t_1 does only depend on the coefficients of the system but not on the initial condition. This implies the assertion.

Proof of Item 2. The proof is done using a classical Picard-iteration. For that sake we introduce the sequence of processes $\bar{X}_t^{[0]} \equiv \bar{X}_0$, and construct the process $\bar{X}^{[n+1]}$ from $\bar{X}^{[n]}$ and $\mu_s^n = P(\bar{X}_s^{[n]} \in \cdot|\mathcal{W})$ in the following way:

- let $0 < \tau_1 < \tau_2 < ...$ be the jump times of

$$t \ge 0 \mapsto \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\left\{z \le f(\bar{X}_{s-}^{[n]})\right\}} \pi(ds, dz),$$

- set $\bar{X}_{\tau_i}^{[n+1]} = 0$ $(i \ge 1)$ and define, $\bar{X}^{[n+1]}$ between the jump times: for all $i \ge 1$, for all $t \in]\tau_i, \tau_{i+1}[$,

$$\bar{X}_t^{[n+1]} = \sigma \int_0^t e^{-\alpha(t-s)} \sqrt{\mu_s^n(f)} dW_s.$$

One can note, using the integration by parts formula, that $\bar{X}^{[n+1]}$ satisfies the following stochastic differential equation:

$$\bar{X}_t^{[n+1]} := \bar{X}_0 - \alpha \int_0^t \bar{X}_s^{[n+1]} ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-}^{[n+1]} \mathbf{1}_{\{z \le f(\bar{X}_{s-}^{[n]})\}} \pi(ds, dz, du) + \sigma \int_0^t \sqrt{\mu_s^n(f)} dW_s.$$

Using the same proof as the one of Lemma 4.3.2, we can show that for all $t \ge 0$,

$$\sup_{n\in\mathbb{N}} \sup_{0\le s\le t} \mathbb{E}\left[\left(\bar{X}_s^{[n]}\right)^2\right] < +\infty. \tag{4.9}$$

Now, we prove the convergence of $\bar{X}_t^{[n]}$. The same strategy as the one of the proof of Item 1. of Theorem 4.3.3 allows to show that

$$\delta_t^n := I\!\!E \sup_{s\le t} |a(\bar{X}_s^{[n]}) - a(\bar{X}_s^{[n-1]})| \text{ satisfies } \delta_t^n \le C(t+\sqrt{t})\delta_t^{n-1},$$

for all $n \ge 1$, for a constant C only depending on the parameters of the model, but not on n, neither on t. Choose t_1 such that

$$C(t_1 + \sqrt{t_1}) \le \frac{1}{2}.$$

Since $\sup_{s\le t_1} |a(\bar{X}_s^{[0]})| = a(\bar{X}_0) \le \|a\|_\infty$, we deduce from this that

$$\delta_{t_1}^n \le 2^{-n}\|a\|_\infty.$$

This implies the almost sure convergence of $a\left(\bar{X}_t^{[n]}\right)_n$ to some random variable Z_t for all $t \in [0, t_1]$. As a is an increasing function, the almost sure convergence of $\bar{X}_t^{[n]}$ to some (possibly infinite) random variable $\bar{X}_t$ follows from this. The almost sure finiteness of $\bar{X}_t$ is then guaranteed by Fatou's lemma and (4.9).

Now let us prove that $\bar{X}$ is solution of the limit equation (4.3) which follows by standard arguments (note that the jump term does not cause troubles because it is of finite activity). The most important point is to notice that

$$\mu_t^n(f) = I\!\!E(f(\bar{X}_t^{[n]})|\mathcal{W}) \to I\!\!E(f(\bar{X}_t)|\mathcal{W})$$

almost surely, which follows from the almost sure convergence of $f(\bar{X}_t^{[n]}) \to f(\bar{X}_t)$, using dominated convergence.

Once the convergence is proven on the time interval $[0, t_1]$, we can proceed iteratively over successive intervals $[kt_1, (k+1)t_1]$ to conclude that $\bar{X}$ is solution of (4.3) on $\mathbb{R}_+$.

Note that, thanks to this construction, we can show that $\mu_s(f)$ is measurable w.r.t. $\sigma(W_r : r \le s)$ and that

$$\mu_s(f) := \mathbb{E}\left[f(\bar{X}_s)\middle|\mathcal{W}\right] = \mathbb{E}\left[f(\bar{X}_s)\middle|\mathcal{W}_s\right].$$

Indeed, we can prove by induction that, for any n, $\mu_s^n(s)$ and $X_s^{[n]}$ are measurable w.r.t. $\sigma(\bar{X}_0)\vee\mathcal{W}_s \vee \sigma(\pi_{|[0,s]})$. For $n = 0$, it is clear since $\bar{X}_s^{[0]} = \bar{X}_0$ and $\mu_s^{[0]} = \mathcal{L}(\bar{X}_0)$. If we assume that the result holds true for some $n \in \mathbb{N}$, then, thanks to the definition of $\bar{X}^{[n+1]}$, we know that $\bar{X}_t^{[n+1]}$

is measurable w.r.t. to $\sigma(W_s : 0 \le s \le t) \vee \sigma(\pi(]u,v]) : 0 \le u < v \le t) \vee \sigma(\bar{X}_0)$. Then, since $\mu_t^{[n+1]}(f) = \mathbb{E}[f(\bar{X}_t^{[n+1]}|\mathcal{W}]$, and because we can write

$$\mathcal{W} = \sigma(W_s : 0 \le s \le t) \vee \sigma(W_r - W_t : r \ge t),$$

we know that $\mu_t^{[n+1]}(f) = \mathbb{E}[f(\bar{X}_t^{[n+1]})|\sigma(W_s : 0 \le s \le t)]$ (recalling that $\sigma(W_s : 0 \le s \le t)$ is independent of $\sigma(W_r - W_t : r \ge t) \vee \sigma(\pi) \vee \sigma(\bar{X}_0)$). □

4.3.3 Properties of the limit system

In the sequel, we shall also use an important property of the limit system (4.2), which is the conditional independence of the processes $\bar{X}^i$ $(i \ge 1)$ given the Brownian motion W. This property implies that $\mu = \mathcal{L}(\bar{X}^1|W)$ is the directing measure of the limit system (4.2).

Proposition 4.3.4. *If Assumption 4.3 holds then, $\bar{X}^1, \dots, \bar{X}^N$ are i.i.d. conditionally to W.*

Proof. The construction of the proof of Item 2. of Theorem 4.3.3, together with the proof of Theorem 1.1 of Chapter IV.1 and of Theorem 9.1 in Chapter IV.9 of Ikeda and Watanabe (1989), imply the existence of a measurable function Φ that does not depend on $k = 1, \dots, N$, and that satisfies, for each $1 \le k \le N$,

$$\bar{X}^k = \Phi(\bar{X}_0^k, \pi^k, W)$$

and for all $t \ge 0$,

$$\bar{X}^k_{|[0,t]} = \Phi_t(\bar{X}_0^k, \pi^k_{|[0,t]\times\mathbb{R}_+\times\mathbb{R}}, (W_s)_{s\le t}); \tag{4.10}$$

in other words, our process is non-anticipative and does only depend on the underlying noise up to time t.

Then we can write, for all continuous bounded functions g, h,

$$\mathbb{E}\left[g(\bar{X}^i)h(\bar{X}^j)\middle|\mathcal{W}\right] = \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, W))h(\Phi(\bar{X}_0^j, \pi^j, W))\middle|\mathcal{W}\right] = \psi(W),$$

where $\psi(w) := \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, w))h(\Phi(\bar{X}_0^j, \pi^j, w))\right] = \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, w))\right]\mathbb{E}\left[h(\Phi(\bar{X}_0^j, \pi^j, w))\right] =: \psi_i(w)\psi_j(w)$. With the same reasoning, we show that $\mathbb{E}\left[g(\bar{X}^i)\middle|\mathcal{W}\right] = \psi_i(W)$ and $\mathbb{E}\left[h(\bar{X}^j)\middle|\mathcal{W}\right] = \psi_j(W)$. The same arguments prove the mutual independence of $\bar{X}^1, \dots \bar{X}^N$ conditionally to W. □

Let us also mention that the random limit measure μ satisfies a nonlinear stochastic PDE in weak form. More precisely,

Corollary 4.3.5. *Grant Assumption 4.3 and suppose that $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$. Then the measure $\mu = P((\bar{X}_t)_{t\ge0} \in \cdot|W)$ satisfies the following nonlinear stochastic PDE in weak form: for any $\varphi \in C_b^2(\mathbb{R})$, for any $t \ge 0$,*

$$\int_{\mathbb{R}} \varphi(x)\mu_t(dx) = \int_{\mathbb{R}} \varphi(x)\nu_0(dx) + \int_0^t \left(\int_{\mathbb{R}} \varphi'(x)\mu_s(dx)\right) \sqrt{\mu_s(f)}\sigma dW_s$$
$$+ \int_0^t \int_{\mathbb{R}} \Big([\varphi(0) - \varphi(x)]f(x) - \alpha\varphi'(x)x + \frac{1}{2}\sigma^2\varphi''(x)\mu_s(f)\Big)\mu_s(dx)ds.$$

Proof. Applying Ito's formula, we have

$$\varphi(\bar{X}_t) = \varphi(\bar{X}_0) + \int_0^t \left(-\alpha\varphi'(\bar{X}_s)\bar{X}_s + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f) \right) ds + \int_0^t \varphi'(\bar{X}_s)\sqrt{\mu_s(f)}dW_s$$
$$+ \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbf{1}_{\{z\le f(\bar{X}_{s-})\}} \left(\varphi(0) - \varphi(\bar{X}_{s-}\right) \pi(ds,dz,du). \quad (4.11)$$

Since φ', φ'' and f are bounded, it follows from Lemma 4.3.2 and Fubini's theorem that

$$I\!E\left(\int_0^t \left(-\alpha\varphi'(\bar{X}_s)\bar{X}_s + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f)\right) ds|W \right) = \int_0^t E\left(-\alpha\varphi'(\bar{X}_s)\bar{X}_s + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f)|W \right) ds$$
$$= \int_0^t \int_{\mathbb{R}} \left(-\alpha\varphi'(x)x + \frac{1}{2}\varphi''(x)\mu_s(f) \right) \mu_s(dx)ds.$$

Moreover, by independence of $\bar{X}_0$ and W, $I\!E(\varphi(\bar{X}_0)|W) = \int_{\mathbb{R}} \varphi(x)\nu_0(dx)$.

To deal with the martingale part in (4.11), we use an Euler scheme to approximate the stochastic integral $I_t := \int_0^t \varphi'(\bar{X}_s)\sqrt{\mu_s(f)}dW_s$. For that sake, let $t_k^n := k2^{-n}t, 0 \le k \le 2^n$, $n \ge 1$, and define

$$I_t^n := \sum_{k=0}^{2^n-1} \varphi'(\bar{X}_{t_k^n})\Delta_k^n, \; \Delta_k^n = \int_{t_k^n}^{t_{k+1}^n} \sqrt{\mu_s(f)}dW_s,$$

then $I\!E(|I_t - I_t^n|^2) \to 0$ as $n \to \infty$, and therefore $I\!E(I_t^n|\mathcal{W}) \to I\!E(I_t|\mathcal{W})$ in $L^2(P)$, as $n \to \infty$. But

$$I\!E(I_t^n|\mathcal{W}) = \sum_{k=0}^{2^n-1} I\!E(\varphi'(\bar{X}_{t_k^n})|\mathcal{W})\Delta_k^n \to \int_0^t I\!E(\varphi'(\bar{X}_s)|\mathcal{W})\sqrt{\mu_s(f)}dW_s$$

in $L^2(P)$, since the sequence of processes $Y_s^n := \sum_{k=0}^{2^n-1} \mathbf{1}_{]t_k^n,t_{k+1}^n]}(s)I\!E(\varphi'(\bar{X}_{t_k^n})|\mathcal{W}), 0 \le s \le t$, converges in $L^2(\Omega \times [0,t])$ to $I\!E(\varphi'(\bar{X}_s)|\mathcal{W})$.

We finally deal with the jump part in (4.11). Since f is bounded, and by independence of W and π, we can rewrite this part in terms of an underlying Poisson process N_t, independent of W and having rate $\|f\|_\infty$, and in terms of i.i.d. variables $(V_n)_{n\ge1}$ uniformly distributed on $[0,1]$, independent of W and of N as follows.

$$\int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbf{1}_{\{z\le f(\bar{X}_{s-})\}} \left(\varphi(0) - \varphi(\bar{X}_{s-}\right) \pi(ds,dz,du) = \sum_{n=1}^{N_t} \mathbf{1}_{\{\|f\|_\infty V_n \le f(\bar{X}_{T_n-})\}}(\varphi(0) - \varphi(\bar{X}_{T_n-})).$$

Taking conditional expectation $E(\cdot|\mathcal{W})$, we obtain

$$I\!E\left(\sum_{n=1}^{N_t} \mathbf{1}_{\{\|f\|_\infty V_n \le f(\bar{X}_{T_n-})\}}(\varphi(0) - \varphi(\bar{X}_{T_n-}))|\mathcal{W} \right)$$
$$= I\!E\left(\sum_{n=1}^{N_t} \frac{f(\bar{X}_{T_n-})}{\|f\|_\infty}(\varphi(0) - \varphi(\bar{X}_{T_n-}))|\mathcal{W} \right)$$
$$= \int_0^t I\!E\left(f(\bar{X}_s)(\varphi(0) - \varphi(\bar{X}_s))|\mathcal{W} \right) ds,$$

where we have used the independence properties of $(V_n)_n, N_t$ and W and the fact that conditionally on $\{N_t = n\}$, the jump times $(T_1, \ldots, T_n)$ are distributed as the order statistics of n i.i.d. times which are uniformly distributed on $[0, t]$. This concludes our proof. □

4.3.4 Another version of the limit system

We have already stated the existence of a unique strong solution $(\bar{X}^i)_{i\geq 1}$ of the system (4.2). In the sequel we also need to show the well-posedness of the following exchangeable system of stochastic differential equations:

$$\begin{cases} \bar{Y}^i_t &= \bar{Y}^i_0 - \alpha \int_0^t \bar{Y}^i_s ds - \int_{[0,t]\times\mathbb{R}_+} \bar{Y}^i_{s-} \mathbf{1}_{\{z\leq f(\bar{Y}^i_{s-})\}} \pi^i(ds,dz) \\ & \quad + \sigma \int_0^t \sqrt{\mu_t(f)} dW_s, \\ (\bar{Y}^i_0)_{i\geq 1} & \text{are i.i.d., distributed } \sim \nu_0, \end{cases} \tag{4.12}$$

where μ is the directing measure of the exchangeable system $(\bar{Y}^i)_{i\geq 1}$ and μ_t its projection onto the $t-$th time coordinate. According to our previous reasoning, any strong solution of (4.2) is also solution of (4.12). But the converse is not obvious, because it is not a priori clear whether the Brownian motion is the only common noise of the system (4.12). We claim it in the next result. Let us note that the notion of strong solution for (4.12) is not clear since the form of the directing measure is not stated. So the solutions of (4.12) have to be understood in a weak sense.

Theorem 4.3.6. *Grant Assumption 4.3 and suppose that $\int_{\mathbb{R}} x^2 d\nu_0(x) < \infty$. Then there exists a unique weak solution $(\bar{Y}^i)_{i\geq 1}$ of (4.12). This solution is given by the unique strong solution of (4.2).*

Proof. Step 1: Let us begin by proving that any solution $(\bar{X}^i)_{i\geq 1}$ of (4.2) is solution of (4.12).

By Proposition 4.3.4.(ii), conditionally to $\mathcal{W}$, the variables $\bar{X}^i_t$ are i.i.d. This implies that the directing measure of $(\bar{X}^i_t)_{i\geq 1}$ is $\mathcal{L}(\bar{X}^i_t|\mathcal{W})$ (see Lemma (2.12) of Aldous (1983)).

Step 2. It is now sufficient to prove that $(\bar{X}^i)_{i\geq 1}$ is the only solution of (4.12) defined w.r.t. the same Brownian motion, Poisson random measures and initial conditions. For that sake, let us consider $(\bar{Y}^i)_{i\geq 1}$ any solution of (4.12), and prove that $(\bar{X}^i)_{i\geq 1} = (\bar{Y}^i)_{i\geq 1}$ almost surely. In the rest of the proof, μ_t denotes only the directing measure of the system $(\bar{Y}^i_t)_{i\geq 1}$. So we want to prove that $\mu_t(f) := \mathbb{E}\left[f(\bar{Y}^1_t) \middle| \mu\right] = \mathbb{E}\left[f(\bar{Y}^1_t) \middle| \mathcal{W}\right]$ *a.s.*.

To begin with, Lemma (2.15) of Aldous (1983) implies that $\mu_t(f)$ is the almost sure limit of $N^{-1}\sum_{j=1}^N f(\bar{Y}^j_t)$. We now prove that this sequence converges to $\mathbb{E}\left[f(\bar{Y}^1_t) \middle| \mathcal{W}\right]$. For this purpose, we introduce an auxiliary system $(\widetilde{X}^{N,i})_{1\leq i\leq N}$, driven by the same Brownian motion W and the same Poisson random measures π^i, with $\bar{Y}^i_0 = \tilde{X}^{N,i}_0$ $(i \geq 1)$, replacing the term $\mu_t(f)$ by the empirical measure:

$$d\widetilde{X}^{N,i}_t = -\alpha \widetilde{X}^{N,i}_t dt + \sqrt{\frac{1}{N}\sum_{j=1}^N f(\widetilde{X}^{N,j}_t)} dW_t - \widetilde{X}^{N,i}_{t-} \int_{\mathbb{R}_+} \mathbb{1}_{\{z\leq f(\widetilde{X}^{N,i}_{t-})\}} \pi^i(dt,dz), \ \widetilde{X}^{N,i}_0 = \bar{Y}^i_0.$$

Notice that $(\bar{X}^i)_{i\geq 1}, (\bar{Y}^i)_{i\geq 1}$ and $(\tilde{X}^{N,i})_{1\leq i\leq N}$ are all defined on the same probability space, driven by the same Brownian motion W and the same Poisson random measures π^i.

It is now sufficient to prove that both for $(\bar{Y}^i)_{i\geq 1}$ and for $(\bar{X}^i)_{i\geq 1}$,

$$\mathbb{E}\left[\left|a(\bar{Y}^i_t) - a(\widetilde{X}^{N,i}_t)\right|\right] + \mathbb{E}\left[\left|a(\bar{X}^i_t) - a(\widetilde{X}^{N,i}_t)\right|\right] \leq C_t N^{-1/2}. \tag{4.13}$$

Indeed, suppose we have already proven the above control (4.13). Then

$$\mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N f(\bar{Y}^j_t) - \mathbb{E}\left[f(\bar{X}^1_t)\middle|\mathcal{W}\right]\right|\right] \leq \frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[|f(\bar{Y}^j_t) - f(\widetilde{X}^{N,j}_t)|\right]$$

$$+\frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[|f(\widetilde{X}^{N,j}_t) - f(\bar{X}^j_t)|\right] + \mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N f(\bar{X}^j_t) - \mathbb{E}\left[f(\bar{X}^1_t)\middle|\mathcal{W}\right]\right|\right].$$

Then, (4.13) and Assumption 4.3 imply that the first and the second term of the sum above are smaller than $C_t N^{-1/2}$ for some $C_t > 0$. In addition, by item (ii) of Proposition 4.3.4, the variables $(\bar{X}^j)_{1\leq j\leq N}$ are i.i.d., conditionally on W. Consequently, and since f is bounded, the third term is smaller than $C_t N^{-1/2}$.

The above implies that, as $N\to\infty$, $\frac{1}{N}\sum_{j=1}^N f(\bar{Y}^j_t)$ converges in $L^1(P)$ to $\mathbb{E}\left[f(\bar{X}^1_t)\middle|\mathcal{W}\right]$. On the other hand, we know this sequence converges almost surely to $\mu_t(f)$. Thus,

$$\mathbb{E}\left[f(\bar{Y}^1_t)\middle|\mu\right] = \mu_t(f) = \mathbb{E}\left[f(\bar{X}^1_t)\middle|\mathcal{W}\right] = \mathbb{E}\left[f(\bar{X}^i_t)\middle|\mathcal{W}\right] \text{ a.s..}$$

As a consequence, $(\bar{Y}^i)_{i\geq 1}$ is solution of the infinite system

$$d\bar{Y}^i_t = -\alpha\bar{Y}^i_t dt + \sigma\sqrt{\mathbb{E}\left[f(\bar{X}^i_t)\middle|\mathcal{W}\right]}dW_t - \bar{Y}^i_{t-}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\leq f(\bar{Y}^i_{t-})\right\}}\pi^i(dt,dz),$$

while $(\bar{X}^i)_{i\geq 1}$ in (4.2) is solution of

$$d\bar{X}^i_t = -\alpha\bar{X}^i_t dt + \sigma\sqrt{\mathbb{E}\left[f(\bar{X}^i_t)\middle|\mathcal{W}\right]}dW_t - \bar{X}^i_{t-}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\leq f(\bar{X}^i_{t-})\right\}}\pi^i(dt,dz),$$

with $\bar{X}^i_0 = \bar{Y}^i_0$, for all $i\geq 1$.

Step 3. In the previous step, we have proved that $\mathbb{E}\left[f(\bar{Y}^i_t)\middle|\mu\right] = \mathbb{E}\left[f(\bar{X}^i_t)\middle|\mathcal{W}\right]$. So, we still have not proved that $\bar{Y}^i$ satisfies the equation (4.2), which would have allowed to conclude that $\bar{Y} = \bar{X}$ almost surely (by Theorem 4.3.3.(i)). We prove it in this step.

For that sake, consider $\tau_M = \inf\{t > 0 : |\bar{X}^i_t| \wedge |\bar{Y}^i_t| > M\}$ for $M > 0$. We prove that $\mathbb{E}\left[|\bar{X}^i_{t\wedge\tau_M} - \bar{Y}^i_{t\wedge\tau_M}|\right] = 0$ for all $M > 0$. Recalling Lemma 4.3.2 and the fact that we can prove a similar control for $\bar{Y}^i$ this implies, by Fatou's lemma, that $\mathbb{E}\left[|\bar{X}^i_t - \bar{Y}^i_t|\right] = 0$.

Let $u_M(t) := \mathbb{E}\left[|\bar{X}^i_{t\wedge\tau_M} - \bar{Y}^i_{t\wedge\tau_M}|\right]$. To see that $u_M(t) = 0$, it is sufficient to apply Grönwall's lemma to the following inequality

$$u_M(t) \leq \alpha\int_0^t u_M(s)ds + \mathbb{E}\left[\int_{[0,t\wedge\tau_M]\times\mathbb{R}_+}\left|\bar{X}^i_{s-}\mathbb{1}_{\left\{z\leq f(\bar{X}^i_{s-})\right\}} - \bar{Y}^i_{s-}\mathbb{1}_{\left\{z\leq f(\bar{Y}^i_{s-})\right\}}\right|\pi^i(ds,dz)\right]$$

implying that

$$u_M(t) \le \alpha \int_0^t u_M(s)ds + \mathbb{E}\left[\int_{[0,t\wedge\tau_M]\times\mathbb{R}_+} \mathbb{1}_{\left\{z\in[0,f(\bar X^i_{s-})\wedge f(\bar Y^i_{s-})]\right\}} \left|\bar X^i_{s-} - \bar Y^i_{s-}\right| \pi^i(ds,dz)\right]$$
$$+ \mathbb{E}\left[\int_{[0,t\wedge\tau_M]\times\mathbb{R}_+} \mathbb{1}_{\left\{z\in]f(\bar X^i_{s-})\wedge f(\bar Y^i_{s-}), f(\bar X^i_{s-})\vee f(\bar Y^i_{s-})]\right\}} |\bar X^i_{s-}| \vee |\bar Y^i_{s-}| \pi^i(ds,dz)\right],$$

whence

$$u_M(t) \le C(1+M)\int_0^t u_M(s)ds$$

and thus $u_M(t) = 0$.

Hence $(\bar Y^i)_{i\ge 1}$ is solution of the infinite system (4.2) and $\mu = \mathcal{L}(\bar Y^1|\mathcal{W})$, its directing measure, is uniquely determined.

Step 4. Finally, let us show (4.13). We only prove it for $\bar Y^i$, the proof for $\bar X^i$ is similar. By exchangeability, it is sufficient to work with $\bar Y^1$. We decompose the evolution of $a(\bar Y^1_t)$ in the following way.

$$a(\bar Y^1_t) = a(\bar Y^1_0) - \alpha\int_0^t a'(\bar Y^1_s)\bar Y^1_s ds + \int_{[0,t]\times\mathbb{R}_+} \left(a(0) - a(\bar Y^1_{s-})\right) \mathbf{1}_{\{z\le f(\bar Y^1_{s-})\}}\pi^1(ds,dz) \tag{4.14}$$
$$+ \frac{\sigma^2}{2}\int_0^t a''(\bar Y^1_s)\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s)ds - B^N_t + \sigma\int_0^t a'(\bar Y^1_s)\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s)}dW_s - M^N_t,$$

where

$$B^N_t = \frac{\sigma^2}{2}\int_0^t a''(\bar Y^1_s)\left(\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s) - \mathbb{E}\left[f(\bar Y^1_s)\middle|\,\mu\right]\right)ds$$

and

$$M^N_t = \sigma\int_0^t a'(\bar Y^1_s)\left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s)} - \sqrt{\mathbb{E}\left[f(\bar Y^1_s)|\mu\right]}\right)dW_s.$$

Clearly,

$$< M^N >_t \le \sigma^2\left(\sup_{x\in\mathbb{R}}|a'(x)^2|\right)\int_0^t\left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s)} - \sqrt{\mathbb{E}\left[f(\bar Y^1_s)|\mu\right]}\right)^2 ds.$$

Recalling that the variables $\bar Y^j_s$ $(1\le j\le N)$ are i.i.d. conditionally to μ, taking conditional expectation $I\!E(\cdot|\mu)$ and using the fact that f is lower bounded such that

$$\left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s)} - \sqrt{\mathbb{E}\left[f(\bar Y^1_s)|\mu\right]}\right)^2 \le C\left(\frac{1}{N}\sum_{j=1}^N f(\bar Y^j_s) - \mathbb{E}\left[f(\bar Y^1_s)|\mu\right]\right)^2,$$

we deduce that

$$\mathbb{E}\left[< M^N >_t\right] \le C_t N^{-1} \text{ and } \mathbb{E}\left[B^N_t\right] \le C_t N^{-1/2}.$$

Then, applying Itô's formula on $a(\tilde{X}^{N,1})$, we obtain the same equation as (4.14), but without the terms B_t^N and M_t^N. Introducing

$$u(t) := \sup_{0\le s\le t} \mathbb{E}\left[\left|a(\bar{Y}_s^1) - a(\tilde{X}_s^{N,1})\right|\right],$$

we can prove with the same reasoning as in the proof of Theorem 4.3.3 that

$$u(t) \le C(t+\sqrt{t})u(t) + \frac{C_t}{\sqrt{N}},$$

where C and C_t are independent of N. Finally, using the arguments of the proof of Theorem 4.3.3, this implies (4.13). To prove it for $\bar{X}^1$ instead of $\bar{Y}^1$, one has to use the fact that the $\bar{X}^j$ are i.i.d. conditionally to W, so one has to replace in *Step 4*, $\mu_s(f)$ by $\mathbb{E}\left[f(\bar{X}_s^j)\middle|\mathcal{W}\right]$, and μ by $\mathcal{W}$ in the conditionings. □

4.4 Convergence of $(X^{N,i})_{1\le i\le N}$ in distribution in Skorohod topology

The main result of this chapter is

Theorem 4.4.1. *Grant Assumptions 4.1, 4.2 and 4.3. $(X^{N,i})_{1\le i\le N}$ converges to $(\bar{X}^i)_{i\ge1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$ endowed with the product topology.*

According to Proposition (7.20) of Aldous (1983), as the systems $(X^{N,i})_{1\le i\le N}$ and $(\bar{X}^i)_{i\ge1}$ are exchangeable, and the directing measure of $(\bar{X}^i)_{i\ge1}$ is given by $\mathcal{L}(\bar{X}^1|W)$, Theorem 4.4.1 is equivalent to

Theorem 4.4.2. *Suppose that Assumptions 4.1, 4.2 and 4.3 hold. Then the empirical measure $\mu^N = \frac{1}{N}\sum_{i=1}^N \delta_{X^{N,i}}$ of the $N-$particle system $(X^{N,i})_{1\le i\le N}$ converges in distribution in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$ to $\mu := \mathcal{L}(\bar{X}^1|\mathcal{W})$, where $(\bar{X}^i)_{i\ge1}$ is solution of (4.2).*

In fact, Theorem 4.4.2 is easier to prove than Theorem 4.4.1, since the trajectories of the system $(X^{N,i})_{1\le i\le N}$ belong to $D(\mathbb{R}_+,\mathbb{R})^N$ that depends on N.

We will prove Theorem 4.4.2 in a two step procedure. Firstly we prove the tightness of the sequence of empirical measures, and then in a second step we identify all possible limits as solutions of a martingale problem, whose solution will be shown to be unique.

4.4.1 Tightness of $(\mu^N)_N$

Proposition 4.4.3. *Grant Assumptions 4.1 and 4.2, and assume that f is bounded. For each $N\ge1$, consider the unique solution $(X_t^N)_{t\ge0}$ to (4.1) starting from some i.i.d. ν_0-distributed initial conditions $X_0^{N,i}$.*

(i) The sequence of processes $(X_t^{N,1})_{t\ge0}$ is tight in $D(\mathbb{R}_+,\mathbb{R})$.

(ii) The sequence of empirical measures $\mu^N = N^{-1}\sum_{i=1}^N \delta_{(X_t^{N,i})_{t\ge0}}$ is tight in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$.

Proof. First, it is well-known that point (ii) follows from point (i) and the exchangeability of the system, see (Sznitman, 1989, Proposition 2.2-(ii)). We thus only prove (i). To show that the family $((X_t^{N,1})_{t\geq 0})_{N\geq 1}$ is tight in $D(\mathbb{R}_+,\mathbb{R})$, we use the criterion of Aldous, see Theorem 4.5 of Jacod and Shiryaev (2003). It is sufficient to prove that

(a) for all $T>0$, all $\varepsilon>0$, $\lim_{\delta\downarrow 0}\limsup_{N\to\infty}\sup_{(S,S')\in A_{\delta,T}} P(|X_{S'}^{N,1}-X_S^{N,1}|>\varepsilon)=0$, where $A_{\delta,T}$ is the set of all pairs of stopping times (S,S') such that $0\leq S\leq S'\leq S+\delta\leq T$ a.s.,

(b) for all $T>0$, $\lim_{K\uparrow\infty}\sup_N P(\sup_{t\in[0,T]}|X_t^{N,1}|\geq K)=0$.

To check (a), consider $(S,S')\in A_{\delta,T}$ and write

$$X_{S'}^{N,1}-X_S^{N,1}=-\int_S^{S'}\int_{\mathbb{R}}\int_0^\infty X_{s-}^{N,1}\mathbf{1}_{\{z\leq f(X_{s-}^{N,1})\}}\pi^1(ds,du,dz)-\alpha\int_S^{S'}X_s^{N,1}ds + \frac{1}{\sqrt{N}}\sum_{j=2}^N\int_S^{S'}\int_{\mathbb{R}}\int_0^\infty u\mathbf{1}_{\{z\leq f(X_{s-}^{N,j})\}}\pi^j(ds,du,dz),$$

implying that

$$|X_{S'}^{N,1}-X_S^{N,1}|\leq\Big|\int_S^{S'}\int_{\mathbb{R}}\int_0^\infty X_{s-}^{N,1}\mathbf{1}_{\{z\leq f(X_{s-}^{N,1})\}}\pi^1(ds,du,dz)\Big| + \delta\alpha\sup_{0\leq s\leq T}\left|X_s^{N,1}\right|+\Big|\frac{1}{\sqrt{N}}\sum_{j=2}^N\int_S^{S'}\int_{\mathbb{R}}\int_0^\infty u\mathbf{1}_{\{z\leq f(X_{s-}^{N,j})\}}\pi^j(ds,du,dz)\Big| =: |I_{S,S'}|+\delta\alpha\sup_{0\leq s\leq T}\left|X_s^{N,1}\right|+|J_{S,S'}|.$$

We first note that $|I_{S,S'}|>0$ implies that $\tilde{I}_{S,S'}:=\int_S^{S'}\int_{\mathbb{R}}\int_0^\infty\mathbf{1}_{\{z\leq f(X_{s-}^{N,1})\}}\pi^i(ds,du,dz)\geq 1$, whence

$$P(|I_{S,S'}|>0)\leq P(\tilde{I}_{S,S'}\geq 1)\leq I\!E[\tilde{I}_{S,S'}]\leq I\!E\Big[\int_S^{S+\delta}f(X_s^{N,1})ds\Big]\leq ||f||_\infty\delta,$$

since f is bounded. We proceed similarly to check that

$$P(|J_{S,S'}|\geq\varepsilon)\leq\frac{1}{\varepsilon^2}I\!E[(J_{S,S'})^2]\leq\frac{\sigma^2}{N\varepsilon^2}\sum_{j=2}^N I\!E\Big[\int_S^{S+\delta}f(X_s^{N,j})ds\Big]\leq\frac{\sigma^2}{\varepsilon^2}\|f\|_\infty\delta.$$

The term $\sup_{0\leq s\leq T}|X_s^{N,1}|$ can be handled using Lemma 4.3.1.(ii).

Finally (b) is a straightforward consequence of Lemma 4.3.1.(ii) and Markov's inequality. □

4.4.2 Martingale problem

We now introduce a new martingale problem, whose solutions are the limits of any converging subsequence of $\mu^N=\frac{1}{N}\sum_{j=1}^N\delta_{X^{N,j}}$. In this martingale problem, we are interested in couples of trajectories to be able to put hands on the correlations between the particles. In particular, this will allow us to show that, in the limit system (4.2), the processes $\bar{X}^i$ $(i \geq 1)$ share the same Brownian motion, but are driven by Poisson measures π^i $(i \geq 1)$ which are independent. The reason why we only need to study the correlation between two particles is the exchangeability of the infinite system.

Let Q be a distribution on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))$. Define a probability measure P on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \times D(\mathbb{R}_+, \mathbb{R})^2$ by

$$P(A \times B) := \int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))} \mathbb{1}_A(m) m \otimes m(B) Q(dm). \tag{4.15}$$

We write any atomic event $\omega \in \Omega := \mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \times D(\mathbb{R}_+, \mathbb{R})^2$ as $\omega = (\mu, Y)$, with $Y = (Y^1, Y^2)$. Thus, the law of the canonical variable μ is Q, and that of $(Y_t)_{t \geq 0}$ is

$$P_Y = \int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))} Q(dm) m \otimes m(\cdot).$$

Moreover we have $P-$ almost surely

$$\mu = \mathcal{L}(Y^1|\mu) = \mathcal{L}(Y^2|\mu) \text{ and } \mathcal{L}(Y|\mu) = \mu \otimes \mu.$$

Writing $\mu_t := \int_{D(\mathbb{R}_+,\mathbb{R})} \mu(d\gamma)\delta_{\gamma_t}$ for the projection onto the $t-$th time coordinate, we introduce the filtration

$$\mathcal{G}_t = \sigma(Y_s, s \leq t) \vee \sigma(\mu_s(f), s \leq t).$$

Definition 4.4.4. *We say that $Q \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})))$ is a solution to the martingale problem $(\mathcal{M})$ if the following holds.*

(i) $Q-$almost surely, $\mu_0 = \nu_0$.

(ii) For all $g \in C_b^2(\mathbb{R}^2)$, $M_t^g := g(Y_t) - g(Y_0) - \int_0^t Lg(\mu_s, Y_s) ds$ is a $(P, (\mathcal{G}_t)_t)-$martingale, where

$$\begin{aligned} Lg(\mu, x) =& -\alpha x^1 \partial_{x^1} g(x) - \alpha x^2 \partial_{x^2} g(x) + \frac{\sigma^2}{2} \mu(f) \sum_{i,j=1}^2 \partial^2_{x^i x^j} g(x) \\ &+ f(x^1)(g(0, x^2) - g(x)) + f(x^2)(g(x^1, 0) - g(x)). \end{aligned}$$

Let $(\bar{X}^i)_{i \geq 1}$ be the solution of the limit system (4.2) and $\mu = \mathcal{L}(\bar{X}^1|\mathcal{W})$. Then Proposition 4.3.4.(ii) and Lemma (2.12).(a) of Aldous (1983) imply that μ is the directing measure of $(\bar{X}^i)_{i \geq 1}$. Thus the law of $(\mu, \bar{X}^1, \bar{X}^2)$ is P given in (4.15). And, by Itô's formula, $(\bar{X}^1, \bar{X}^2)$ satisfies the martingale property of Definition 4.4.4. In other words, $\mathcal{L}(\mu)$ is a solution of $(\mathcal{M})$.

Let us now characterize any possible solution of $(\mathcal{M})$, which is the first step to prove uniqueness of the solution of $(\mathcal{M})$.

Lemma 4.4.5. *Let $Q \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})))$. Assume that Q is a solution of $(\mathcal{M})$ and that f is bounded. Let (μ, Y) be the canonical variable defined above, and write $Y = (Y^1, Y^2)$. Then there exists a standard $(\mathcal{G}_t)_t-$Brownian motion W and on an extension $(\tilde{\Omega}, (\tilde{\mathcal{G}}_t)_t, \tilde{P})$ of $(\Omega, (\mathcal{G}_t)_t, P)$ there exist $(\tilde{\mathcal{G}}_t)_t-$ Poisson random measures π^1, π^2 on $\mathbb{R}_+ \times \mathbb{R}_+$ having Lebesgue intensity such that W, π^1 and π^2 are independent and*

$$\begin{aligned} dY_t^1 =& -\alpha Y_t^1 dt + \sigma \sqrt{\mu_t(f)} dW_t - Y_{t-}^1 \int_{\mathbb{R}_+} \mathbb{1}_{\{z \leq f(Y_{t-}^1)\}} \pi^1(dt, dz), \\ dY_t^2 =& -\alpha Y_t^2 dt + \sigma \sqrt{\mu_t(f)} dW_t - Y_{t-}^2 \int_{\mathbb{R}_+} \mathbb{1}_{\{z \leq f(Y_{t-}^2)\}} \pi^2(dt, dz). \end{aligned}$$

Proof. Item (ii) of of $(\mathcal{M})$ together with Theorem II.2.42 of Jacod and Shiryaev (2003) imply that Y is a semimartingale with characteristics (B, C, ν) given by

$$B_t^i = -\alpha \int_0^t Y_s^i ds - \int_0^t Y_s^i f(Y_s^i) ds, \quad 1 \le i \le 2,$$
$$C_t^{i,j} = \int_0^t \mu_s(f) ds, \quad 1 \le i, j \le 2,$$
$$\nu(dt, dy) = dt(f(Y_{t-}^1)\delta_{(-Y_{t-}^1, 0)}(dy) + f(Y_{t-}^2)\delta_{(0, -Y_{t-}^2)}(dy)).$$

Then we can use the canonical representation of Y (see Theorem II.2.34 of Jacod and Shiryaev (2003)) with the truncation function $h(y) = y$ for every y: $Y_t - Y_0 - B_t = M_t^c + M_t^d$, where M^c is a continuous local martingale and M^d a purely discontinuous local martingale. By definition of the characteristics, $\langle M^{c,i}, M^{c,j} \rangle_t = C_t^{i,j}$. In particular, $\langle M^{c,i} \rangle_t = \int_0^t \mu_s(f) ds$ $(i = 1, 2)$. Consequently, applying Theorem II.7.1' of Ikeda and Watanabe (1989) to the 2-dimensional martingale $(M^{c,1}, M^{c,2})$, we know that there exists a Brownian motion W such that

$$M_t^{c,i} = \int_0^t \sqrt{\mu_s(f)} dW_s, \quad i = 1, 2.$$

We now prove the existence of the independent Poisson measures π^1, π^2. We know that $M^d = h * (\mu^Y - \nu)$, where $\mu^Y = \sum_s \mathbb{1}_{\{\Delta Y_s \neq 0\}} \delta_{(s, Y_s)}$ is the jump measure of Y and ν is its compensator. We rely on Theorem II.7.4 of Ikeda and Watanabe (1989). Using the notation therein, we introduce $Z = \mathbb{R}_+$, m Lebesgue measure on Z and

$$\theta(t, z) := (-Y_{t-}^1, 0)\mathbb{1}_{\left\{z \le f(Y_{t-}^1)\right\}} + (0, -Y_{t-}^2)\mathbb{1}_{\left\{||f||_\infty < z \le ||f||_\infty + f(Y_{t-}^2)\right\}}.$$

According to Theorem II.7.4 of Ikeda and Watanabe (1989), there exists a Poisson measure π on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity $dt \cdot dz$ such that, for all $E \in \mathcal{B}(\mathbb{R}^2)$,

$$\mu^Y([0, t] \times E) = \int_0^t \int_0^\infty \mathbb{1}_{\{\theta(s,z) \in E\}} \pi(ds, dz). \tag{4.16}$$

In what follows we show how to construct two independent Poisson random measures π^1 and π^2 from π with the desired representation property, using two disjoint parts of π. For π^1 we use $\pi_{|\mathbb{R}_+ \times [0, ||f||_\infty]}$, and for π^2 we use $\pi_{|\mathbb{R}_+ \times [||f||_\infty, 2||f||_\infty]}$ such that the Poisson measures π^1 and π^2 will be independent.

To construct π_1 and π_2, we also consider two independent Poisson measures $\tilde{\pi}^1, \tilde{\pi}^2$ (independent of everything else) on $[||f||_\infty, \infty[$ having Lebesgue intensity. We then define π^1 in the following way: for any $A \in \mathcal{B}(\mathbb{R}_+ \times [0, ||f||_\infty])$, $\pi^1(A) = \pi(A)$, and for $A \in \mathcal{B}(\mathbb{R}_+ \times]||f||_\infty, \infty[)$, $\pi^1(A) = \tilde{\pi}^1(A)$. We define π^2 in a similar way: For $A \in \mathcal{B}(\mathbb{R}_+ \times [0, ||f||_\infty])$, $\pi^2(A) = \pi(\{(t, ||f||_\infty + z) : (t, z) \in A\})$, and for $A \in \mathcal{B}(\mathbb{R}_+ \times]||f||_\infty, \infty[)$, $\pi^2(A) = \tilde{\pi}^2(A)$. By definition of Poisson measures, π^1 and π^2 are independent Poisson measures on $\mathbb{R}_+^2$ having Lebesgue intensity, and together with (4.16), we have

$$M_t^{d,i} = -\int_{[0,t] \times \mathbb{R}_+} Y_{s-}^i \mathbb{1}_{\left\{z \le f(Y_{s-}^i)\right\}} \pi^i(ds, dz) + \int_0^t Y_s^i f(Y_s^i) ds, \quad 1 \le i \le 2.$$

□

In the next step we prove that there exists at most one (and thus exactly one) solution for the martingale problem $(\mathcal{M})$ using Lemma 4.4.5 and Theorem 4.3.6.

Theorem 4.4.6. *Grant Assumptions 4.2, 4.3 and 4.4. Then there is a unique solution Q of the martingale problem $(\mathcal{M})$. This solution can be written as $Q = \mathcal{L}(\mu)$, with $\mu = \mathcal{L}(\bar{X}|\mathcal{W})$, where $\bar{X}$ is the unique strong solution of (4.3).*

The main idea of the proof is to apply Lemma 4.4.5 to recover the system of equations (4.12) and then to rely on Theorem 4.3.6.

Proof. Let $Q \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})))$ be a solution of $(\mathcal{M})$ and write $Q = \mathcal{L}(\mu)$. The proof consists in showing that μ is the distribution of the directing measure of the system (4.12), which is unique by Theorem 4.3.6.

To begin with, we can assume that μ is the directing measure of some exchangeable system $(\bar{Y}^i)_{i\geq 1}$. Indeed, it is sufficient to work on the canonical space

$$\Omega' = \mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \times D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$$

endowed with the probability measure P' defined as follows. For all $A \in \mathcal{B}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))), B_k \in \mathcal{B}(D(\mathbb{R}_+, \mathbb{R}))$ $(k \geq 1)$ where at most a finite number of sets B_k $(k \geq 1)$ are different from $D(\mathbb{R}_+, \mathbb{R})$,

$$P'(A \times B_1 \times ...B_k \times ...) = \int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))} \mathbb{1}_A(m) m^{\otimes \mathbb{N}^*}(B_1 \times ... \times B_k \times ...)Q(dm).$$

Then, noting $(\mu, \bar{Y}^1, \bar{Y}^2, ..., \bar{Y}^k, ...)$ the canonical random variables on Ω', we know that μ is the directing measure of the exchangeable system $(\bar{Y}^i)_{i\geq 1}$. In particular, for all $i \neq j$,

$$\mathcal{L}(\mu, (\bar{Y}^i, \bar{Y}^j)) = P,$$

where P is given by (4.15), with $Q = \mathcal{L}(\mu)$.

Thanks to Lemma 4.4.5 we know that there exist Brownian motions $W^{(i,j)}$ $(i, j \geq 1)$ and Poisson random measures $\pi^{(i,j),1}, \pi^{(i,j),2}$ $(i, j \geq 1)$ such that for all pairs $(i, j), i \neq j$, $\pi^{(i,j),1}$ is independent of $\pi^{(i,j),2}$ and such that

$$\begin{aligned} d\bar{Y}^i_t &= -\alpha \bar{Y}^i_t dt + \sigma\sqrt{\mu_t(f)}dW^{(i,j)}_t - \bar{Y}^i_{t-}\int_{\mathbb{R}_+} \mathbb{1}_{\{z\leq f(\bar{Y}^i_{t-})\}}\pi^{(i,j),1}(dt, dz),\\ d\bar{Y}^j_t &= -\alpha \bar{Y}^j_t dt + \sigma\sqrt{\mu_t(f)}dW^{(i,j)}_t - \bar{Y}^j_{t-}\int_{\mathbb{R}_+} \mathbb{1}_{\{z\leq f(\bar{Y}^j_{t-})\}}\pi^{(i,j),2}(dt, dz). \end{aligned}$$

The exchangeability of the system $(\bar{Y}^i)_{i\geq 1}$ implies that for all distinct $i, j \geq 1$, $j > 1$, $\mathcal{L}(\bar{X}^1, \bar{X}^2) = \mathcal{L}(\bar{X}^1, \bar{X}^j) = \mathcal{L}(\bar{X}^i, \bar{X}^j)$. Thus $W := W^{(1,2)} = W^{(1,j)} = W^{(i,j)}$ for all $i \neq j$. Besides, as for all distinct $i, j, k \geq 1$, $\mathcal{L}(\bar{X}^i, \bar{X}^j) = \mathcal{L}(\bar{X}^i, \bar{X}^k) = \mathcal{L}(\bar{X}^j, \bar{X}^i)$, we also know that $\pi^i := \pi^{(i,j),1} = \pi^{(i,k),1} = \pi^{(j,i),2}$. As a consequence, the Poisson measures π^i $(i \geq 1)$ are pairwise independent, and so they are also mutually independent, since the independence between Poisson measures is characterized by the fact that their supports (i.e. the sets $\{t \geq 0 : \pi^i(\{t\} \times \mathbb{R}_+) \neq 0\}$) are disjoint (see Theorem II.6.3 of Ikeda and Watanabe (1989)).

We resume the above step. We have just shown that there exist a Brownian motion W and independent Poisson random measures π^i $(i \geq 1)$ with Lebesgue intensity, independent of W, such that, for all $i \geq 1$,

$$d\bar{Y}^i_t = -\alpha\bar{Y}^i_t dt + \sigma\sqrt{\mu_t(f)}dW_t - \bar{Y}^i_{t-}\int_{\mathbb{R}_+} \mathbb{1}_{\{z\leq f(\bar{Y}^i_{t-})\}}\pi^i(dt, dz).$$

As a consequence, $(\bar{Y}^i)_{i\geq 1}$ is solution to (4.12), and Theorem 4.3.6 allows to conclude. $\square$

The last missing point to prove our main result, Theorem 4.4.2, is the following

Theorem 4.4.7. *Assume that Assumptions 4.1, 4.2 and 4.4 hold, and that f is bounded. Then the distribution of any limit μ of the sequence $\mu^N := \frac{1}{N}\sum_{j=1}^N \delta_{X^{N,j}}$ is solution of item (ii) of $(\mathcal{M})$.*

Proof. Step 1. We first check that for any $t\geq 0$, a.s., $\mu(\{\gamma : \Delta\gamma(t)\neq 0\}) = 0$, where $\Delta\gamma(t) := \gamma(t)-\gamma(t-)$. We assume by contradiction that there exists $t>0$ such that $\mu(\{\gamma : \Delta\gamma(t)\neq 0\}) > 0$ with positive probability. Hence there are $a,b>0$ such that the event $E := \{\mu(\{\gamma : |\Delta\gamma(t)| > a\}) > b\}$ has a positive probability. For every $\varepsilon>0$, we have $E\subset\{\mu(\mathcal{B}_a^\varepsilon) > b\}$, where $\mathcal{B}_a^\varepsilon := \{\gamma : \sup_{s\in(t-\varepsilon,t+\varepsilon)}|\Delta\gamma(s)| > a\}$, which is an open subset of $D(\mathbb{R}_+,\mathbb{R})$. Thus $\mathcal{P}_{a,b}^\varepsilon := \{\mu\in\mathcal{P}(D(\mathbb{R}_+,\mathbb{R})) : \mu(\mathcal{B}_a^\varepsilon) > b\}$ is an open subset of $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$. The Portmanteau theorem implies then that for any $\varepsilon>0$,

$$\liminf_{N\to\infty} P(\mu^N\in\mathcal{P}_{a,b}^\varepsilon) \geq P(\mu\in\mathcal{P}_{a,b}^\varepsilon) \geq P(E) > 0. \tag{4.17}$$

Firstly, we can write

$$J^{N,\varepsilon,i} := \sup_{t-\varepsilon<s<t+\varepsilon}\left|\Delta X_s^{N,i}\right| = G_N^{\varepsilon,i}\vee S_N^\varepsilon,$$

where $G_N^{\varepsilon,i} := \max_{s\in D_N^{\varepsilon,i}}|X_{s-}^{N,i}|$ is the maximal height of the big jumps of $X^{N,i}$, with $D_N^{\varepsilon,i} := \{t-\varepsilon\leq s\leq t+\varepsilon : \pi^i(\{s\}\times[0,f(X_{s-}^{N,i})]\times\mathbb{R}_+)\neq 0\}$. Moreover, $S_N^\varepsilon := \max\{|U^j(s)|/\sqrt{N} : s\in\bigcup_{1\leq j\leq N} D_N^{\varepsilon,j}\}$ is the maximal height of the small jumps of $X^{N,i}$, where $U^j(s)$ is defined for $s\in D_N^{\varepsilon,j}$, almost surely, as the only real number that satisfies $\pi^j(\{s\}\times[0,f(X_{s-}^{N,j})]\times\{U^j(s)\}) = 1$.

We have that

$$\left\{\mu^N(\mathcal{B}_a^\varepsilon) > b\right\} = \left\{\frac{1}{N}\sum_{j=1}^N \mathbb{1}_{\{J^{N,\varepsilon,j}>a\}} > b\right\}.$$

Consequently, by exchangeability and Markov's inequality,

$$\mathbb{P}\left(\mu^N(\mathcal{B}_a^\varepsilon) > b\right) \leq \frac{1}{b}\mathbb{E}\left[\mathbb{1}_{\{J^{N,\varepsilon,1}>a\}}\right] = \frac{1}{b}\mathbb{P}\left(J^{N,\varepsilon,1} > a\right) \leq \frac{1}{b}\left(\mathbb{P}\left(G_N^{\varepsilon,1} > a\right) + \mathbb{P}\left(S_N^\varepsilon > a\right)\right). \tag{4.18}$$

The number of big jumps of $X^{N,1}$ in $]t-\varepsilon,t+\varepsilon[$ is smaller than a random variable ξ having Poisson distribution with parameter $2\varepsilon||f||_\infty$. Hence

$$\mathbb{P}\left(G_{N,1}^\varepsilon > a\right) \leq \mathbb{P}\left(\xi\geq 1\right) = 1 - e^{2\varepsilon||f||_\infty} \leq 2\varepsilon||f||_\infty. \tag{4.19}$$

The small jumps that occur in $]t-\varepsilon,t+\varepsilon[$ are included in $\{U_1/\sqrt{N},...,U_K/\sqrt{N}\}$ where K is a $\mathbb{N}-$valued random variable having Poisson distribution with parameter $2\varepsilon N||f||_\infty$, which is independent of the variables U_i $(i\geq 1)$ that are i.i.d. with distribution ν. Hence,

$$\mathbb{P}\left(S_N^\varepsilon > a\right) \leq \mathbb{P}\left(\max_{1\leq i\leq K}\frac{|U_i|}{\sqrt{N}} > a\right) \leq \mathbb{E}\left[\mathbb{P}\left(\max_{1\leq i\leq K}\frac{|U_i|}{\sqrt{N}} > a\,\middle|\,K\right)\right] = \mathbb{E}\left[\psi(K)\right],$$

where $\psi(k) = \mathbb{P}\left(\max_{1\leq i\leq k}|U_i| > a\sqrt{N}\right) \leq k\mathbb{P}\left(|U_1| > a\sqrt{N}\right) \leq ka^{-2}N^{-1}\mathbb{E}\left[U_1^2\right]$. Hence

$$\mathbb{P}\left(S_N^\varepsilon > a\right) \leq \frac{\mathbb{E}\left[U_1^2\right]}{Na^2}\mathbb{E}\left[K\right] \leq 2||f||_\infty\mathbb{E}\left[U_1^2\right]\frac{1}{a}\varepsilon. \tag{4.20}$$

Inserting the bounds (4.19) and (4.20) in (4.18), we have

$$\mathbb{P}\left(\mu^N(\mathcal{B}_a^\varepsilon > b)\right) \leq C\varepsilon,$$

where C does not depend on N nor ε. This last inequality is in contradiction with (4.17) since $P(E)$ does not depend on ε.

Step 2. In the following, we note $\partial^2\varphi := \sum_{i,j=1}^2 \partial^2_{x^i x^j}\varphi$. For any $0 \leq s_1 < \ldots < s_k < s < t$, any $\varphi_1, \ldots, \varphi_k, \psi_1, \ldots, \psi_k \in C_b(\mathbb{R})$, any $\varphi \in C_c^3(\mathbb{R}^2)$, we introduce

$$\begin{aligned}
F(\mu) := \psi_1(\mu_{s_1}(f)) \ldots \psi_k(\mu_{s_k}(f)) \int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu \otimes \mu(d\gamma)\varphi_1(\gamma_{s_1}) \ldots \varphi_k(\gamma_{s_k}) \\
\Big[\varphi(\gamma_t) - \varphi(\gamma_s) + \alpha \int_s^t \gamma_r^1 \partial_{x^1}\varphi(\gamma_r)dr + \alpha \int_s^t \gamma_r^2 \partial_{x^2}\varphi(\gamma_r)dr - \frac{\sigma^2}{2}\int_s^t \mu_r(f)\partial^2\varphi(\gamma_r)dr \\
- \int_s^t f(\gamma_r^1)(\varphi(0,\gamma_r^2) - \varphi(\gamma_r))dr - \int_s^t f(\gamma_r^2)(\varphi(\gamma_r^1, 0) - \varphi(\gamma_r))dr\Big].
\end{aligned}$$

Let us note that F is bounded, because we have chosen a compactly supported test function φ. To show that $\mathcal{L}(\mu)$ is solution of item (ii) of the martingale problem $(\mathcal{M})$, by a classical density argument, it is sufficient to prove that $\mathbb{E}\left[F(\mu)\right] = 0$. We have

$$\begin{aligned}
F(\mu^N) = \psi_1(\mu^N_{s_1}(f)) \ldots \psi_k(\mu^N_{s_k}(f)) \frac{1}{N^2}\sum_{i=1}^N \sum_{j=1}^N \varphi_1(X^{N,i}_{s_1}, X^{N,j}_{s_1}) \ldots \varphi_k(X^{N,i}_{s_k}, X^{N,j}_{s_k}) \cdot \\
\Big[\varphi(X_t^{N,i}, X_t^{N,j}) - \varphi(X_s^{N,i}, X_s^{N,j}) + \alpha \int_s^t X_r^{N,i}\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,j})dr \\
+ \alpha \int_s^t X_r^{N,j}\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,j})dr - \frac{\sigma^2}{2}\int_s^t \mu_r^N(f)\partial^2\varphi(X_r^{N,i}, X_r^{N,j})dr \\
- \int_s^t f(X_r^{N,i})(\varphi(0, X_r^{N,j}) - \varphi(X_r^{N,i}, X_r^{N,j}))dr - \int_s^t f(X_r^{N,j})(\varphi(X_r^{N,i}, 0) - \varphi(X_r^{N,i}, X_r^{N,j}))dr\Big].
\end{aligned}$$

But recalling (4.1) and using Itô's formula, for any $i \neq j$, we have

$$\begin{aligned}
&\varphi(X_t^{N,i}, X_t^{N,j}) \\
&= \varphi(X_s^{N,i}, X_s^{N,j}) - \alpha \int_s^t X_r^{N,i}\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,j})dr - \alpha\int_s^t X_r^{N,j}\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,j})dr \\
&+ \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z \leq f(X^{N,i}_{r-})\right\}} \left[\varphi\left(0, X_{r-}^{N,j} + \frac{u}{\sqrt{N}}\right) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\right] \pi^i(dr,dz,du) \\
&+ \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z \leq f(X^{N,j}_{r-})\right\}} \left[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, 0\right) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\right] \pi^j(dr,dz,du) \\
&+ \sum_{\substack{k=1 \\ k\notin\{i,j\}}}^N \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z \leq f(X^{N,k}_{r-})\right\}} \Big[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, X_{r-}^{N,j} + \frac{u}{\sqrt{N}}\right) \\
&\qquad - \varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\Big] \pi^k(dr,dz,du).
\end{aligned}$$

We use the notation $\tilde{\pi}^j(dr, dz, du) = \pi^j(dr, dz, du) - drdz\nu(du)$ and set

$$
\begin{aligned}
M_{s,t}^{N,i,j,1} :=& \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,i})\right\}} \left[\varphi\left(0, X_{r-}^{N,j} + \frac{u}{\sqrt{N}}\right) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\right] \tilde{\pi}^i(dr, dz, du),\\
M_{s,t}^{N,i,j,2} :=& \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,j})\right\}} \left[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, 0\right) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\right] \tilde{\pi}^j(dr, dz, du),\\
W_{s,t}^{N,i,j} :=& \sum_{\substack{k=1\\ j\notin\{i,j\}}}^{N} \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,k})\right\}} \left[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, X_{r-}^{N,j} + \frac{u}{\sqrt{N}}\right)\right.\\
&\qquad\qquad\left.-\varphi(X_{r-}^{N,i}, X_{r-}^{N,j})\right] \tilde{\pi}^k(dr, dz, du),\\
\Delta_{s,t}^{N,i,j,1} :=& \int_s^t \int_{\mathbb{R}} f(X_r^{N,i}) \left[\varphi\left(0, X_r^{N,j} + \frac{u}{\sqrt{N}}\right) - \varphi(0, X_r^{N,j})\right] d\nu(u)dr,\\
\Delta_{s,t}^{N,i,j,2} :=& \int_s^t \int_{\mathbb{R}} f(X_r^{N,j}) \left[\varphi\left(X_r^{N,i} + \frac{u}{\sqrt{N}}, 0\right) - \varphi(X_r^{N,i}, 0)\right] d\nu(u)dr,\\
\Gamma_{s,t}^{N,i,j} :=& \sum_{\substack{k=1\\ k\notin\{i,j\}}}^{N} \int_s^t \int_{\mathbb{R}} f(X_r^{N,k}) \left[\varphi\left(X_r^{N,i} + \frac{u}{\sqrt{N}}, X_r^{N,j} + \frac{u}{\sqrt{N}}\right) - \varphi(X_r^{N,i}, X_r^{N,j})\right.\\
&\left.-\frac{u}{\sqrt{N}}\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,j}) - \frac{u}{\sqrt{N}}\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,j})\right] d\nu(u)dr\\
&- \int_s^t \int_{\mathbb{R}} \frac{u^2}{2}\partial^2\varphi(X_r^{N,i}, X_r^{N,j})\frac{1}{N} \sum_{\substack{k=1\\ k\notin\{i,j\}}}^{N} f(X_r^{N,k})d\nu(u)dr,\\
R_{s,t}^{N,i,j} :=& \frac{\sigma^2}{2} \int_s^t \partial^2\varphi(X_r^{N,i}, X_r^{N,j}) \left(\frac{1}{N} \sum_{\substack{k=1\\ k\notin\{i,j\}}}^{N} f(X_r^{N,k}) - \frac{1}{N}\sum_{k=1}^{N} f(X_r^{N,k})\right) dr,
\end{aligned}
$$

where the two terms of the second line in the expression of $\Gamma_{s,t}^{N,i}$ can be artificially introduced since $\int_{\mathbb{R}} ud\nu(u) = 0$.

Finally, for $i = j$, we have

$$
\begin{aligned}
\varphi(X_t^{N,i}, X_t^{N,i}) =& \varphi(X_s^{N,i}, X_s^{N,i})\\
&- \alpha \int_s^t X_r^{N,i}\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,i})dr - \alpha \int_s^t X_r^{N,i}\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,i})dr\\
&+ \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,i})\right\}} \left[\varphi(0,0) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,i})\right] \pi^i(dr, dz, du)\\
+ \sum_{\substack{k=1\\ k\neq i}}^{N} \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} &\mathbb{1}_{\left\{z\le f(X_{r-}^{N,k})\right\}} \left[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, X_{r-}^{N,i} + \frac{u}{\sqrt{N}}\right) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,i})\right] \pi^k(dr, dz, du).
\end{aligned}
$$

The associated martingales and error terms are given by

$$
\begin{aligned}
M_{s,t}^{N,i} &:= \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,i})\right\}} \left[\varphi(0,0) - \varphi(X_{r-}^{N,i}, X_{r-}^{N,i})\right] \tilde{\pi}^i(dr,dz,du),\\
W_{s,t}^{N,i} &:= \sum_{\substack{k=1\\k\neq i}}^N \int_{]s,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\left\{z\le f(X_{r-}^{N,k})\right\}} \left[\varphi\left(X_{r-}^{N,i} + \frac{u}{\sqrt{N}}, X_{r-}^{N,i} + \frac{u}{\sqrt{N}}\right)\right.\\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. -\varphi(X_{r-}^{N,i}, X_{r-}^{N,i})\right] \tilde{\pi}^k(dr,dz,du),\\
\Delta_{s,t}^{N,i} &:= \int_s^t \int_{\mathbb{R}} f(X_r^{N,i}) \left[\varphi(0,0) - \varphi(0, X_r^{N,i}) - \varphi(X_r^{N,i}, 0) + \varphi(X_r^{N,i}, X_r^{N,i})\right] d\nu(u)dr,\\
\Gamma_{s,t}^{N,i} &:= \sum_{\substack{k=1\\k\neq i}}^N \int_s^t \int_{\mathbb{R}} f(X_r^{N,k}) \left[\varphi\left(X_r^{N,i} + \frac{u}{\sqrt{N}}, X_r^{N,i} + \frac{u}{\sqrt{N}}\right) - \varphi(X_r^{N,i}, X_r^{N,i})\right.\\
&\qquad \left. -\frac{u}{\sqrt{N}}\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,i}) - \frac{u}{\sqrt{N}}\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,i})\right] d\nu(u)dr\\
&\qquad - \int_s^t \int_{\mathbb{R}} \frac{u^2}{2}\partial^2\varphi(X_r^{N,i}, X_r^{N,i}) \frac{1}{N}\sum_{\substack{k=1\\k\neq i}}^N f(X_r^{N,k}) d\nu(u)dr,\\
R_{s,t}^{N,i} &:= \frac{\sigma^2}{2}\int_s^t \partial^2\varphi(X_r^{N,i}, X_r^{N,i}) \left(\frac{1}{N}\sum_{\substack{k=1\\k\neq i}}^N f(X_r^{N,k}) - \frac{1}{N}\sum_{k=1}^N f(X_r^{N,k})\right) dr,
\end{aligned}
$$

Then, since $\int_{\mathbb{R}} u d\nu(u) = 0$, we obtain

$$
\begin{aligned}
F(\mu^N) = \psi_1(\mu_{s_1}^N(f))\dots\psi_k(\mu_{s_k}^N(f)) \frac{1}{N^2} &\sum_{i,j=1, i\neq j}^N \varphi_1(X_{s_1}^{N,i}, X_{s_1}^{N,j})\dots\varphi_k(X_{s_k}^{N,i}, X_{s_k}^{N,j})\\
&\left[M_{s,t}^{N,i,j,1} + M_{s,t}^{N,i,j,2} + W_{s,t}^{N,i,j} + \Delta_{s,t}^{N,i,j,1} + \Delta_{s,t}^{N,i,j,2} + \Gamma_{s,t}^{N,i,j} + R_{s,t}^{N,i,j}\right]\\
&+ \psi_1(\mu_{s_1}^N(f))\dots\psi_k(\mu_{s_k}^N(f)) \frac{1}{N^2}\sum_{i=1}^N \varphi_1(X_{s_1}^{N,i}, X_{s_1}^{N,i})\dots\varphi_k(X_{s_k}^{N,i}, X_{s_k}^{N,i})\\
&\qquad\qquad\qquad\qquad \left[M_{s,t}^{N,i} + W_{s,t}^{N,i} + \Delta_{s,t}^{N,i} + \Gamma_{s,t}^{N,i} + R_{s,t}^{N,i}\right].
\end{aligned}
$$

Using exchangeability and the boundedness of the φ_j, ψ_j $(1 \le j \le k)$ and the fact that $M^{N,i,j,1}, \dots, W^{N,i}$ are martingales, this implies

$$
|\mathbb{E}\left[F(\mu^N)\right]| \le C\mathbb{E}\left[|\Delta_{s,t}^{N,i,j,1}| + |\Delta_{s,t}^{N,i,j,2}| + |\Gamma_{s,t}^{N,i,j}| + |R_{s,t}^{N,i,j}| + \frac{|\Delta_{s,t}^{N,i}| + |\Gamma_{s,t}^{N,i}| + |R_{s,t}^{N,i}|}{N}\right].
$$

Since f is bounded and $\varphi \in C_b^3(\mathbb{R}^2)$, Taylor-Lagrange's inequality implies then that

$$
|\mathbb{E}\left[F(\mu^N)\right]| \le \frac{C}{\sqrt{N}}.
$$

Finally, using that F is bounded and almost surely continuous at μ (see *Step 1*), we have

$$\mathbb{E}\left[F(\mu)\right] = \lim_{N\to\infty} \mathbb{E}\left[F(\mu^N)\right] = 0,$$

concluding our proof. □

Now we have all elements to give the

Proof of Theorem 4.4.2. According to Proposition 4.4.3, the sequence $(\mu^N)_N$ is tight. Besides, thanks to Theorem 4.4.7, any limit Q of a converging subsequence of $(\mathcal{L}(\mu^N))_N$ is solution to the martingale problem.

By Theorem 4.4.6, there is a unique such distribution Q which can be written as $Q = \mathcal{L}(\mu)$, with $\mu = \mathcal{L}(\bar{X}^1|\mathcal{W})$, where $(\bar{X}^j)_{j\geq 1}$ is solution of (4.2). This implies the result. □

Part II

Propagation of chaos for McKean-Vlasov systems

Introduction of the part

In the previous chapter, we have studied the following nonlinear stochastic differential equation

$$d\bar{X}_t = -\alpha \bar{X}_t dt + \sigma\sqrt{\mu_t(f)} dW_t - \bar{X}_{t-} dZ_t,$$

where $\mu_t(f) = \mathbb{E}\left[f(\bar{X}_t) \middle| W \right]$ and Z is point process with intensity $f(\bar{X}_{t-})$. This equation is not a classical stochastic differential equation. Such an equation can be qualified of conditional McKean-Vlasov equation. In the classical framework of McKean-Vlasov equation, the coefficients of the equation depend both on the solution of the equation and on the law of this solution. Here we talk about conditional McKean-Vlasov equations since μ_t is the conditional law of the solution. McKean-Vlasov equations arise naturally in the framework of $N-$particle systems where the particles interact in a mean field way: this phenomenon can be seen as a generalization of the law of large numbers. Indeed, in examples where the dynamic of the $N-$particle system is described by a stochastic differential equation, the mean field interactions can be expressed as a dependency of the coefficients on the empirical measure of the system. And as N goes to infinity, this empirical measure $\mu^N = N^{-1}\sum_{j=1}^N \delta_{X^{N,j}}$ converges to the law of any particle of the limit system $\mathcal{L}(\bar{X}^1)$. This entails natural dependencies of the coefficients on the law of the solution of the limit equation. For instance, see De Masi et al. (2015) and Fournier and Löcherbach (2016) for examples in neural network modeling, Fischer and Livieri (2016) for an example in portfolio modeling, and Carmona et al. (2016) for an application in mean field games.

The aim of this part is to generalize this type of question. We study the large scale limit of McKean-Vlasov particle systems (i.e. particle systems directed by some stochastic differential equations whose coefficients depend on the empirical measure of the system). In Chapter 5, we study McKean-Vlasov systems in a diffusive regime. The model studied is a generalization of the one studied in Chapter 4. More precisely, we study the large scale limit of the following particle systems.

$$\begin{aligned} dX_t^{N,i} &= b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)d\beta_t^i \\ &+ \frac{1}{\sqrt{N}} \sum_{k=1,k\neq i}^N \int_{\mathbb{R}_+\times E} \Psi(X_{t-}^{N,k}, X_{t-}^{N,i}, \mu_{t-}^N, u^k, u^i)\mathbb{1}_{\left\{z\leq f(X_{t-}^{N,k},\mu_{t-}^N)\right\}} d\pi^k(t,z,u), 1\leq i\leq N, \end{aligned}$$

where $\mu_t^N = N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$, $E = \mathbb{R}^{\mathbb{N}^*}$, with β^i $(i\geq 1)$ are i.i.d. one-dimensional standard Brownian motions and π^k $(k\geq 1)$ are i.i.d. Poisson measures with intensity $ds\cdot dz\cdot \nu(du)$, with ν a product probability measure on $E = \mathbb{R}^{\mathbb{N}^*}$.

To prove the convergence of the systems above, we use similar technics as the ones used in Chapter 4: the main argument of the proof is a particular martingale problem. However there are many additional difficulties in this model compared to the previous one. The major difference is the form of the limit system, where some white noises appear instead of one Brownian motion. Note that, as in Chapter 4, the limit system is characterized by conditional McKean-Vlasov equations (this conditioning appears because of the diffusive scaling).

The last chapter of this part concerns the large scale limit of McKean-Vlasov system, but in the linear scaling (the difference with the equation above is that the scaling $N^{-1/2}$ is replaced by N^{-1}) for coefficients that are only locally Lipschitz. In this case, the limit equation is a classical McKean-Vlasov equation. The technics used in the proofs are very different from the technics of the other chapters, since in this last chapter, we prove a L^1 convergence and not a convergence in distribution.

We deal with the fact that the coefficients are only locally Lipschitz using Osgood's lemma instead of Grönwall's lemma. The other difficulties that appear in this framework are explained at the beginning of Chapter 6.

Note that the question of propagation of chaos in this model has already been studied by Andreis et al. (2018). The novelty of our results is that we do not work with globally Lipschitz coefficients. The differences between Andreis et al. (2018) and Chapter 6 are explained in details at the beginning of the chapter. In Chapter 4, we already manipulated locally Lipschitz coefficients of the form $x \mapsto -xf(x)$ in the compensator of the reset jump term. We overcomed this difficulty using an appropriate metric and an appropriate assumption on the function f. In Chapter 6, we work with generic assumptions.

Chapter 5

White-noise driven conditional McKean-Vlasov limits for systems of particles with simultaneous and random jumps

This chapter is based on Erny et al. (2021).

In this chapter, we prove the conditional propagation of chaos of the following system

$$
\begin{aligned}
dX_t^{N,i} &= b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)d\beta_t^i \\
&+ \frac{1}{\sqrt{N}} \sum_{k=1, k\neq i}^{N} \int_{\mathbb{R}_+ \times E} \Psi(X_{t-}^{N,k}, X_{t-}^{N,i}, \mu_{t-}^N, u^k, u^i) \mathbb{1}_{\left\{z \leq f(X_{t-}^{N,k}, \mu_{t-}^N)\right\}} d\pi^k(t,z,u), 1 \leq i \leq N,
\end{aligned}
\tag{5.1}
$$

starting from the initial condition

$$
(X_0^{N,i})_{1\leq i\leq N} \sim \nu_0^{\otimes N},
$$

where $\mu_t^N = N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$, $E = \mathbb{R}^{\mathbb{N}^*}$, with $\mathbb{N}^* = \{1, 2, 3, \ldots\}$, β^i $(i \geq 1)$ are i.i.d. one-dimensional standard Brownian motions and π^k $(k \geq 1)$ are i.i.d. Poisson measures with intensity $ds \cdot dz \cdot \nu(du)$, with ν a product probability measure on $E = \mathbb{R}^{\mathbb{N}^*}$, and ν_0 a distribution on $\mathbb{R}$ having a finite second moment. We assume that the Poisson measures, the Brownian motions and the initial conditions are independent. Since the jumps are scaled in $N^{-1/2}$, an important assumption of our model (see Assumption 5.3 below) is that the height of the jump term above is centered. That is, for all $x, y \in \mathbb{R}$, for all probability measures m on $\mathbb{R}$,

$$
\int_E \Psi(x, y, m, u^1, u^2) d\nu(u) = 0.
$$

So, in this model, the heights of the jumps depend on the states of both the giving and the receiving particles, on the average state of the system and also on their "noise" parameters.

Notice that contrarily to our previous model studied in Chapter 4 we do only consider small jumps affecting the other particles; that is, we decided not to include auto-interactions induced by jumps, i.e. terms of the type

$$\int_{\mathbb{R}_+\times E} \Theta(X_{t-}^{N,i}, \mu_{t-}^N, u^i) \mathbb{1}_{\{z \le f(X_{t-}^{N,i}, \mu_{t-}^N)\}} d\pi^i(t,z,u)$$

in (5.1). Indeed, such terms would survive in the large population limit leading to discontinuous trajectories, and the presence of the indicator $\mathbb{1}_{\{z \le f(X_{t-}^{N,i}, \mu_{t-}^N)\}}$ requires to work both in L^1 and in L^2 (see Graham (1992) or Chapter 4). In the present chapter, we decided to disregard these big jumps to focus on the very specific form of the limit process given in (5.3) below.

It has already been observed that this diffusive scaling gives rise to the conditional propagation of chaos (see Chapter 4), where the limit of the empirical measures is shown to be the conditional distribution of any coordinate of the limit system. It turns out that to describe the precise dynamic of the limit, we need to rely on martingale measures and white noises instead of Brownian motions (see Walsh (1986) and El Karoui and Méléard (1990)) as driving measures. More precisely, the limit system will be shown to be solution of a nonlinear stochastic differential equation driven by a white noise having an intensity that depends on the conditional law of the system itself.

Such processes have already appeared in the literature related to particle approximations of Boltzmann's equation, starting with the classical article by Tanaka (1978) that gave rise to a huge literature (to cite juste a few, see Graham and Méléard (1997), Méléard (1998), Fournier and Meleard (2002), Fournier and Mischler (2016)). In these papers, the underlying random measure is Poisson, and the dependence on the law arises at the level of the particle system that is designed to approximate Boltzmann's equation. In our work, the underlying random measure is white noise since jumps disappear in the limit, and the dependence on the (conditional) law of the process does only appear in the limit, as an effect of the conditional propagation of chaos.

To understand the form of the limit system, we start investigating the possible form of the limits of the correlations of the finite systems, that is, of the predictable quadratic covariation of the particles. In (5.1), the only term that we have to study is the jump term. Hence, in the following heuristics, we only consider the predictable quadratic covariation of this jump term in order to explain the form of the limit system. Let us note this term $J_t^{N,i}$, it is given by

$$J_t^{N,i} = \frac{1}{\sqrt{N}} \sum_{k=1, k\neq i}^N \int_{[0,t]\times\mathbb{R}_+\times E} \Psi(X_{s-}^{N,k}, X_{s-}^{N,i}, \mu_{s-}^N, u^k, u^i) \mathbb{1}_{\{z \le f(X_{s-}^{N,k}, \mu_{s-}^N)\}} d\pi^k(s,z,u).$$

In what follows we consider some examples to have a better understanding of the limit system. We shall always assume that the jump rate function f is bounded. Let us begin with a situation close to that of Chapter 4, where in the limit system each coordinate shares a common Brownian motion W.

Example 5.0.1. *Suppose that $\Psi(x,y,m,u^1,u^2) = \Psi(u^1)$. Then we have, for all $1 \le i,j \le N$,*

$$\langle J^{N,i}, J^{N,j}\rangle_t = \frac{1}{N} \sum_{k=1, k\neq i,j}^N \int_0^t \int_{\mathbb{R}} \Psi(u^k)^2 f(X_s^{N,k}, \mu_s^N) d\nu_1(u^k) ds$$

$$= \varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s^N) \mu_s^N(dx) ds + O(\frac{t}{N}),$$

since f is bounded, with $\varsigma^2 := \int_{\mathbb{R}} \Psi(u^1)^2 d\nu_1(u^1)$ and ν_1 the projection of ν on the first coordinate. These angle brackets should converge as N goes to infinity to

$$\varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s)\mu_s(dx)ds,$$

where μ is the limit of the empirical measures μ^N. As the quadratic variations are non-null, there should be, in the limit system, a common Brownian motion W underlying each particle's motion. Thus, the limit system is given by

$$d\bar{X}_t^i = b(\bar{X}_t^i, \mu_t)dt + \sigma(\bar{X}_t^i, \mu_t)d\beta_t^i + \varsigma\sqrt{\int_{\mathbb{R}} f(x, \mu_t)\mu_t(dx)}dW_t,$$

where W is a standard one-dimensional Brownian motion. We will also show that $\mu = \mathcal{L}(\bar{X}^1|W)$, since μ is necessarily the directing measure of $(\bar{X}^i)_{i\geq 1}$. In particular, the conditioning in μ reflects the presence of some common noise, which is W here.

Now, let us consider an opposite situation where, in the limit system, each coordinate has its own Brownian motion W^i, and where these Brownian motions are independent.

Example 5.0.2. *In this example, we assume that $\Psi(x, y, m, u^1, u^2) = \Psi(u^2)$. As in the previous example, we begin by computing the angle brackets of the jump terms between particles i and j. Here we distinguish two cases: $i \neq j$ and $i = j$. If $i \neq j$, since ν is of product form and since jumps are centered,*

$$\langle J^{N,i}, J^{N,j}\rangle_t = \frac{1}{N} \sum_{k=1, k\neq i,j}^{N} \int_0^t \int_E \Psi(u^i)\Psi(u^j) f(X_s^{N,k}, \mu_s^N) d\nu(u)ds = 0.$$

Moreover, if $i = j$,

$$\langle J^{N,i}\rangle_t = \frac{1}{N} \sum_{k=1, k\neq i}^{N} \int_0^t \int_{\mathbb{R}} \Psi(u^i)^2 f(X_s^{N,k}, \mu_s^N) d\nu_1(u^i)ds = \varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s^N)\mu_s^N(dx)ds + O(\frac{t}{N}).$$

As the quadratic variation between different particles is null, there will be no common noise in the limit system. So, instead of having one common Brownian motion W as in the previous example, here, each particle is driven by its own Brownian motion. More precisely, in this example the limit system is

$$d\bar{X}_t^i = b(\bar{X}_t^i, \mu_t)dt + \sigma(\bar{X}_t^i, \mu_t)d\beta_t^i + \varsigma\sqrt{\int_{\mathbb{R}} f(x, \mu_t)\mu_t(dx)}dW_t^i,$$

where W^i $(i \geq 1)$ are independent standard one-dimensional Brownian motions, independent of β^i $(i \geq 1)$, and where $\mu = \mathcal{L}(\bar{X}^1)$ is deterministic in this particular case.

Finally let us show an example where, as in Example 5.0.1, each particle shares a common Brownian motion W, and, as in Example 5.0.2, each particle has also its own Brownian motion W^i, and where both W and W^i are produced by the small jumps.

Example 5.0.3. *Here we assume that* $\Psi(x,y,m,u^1,u^2)=\Psi(u^1,u^2)$. *The angle brackets of the jump terms of the particles* i *and* j *are, if* $i\neq j$,

$$\langle J^{N,i},J^{N,j}\rangle_t=\frac{1}{N}\sum_{k=1,k\neq i,j}^{N}\int_0^t\int_E\Psi(u^k,u^i)\Psi(u^k,u^j)f(X_s^{N,k},\mu_s^N)d\nu(u)ds$$
$$=\kappa^2\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\mu_s^N(dx)ds+O(\frac{t}{N}),$$

where we know that $\kappa^2:=\int_E\Psi(u^1,u^2)\Psi(u^1,u^3)d\nu(u)\geq 0$ *since it is the covariance of the infinite exchangeable sequence* $(\Psi(U^1,U^k))_{k\geq 2}$, *where* $(U^k)_{k\geq 1}\sim\nu$. *And if* $i=j$,

$$\langle J^{N,i}\rangle_t=\frac{1}{N}\sum_{k=1,k\neq i}^{N}\int_0^t\int_E\Psi(u^i,u^k)^2f(X_s^{N,k},\mu_s^N)d\nu(u)ds=\varsigma^2\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\mu_s^N(dx)ds+O(\frac{t}{N}),$$

where $\varsigma^2=\int_E\Psi(u^1,u^2)^2d\nu(u)$.

As in Example 5.0.1, there must be a common Brownian motion since the quadratic variation between different particles is not zero. But, here $\langle J^{N,i},J^{N,j}\rangle_t\neq\langle J^{N,i}\rangle_t$ *if* $j\neq i$. *That is why there must be additional Brownian motions. Formally, the limit system in this example is*

$$\begin{aligned}d\bar{X}_t^i=&b(\bar{X}_t^i,\mu_t)dt+\sigma(\bar{X}_t^i,\mu_t)d\beta_t^i\\&+\kappa\sqrt{\int_{\mathbb{R}}f(x,\mu_t)\mu_t(dx)}dW_t+\sqrt{(\varsigma^2-\kappa^2)\int_{\mathbb{R}}f(x,\mu_t)\mu_t(dx)}dW_t^i,\end{aligned}\tag{5.2}$$

where we know that $\varsigma^2\geq\kappa^2$ *by Cauchy-Schwarz's inequality. As before,* W,W^i $(i\geq 1)$ *are independent standard one-dimensional Brownian motions.*

Before defining the limit system in the general case, let us explain the main difficulty that arises. If we apply the same reasoning as in Examples 5.0.1, 5.0.2 and 5.0.3 to the general model given in (5.1), we obtain a compensator of the jump terms between different particles i and j of the form

$$\frac{1}{N}\sum_{k=1,k\neq i,j}^{N}\int_0^t f(X_s^{N,k},\mu_s^N)\int_E\Psi(X_s^{N,k},X_s^{N,i},\mu_s^N,u^k,u^i)\Psi(X_s^{N,k},X_s^{N,j},\mu_s^N,u^k,u^j)\nu(du)ds$$
$$=\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\int_E\Psi(x,X_s^{N,i},\mu_s^N,u^1,u^2)\Psi(x,X_s^{N,j},\mu_s^N,u^1,u^3)\nu(du)\mu_s^N(dx)ds+O(\frac{t}{N}),$$

under appropriate conditions on ψ, see Assumption 5.3 below, and for the same particle

$$\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\int_E\Psi(x,X^{N,i},\mu_s^N,u^1,u^2)^2\nu(du)\mu_s^N(dx)ds,$$

still up to an error term of order $1/N$.

Contrarily to the situation of the previous examples, the quadratic variations depend on the positions of the particles i and j and cannot be written as products, but as integrals of products. This integration involves, among others, the empirical measure of the process. This is the reason why we need to use martingale measures and white noises instead of Brownian motions, as introduced in Walsh (1986), confer also to El Karoui and Méléard (1990).

Let us briefly explain why using martingale measures is well adapted to our problem. If M is a martingale measure on $\mathbb{R}_+ \times F$ (with $(F, \mathcal{F})$ some measurable space), having intensity $dt \cdot m_t(dy)$, then for all $A, B \in \mathcal{F}$,

$$\langle M_\cdot(A), M_\cdot(B)\rangle_t = \int_0^t \int_F \mathbb{1}_{A\cap B}(y) m_s(dy) ds.$$

Having this remark in mind, it is natural to write the limit system in a similar way as (5.2), but replacing the Brownian motions by martingale measures. More precisely, under appropriate conditions on the coefficients, the limit system $(\bar{X}^i)_{i\geq 1}$ of this chapter will be shown to be of the form

$$\begin{aligned} d\bar{X}^i_t &= b(\bar{X}^i_t, \mu_t)dt + \sigma(\bar{X}^i_t, \mu_t)d\beta^i_t + \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\tilde{\Psi}(x, \bar{X}^i_t, \mu_t, v)dM(t,x,v), \quad (5.3)\\ &\quad + \int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\kappa(x, \bar{X}^i_t, \mu_t)dM^i(t,x), \ i\geq 1,\\ (\bar{X}^i_0)_{i\geq 1} &\sim \nu_0^{\otimes \mathbb{N}^*}. \end{aligned}$$

In the above formula,

$$\mu_t := \mathcal{L}(\bar{X}^i_t | W), \quad (5.4)$$

$$\tilde{\Psi}(x,y,m,v) := \int_{\mathbb{R}} \Psi(x,y,m,v,u^2)d\nu_1(u^2), \quad (5.5)$$

$$\begin{aligned} \kappa(x,y,m)^2 &:= \int_E \Psi(x,y,m,u^1,u^2)^2 d\nu(u) - \int_{\mathbb{R}} \tilde{\Psi}(x,y,m,v)^2 d\nu_1(v)\\ &= \int_E \Psi(x,y,m,u^1,u^2)^2 d\nu(u) - \int_E \Psi(x,y,m,u^1,u^2)\Psi(x,y,m,u^1,u^3)d\nu(u). \quad (5.6) \end{aligned}$$

Notice that the expression (5.6) is positive by Cauchy-Schwarz's inequality.

In the above equations, $M(dt,dx,dv)$ and $M^i(dt,dx)$ are orthogonal martingale measures on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$ ($\mathbb{R}_+ \times \mathbb{R}$ respectively) with respective intensities $dt \cdot \mu_t(dx) \cdot \nu_1(dv)$ and $dt \cdot \mu_t(dx)$, defined as

$$M^i_t(A) := \int_0^t \mathbb{1}_A(F_s^{-1}(p))dW^i(s,p) \text{ and } M_t(A\times B) := \int_0^t \mathbb{1}_A(F_s^{-1}(p))\mathbb{1}_B(v)dW(s,p,v), \quad (5.7)$$

with W a white noise on $\mathbb{R}_+^2 \times \mathbb{R}$ with intensity $dt \cdot dp \cdot d\nu_1(v)$, and W^i $(i \geq 1)$ independent white noises on $\mathbb{R}_+^2$, also independent from W, with intensity $dt \cdot dp$. In the above formula, $F_s(x) := \mathbb{P}\left(\bar{X}^i_s \leq x | W\right)$ and F_s^{-1} is the generalized inverse of F_s. As in (5.1), we assume that the Brownian motions, the white noises and the initial conditions are independent.

Let us give some comments on the above system of equations. We have already argued that, in general, μ_t is a random measure because of the scaling $N^{-1/2}$. We shall prove that μ_t is actually the law of $\bar{X}^1$ conditionally on the common noise of the system. This common noise is the white noise W underlying the martingale measure M. Therefore, it is not obvious that the definition of the martingale measures M and M^i in (5.7) and the limit system (5.3) are well-posed. In what follows, we shall give conditions ensuring that equation (5.3) admits a unique strong solution. This is the content of our first main theorem, Theorem 5.3.1. To prove this theorem, we propose a Picard iteration in which we construct a sequence of martingale measures whose intensities depend on the conditional law of the instance of the process within the preceding step. One main ingredient of the proof is the well-known fact that the Wasserstein-2−distance of the laws of two real-valued random variables is given by the L^2- distance of their inverse distribution functions - we apply this fact here to the conditional distribution functions.

Using arguments that are close to those of Chapter 4, we then show in our second main theorem, Theorem 5.4.2, that the finite particle system converges to the limit system, that is, $(X^{N,i})_{1\leq i\leq N}$ converges to $(\bar{X}^i)_{i\geq 1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$. This convergence is the consequence of the well-posedness of an associated martingale problem.

5.1 Assumptions and main results.

We start imposing a hypothesis under which equation (5.1) admits a unique strong solution and which grants a Lipschitz condition on the coefficients of the stochastic differential equation.

Assumption 5.1.

i) For all $x,y\in\mathbb{R}, m,m'\in\mathcal{P}_1(\mathbb{R})$,

$$|b(x,m)-b(y,m')|+|\sigma(x,m)-\sigma(y,m')|\leq C(|x-y|+W_1(m,m')).$$

ii) f *is bounded and strictly positive, and* $\sqrt{f}$ *is Lipschitz, that is, for all* $x,y\in\mathbb{R}, m,m'\in\mathcal{P}_1(\mathbb{R})$,

$$|\sqrt{f}(x,m)-\sqrt{f}(y,m')|\leq C(|x-y|+W_1(m,m')).$$

iii) For all $x,x',y,y',u,v\in\mathbb{R}, m,m'\in\mathcal{P}_1(\mathbb{R})$,

$$|\Psi(x,y,m,u,v)-\Psi(x',y',m',u,v)|\leq M(u,v)(|x-x'|+|y-y'|+W_1(m,m')),$$

where $M:\mathbb{R}^2\to\mathbb{R}_+$ *satisfies* $\int_E M(u^1,u^2)^2 d\nu(u)<\infty$.

iv)

$$\sup_{x,y,m}\int_E|\Psi(x,y,m,u^1,u^2)|d\nu(u)<\infty.$$

Notice that f bounded together with $\sqrt{f}$ Lipschitz implies that f is Lipschitz as well. As a consequence, relying on Theorem 2.1 of Graham (1992), Assumption 5.1 implies that equation (5.1) admits a unique strong solution.

In order to prove the well-posedness of the limit equation (5.3), we need additional assumptions. Recall that κ^2 has been introduced in (5.6) above.

Assumption 5.2.

i)

$$\inf_{x,y,m}\kappa(x,y,m)>0,$$

ii)

$$\sup_{x,y,m} \int_E \Psi(x,y,m,u^1,u^2)^2 d\nu(u) < \infty.$$

Remark 5.1.1. *Using the third point of Assumption 5.1 we can prove that κ^2 is Lipschitz continuous with Lipschitz constant proportional to*

$$\left(\int_E M^2(u,v) d\nu(u,v) \times \sup_{x,y,m} \int_E \Psi^2(x,y,m,u,v) d\nu(u,v) \right)^{1/2}.$$

Assumption 5.2.i) allows then to prove that κ is Lipschitz continuous. Assumption 5.2.ii) gives that $\|\kappa\|_\infty := \sup_{x,y,m} \kappa(x,y,m) < \infty$.

To prove the convergence of the particle system $(X^{N,i})_{1 \le i \le N}$ to the limit system, we need further assumptions on the function Ψ.

Assumption 5.3.

i) For all $x, y \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})$,

$$\int_E \Psi(x,y,m,u^1,u^2) d\nu(u) = 0,$$

ii)

$$\int_E \sup_{x,y,m} |\Psi(x,y,m,u^1,u^2)|^3 d\nu(u) < \infty.$$

iii) b and σ are bounded.

Remark 5.1.2. *We assume the functions b and σ to be bounded to simplify the proofs of Lemmas 5.2.2 and 5.4.7. However the result of these lemmas still hold true under the following weaker assumption: there exists $C > 0$ such that, for all $x \in \mathbb{R}, m \in \mathcal{P}(\mathbb{R})$,*

$$|b(x,m)| + |\sigma(x,m)| \le C(1 + |x|).$$

In other words, b and σ are bounded w.r.t. the measure variable and sublinear w.r.t. the space variable.

Let us give an example of a function Ψ that satisfies all our assumptions and where the random quantity Ψ depends on the difference of the states of the jumping and the receiving particle as well as on the average state of the system as follows

$$\Psi(x,y,m,u,v) = uv \left(\varepsilon + \frac{\pi}{2} + \arctan(x - y + \int_{\mathbb{R}} z dm(z)) \right),$$

with $\nu = \mathcal{R}^{\otimes \mathbb{N}^*}$ and $\mathcal{R} = \frac{1}{2}(\delta_{-1} + \delta_{+1})$ the Rademacher distribution. In the formula above, the variables u and v can be seen as spins such that the receiving particle is excited if the orientation of its spin is the same as the spin of the sending particle, and inhibited otherwise. Ψ satisfies the Lipschitz condition of Assumption 5.1 because arctan is Lipschitz continuous. The hypothesis on the moments of Ψ are also satisfied since $\mathcal{R}$ has finite third moments and is centered. Finally, the first point of Assumption 5.2 holds true, because

$$
\begin{aligned}
\kappa(x,y,m)^2 &= \int_E \Psi(x,y,m,u^1,u^2)^2 d\nu(u) - \int_E \Psi(x,y,m,u^1,u^2)\Psi(x,y,m,u^1,u^3) d\nu(u) \\
&= \left(\varepsilon + \frac{\pi}{2} + \arctan(x-y+\int_{\mathbb{R}} z dm(z))\right)^2 \left(\int_E (u^1u^2)^2 d\nu(u) - \int_E u^1u^2u^1u^3 d\nu(u)\right) \\
&= \left(\varepsilon + \frac{\pi}{2} + \arctan(x-y+\int_{\mathbb{R}} z dm(z))\right)^2 .
\end{aligned}
$$

5.2 Auxiliary results

Let us begin with a result that we use implicitly in the computations. The following Lemma shows that the definition (5.7) indeed defines the right martingale measures.

Lemma 5.2.1. *Let ν_1 be a probability measure on $\mathbb{R}$. Moreover, let $(\Omega, \mathcal{A}, P)$ be a probability space, $(\mathcal{G}_t)_t$ be a filtration on it, $(\mathcal{F}_t)_t$ be a sub-filtration of $(\mathcal{G}_t)_t$ and W (resp. W^1) be a $(\mathcal{G}_t)_t-$white noise on $\mathbb{R}_+ \times [0,1] \times \mathbb{R}$ (resp. $\mathbb{R}_+ \times [0,1]$) of intensity $dt \cdot dp \cdot \nu_1(dv)$ (resp. $dt \cdot dp$). Let X be a continuous $\mathbb{R}-$valued process which is $(\mathcal{G}_t)_t$-adapted and $F_s(x) := \mathbb{P}(X_s \leq x | \mathcal{F}_s)$. Moreover, we suppose that for all $s > 0$, $\mathbb{P}(X_s \leq x|\mathcal{F}_s) = \mathbb{P}(X_s \leq x | \mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \sigma\{\mathcal{F}_t,\ t \geq 0\}$.*

Define for any $A, B \in \mathcal{B}(\mathbb{R})$,

$$
M^1_t(A) := \int_0^t \int_0^1 \mathbb{1}_A((F_s)^{-1}(p)) dW^1(s,p), \quad M_t(A \times B) := \int_0^t \int_0^1 \int_{\mathbb{R}} \mathbb{1}_A((F_s)^{-1}(p)) \mathbb{1}_B(v) dW(s,p,v)
$$

Then M^1 and M are martingale measures with respective intensities $dt \cdot \mu_t(dx)$ and $dt \cdot \mu_t(dx) \cdot \nu_1(dv)$, where $\mu_t := \mathcal{L}(X_t|\mathcal{F}_t)$.

Proof. We only show the result for the martingale measure M^1. The main part of the proof consists in showing that the process $(\omega, s, p) \in \Omega \times \mathbb{R}_+ \times [0,1] \mapsto (F_s)^{-1}(p)$ is $\mathcal{P} \otimes \mathcal{B}([0,1])-$measurable, with $\mathcal{P}$ the predictable sigma field related to the filtration $(\mathcal{G}_t)_t$.

To begin with, let us prove that $(\omega, s, x) \mapsto F_s(x)$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})-$measurable. We write

$$
F_s(x) = P(X_s \leq x | \mathcal{F}_s) = \mathbb{E}\left[\varphi(x, X_s) | \mathcal{F}_s\right],
$$

where $\varphi(x,y) := \mathbb{1}_{\{x \leq y\}}$. As φ is product measurable and bounded, it is the limit of functions of the form

$$
\sum_{k=1}^n c_k \varphi_k(x) \psi_k(y),
$$

where the functions φ_k, ψ_k $(1 \leq k \leq n)$ are Borel measurable and bounded. This limit can be taken to be increasing such that, by monotone convergence,

$$
\mathbb{E}\left[\varphi(x, X_s)|\mathcal{F}_s\right] = \lim_n \sum_{k=1}^n c_k \varphi_k(x) \mathbb{E}\left[\psi_k(X_s)|\mathcal{F}_s\right].
$$

Then, as ψ_k is a bounded and Borel function, it can be approximated by an increasing sequence of bounded and continuous functions $\psi_{k,m}$. Then, as for every n, m, the process

$$
(\omega, s, x) \mapsto \sum_{k=1}^n c_k \varphi_k(x) \mathbb{E}\left[\psi_{k,m}(X_s)|\mathcal{F}_s\right] = \sum_{k=1}^n c_k \varphi_k(x) \mathbb{E}\left[\psi_{k,m}(X_s)|\mathcal{F}_\infty\right]
$$

is continuous in s and $(\mathcal{F}_s)_s-$, whence $(\mathcal{G}_s)_s-$adapted, it is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})-$measurable.

Let $x\in\mathbb{R}$ be fixed. It is sufficient to show that $\{(\omega,s,p):(F_s)^{-1}(p)\geq x\}$ is measurable. Let us write

$$\{(\omega,s,p):(F_s)^{-1}(p)\geq x\}=\{(\omega,s,p):F_s(x)\leq p\}=\{(\omega,s,p):\varphi(F_s(x),p)>0\},$$

where $\varphi(x,p):=\mathbb{1}_{\{x\leq p\}}$ is product measurable.

Then, the measurability of $\{(\omega,s,p):(F_s)^{-1}(p)\geq x\}$ is a consequence of that of $(\omega,s,p)\mapsto\varphi(F_s(x),p)$ w.r.t. $\mathcal{P}\otimes\mathcal{B}([0,1])$.

As a consequence, the process $(\omega,s,p)\in\Omega\times\mathbb{R}_+\times[0,1]\mapsto(F_s)^{-1}(p)$ is $\mathcal{P}\otimes\mathcal{B}([0,1])-$measurable. The rest of the proof consists in writing

$$\langle M^1_\cdot(A)\rangle_t=\int_0^t\int_0^1\mathbb{1}_A((F_s)^{-1}(p))dpds=\int_0^t\mu_s(A)ds.$$

The last inequality above is a classical property of the generalized inverse of the distribution function (see e.g. Fact 1 in Section 8 of Major (1978)). □

Now we prove some a priori estimates for our model.

Lemma 5.2.2. *Grant Assumptions 5.1, 5.2 and 5.3. For all $T>0$,*

$$\sup_{N\in\mathbb{N}^*}\mathbb{E}\left[\sup_{t\leq T}\left|X^{N,1}_t\right|^2\right]<\infty.$$

Proof. Notice that

$$\sup_{0\leq s\leq t}|X^{N,1}_s|\leq(X^{N,1}_0)+||b||_\infty t+\sup_{0\leq s\leq t}\left|\int_0^s\sigma(X^{N,1}_r,\mu^N_r)d\beta^1_r\right|+\frac{1}{\sqrt{N}}\sup_{0\leq s\leq t}|M^N_s|,$$

where M^N is the local martingale

$$M^N_t:=\sum_{k=2}^N\int_{[0,t]\times\mathbb{R}_+\times E}\Psi(X^{N,k}_{s-},X^{N,1}_{s-},\mu^N_{s-},u^k,u^1)\mathbb{1}_{\left\{z\leq f(X^{N,k}_{s-},\mu^N_{s-})\right\}}d\pi^k(s,z,u).$$

Consequently, by Burkholder-Davis-Gundy's inequality and Assumption 5.3,

$$\mathbb{E}\left[\sup_{0\leq s\leq t}|X^{N,1}_s|^2\right]\leq C+C||b||^2_\infty t^2+||\sigma||^2_\infty t+t||f||_\infty\frac{N-1}{N}\int_E\sup_{x,y,m}\Psi(x,y,m,u^1,u^2)^2d\nu(u).$$

This proves the result. □

5.3 Well-posedness of the limit equation

The aim of this section is to prove

Theorem 5.3.1. *Under Assumptions 5.1 and 5.2, equation (5.3) admits a unique strong solution $\bar{X}^i$ that possesses finite second moments. This solution also has finite fourth moments.*

We prove Theorem 5.3.1 in the two following sections. In the first one, we construct a strong solution of (5.3), and in the second one, we prove the pathwise uniqueness property for this equation.

5.3.1 Construction of a strong solution of (5.3) - proof of Theorem 5.3.1

We construct a strong solution of (5.3) using a Picard iteration. Let $\beta^i, i \in \mathbb{N}^*$, be independent one-dimensional Brownian motions. Let $W, W^i,\ i \in \mathbb{N}^*$, be independent white noises on respectively $\mathbb{R}_+ \times [0,1] \times \mathbb{R}$ and $\mathbb{R}_+ \times [0,1]$ with respective intensities $dt \cdot dp \cdot \nu_1(dv)$ and $dt \cdot dp$, independent of the β^i. We suppose that all these processes are defined on the same probability space $(\Omega, \mathcal{F}, P)$ carrying also i.i.d. random variables $X_0^i,\ i \in \mathbb{N}^*$, which are independent of the β^i, W, W^i. Define for all $t \geq 0$,

$$\mathcal{G}_t := \sigma\{\, \beta_v^i;\ W(]u,v] \times A \times B);\ W^i(]u,v] \times A);\ 0 < u < v \leq t;\ A \in \mathcal{B}(\mathbb{R});\ B \in \mathcal{B}([0,1]);\ i \in \mathbb{N}^* \,\};$$

$$\mathcal{W}_t := \sigma\{\, W(]u,v] \times A \times B); u < v \leq t;\ \ A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{B}([0,1]) \,\};$$

$$\mathcal{W} := \sigma\{W_t;\ t \geq 0\}.$$

Step 1. Fix an $i \in \mathbb{N}^*$, and introduce

$$\begin{aligned}
X_t^{i,[0]} :=& X_0^i,\\
\mu_t^{[0]} :=& \mathcal{L}(X_0^i),\ F_t^{[0]}(x) := \mathbb{P}\left(X_0 \leq x | \mathcal{W}\right),\\
M^{[0]}([0,t] \times A \times B) :=& \int_0^t \int_0^1 \int_{\mathbb{R}} \mathbb{1}_A((F_s^{[0]})^{-1}(p)) \mathbb{1}_B(v) dW(s,p,v)\\
M^{i,[0]}([0,t] \times A) :=& \int_0^t \int_0^1 \mathbb{1}_A((F_s^{[0]})^{-1}(p)) dW^i(s,p).
\end{aligned}$$

Assuming everything is defined at order $n \in \mathbb{N}$, we introduce

$$\begin{aligned}
X_t^{i,[n+1]} :=& \int_0^t b(X_s^{i,[n]}, \mu_s^{[n]}) ds + \int_0^t \sigma(X_s^{i,[n]}, \mu_s^{[n]}) d\beta_s^i \qquad (5.8)\\
&+ \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} \sqrt{f(x, \mu_s^{[n]})}\, \tilde{\Psi}(x, X_s^{i,[n]}, \mu_s^{[n]}, v) dM^{[n]}(s,x,v)\\
&+ \int_0^t \int_{\mathbb{R}} \sqrt{f(x, \mu_s^{[n]})}\, \kappa(x, X_s^{i,[n]}, \mu_s^{[n]}) dM^{i,[n]}(s,x),\\
\mu_s^{[n+1]} :=& \mathcal{L}(X_s^{i,[n+1]} | \mathcal{W}),\ F_s^{[n+1]}(x) := \mathbb{P}(X_s^{i,[n+1]} \leq x | \mathcal{W}),\\
M^{[n+1]}([0,t] \times A \times B) :=& \int_0^t \int_0^1 \int_{\mathbb{R}} \mathbb{1}_A((F_s^{[n+1]})^{-1}(p)) \mathbb{1}_B(v) dW(s,p,v),\\
M^{i,[n+1]}([0,t] \times A) :=& \int_0^t \int_0^1 \mathbb{1}_A((F_s^{[n+1]})^{-1}(p)) dW^i(s,p).
\end{aligned}$$

Note that $\forall t > 0$,

$$\mathcal{W} = \sigma(\mathcal{W}_t;\ \mathcal{W}_{]t;\infty[})$$

where

$$\mathcal{W}_{]t;\infty[} := \sigma\{]W(]u,v] \times A);\ \ , t < u < v;\ A \in \mathcal{B}(\mathbb{R})\}.$$

Remember also that white noises are processes with independent increments, more precisely, for all A, A' in $\mathcal{B}([0,1])$, for all B, B' in $\mathcal{B}(\mathbb{R})$, $W(]u,v] \times A \times B)$ and $W(]u',v'] \times A' \times B')$ are independent if $]u;v] \cap]u';v'] = \emptyset$.

Using this last remark, we see that by construction $X_t^{i,[n+1]}$ is independent from $\mathcal{W}_{]t;\infty[}$, and as a consequence, $\mathbb{P}(X_s^{i,[n+1]} \leq x|\mathcal{W}) = \mathbb{P}(X_s^{i,[n+1]} \leq x|\mathcal{W}_s)$. Taking $\mathcal{F}_t = \mathcal{W}_t$ and $\mathcal{F}_\infty = \mathcal{W}$ we see that all assumptions of Lemma (5.2.1) are satisfied. Hence, for each $n \in \mathbb{N}$; $i \in \mathbb{N}^*$ the martingale measures $M^{[n]}$ and $M^{i,[n]}$ are well defined and have respectively the intensities $dt \cdot \mu_t^{[n]}(dx) \cdot \nu_1(dv)$ and $dt \cdot \mu_t^{[n]}(dx)$, where $\mu_t^{[n]} := \mathcal{L}(X_t^{i,[n]}|\mathcal{W}) = \mathcal{L}(X_t^{i,[n]}|\mathcal{W}_t)$.

In what follows, we shall consider $u_t^{[n]} := \mathbb{E}\left[\left(X_t^{i,[n+1]} - X_t^{i,[n]}\right)^2\right]$. Let us introduce

$$h(x, y, m, v) := \sqrt{f(x, m)}\tilde{\Psi}(x, y, m, v) \text{ and } g(x, y, m) := \sqrt{f(x, m)}\kappa(x, y, m).$$

Note that the assumptions of the theorem guarantee that h and g are Lipschitz continuous. Indeed, using Assumption 5.1 (ii) and (iii), for all $x, y, x', y', v \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|h(x, y, m, v) - h(x', y', m', v)| \leq C(v)(|x - x'| + |y - y'| + W_1(m, m')),$$

where

$$C(v) := \int_E M(u, v)\nu_1(du) + \int_E |\Psi(x, y, m, v, u)|\nu_1(du).$$

Using Jensen's inequality together with Assumptions 5.1 (iii) and 5.2 (ii) we see that C satisfies $\int_{\mathbb{R}} C(v)^2 dv < \infty$. Moreover, using Assumption 5.1 (ii) together with Remark 5.1.1, for all $x, x', y, y' \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|g(x, y, m) - g(x', y', m')| \leq K(|x - x'| + |y - y'| + W_1(m, m')),$$

where $K \leq C(\|\kappa\|_\infty + \sqrt{\|f\|_\infty})$.

Step 2. We now prove that our Picard scheme converges. Classical arguments imply the existence of a constant $C > 0$ such that

$$\begin{aligned}
&\frac{1}{C}\left(X_t^{i,[n+1]} - X_t^{i,[n]}\right)^2 \leq \\
&\left(\int_0^t (b(X_s^{i,[n]}, \mu_s^{[n]}) - b(X_s^{i,[n-1]}, \mu_s^{[n-1]}))ds\right)^2 + \left(\int_0^t (\sigma(X_s^{i,[n]}, \mu_s^{[n]}) - \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]}))d\beta_s^i\right)^2 \\
&+ \left(\int_0^t \int_{\mathbb{R}}\int_{\mathbb{R}} h(x, X_s^{i,[n]}, \mu_s^{[n]}, v)dM^{[n]}(s, x, v) - \int_0^t \int_{\mathbb{R}}\int_{\mathbb{R}} h(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}, v)dM^{[n-1]}(s, x, v)\right)^2 \\
&\quad + \left(\int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n]}, \mu_s^{[n]})dM^{i,[n]}(s, x) - \int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n-1]}, \mu_s^{[n-1]})dM^{i,[n-1]}(s, x)\right)^2 \\
&\leq t\int_0^t \left(b(X_s^{i,[n]}, \mu_s^{[n]}) - b(X_s^{i,[n-1]}, \mu_s^{[n-1]})\right)^2 ds + \left(\int_0^t (\sigma(X_s^{i,[n]}, \mu_s^{[n]}) - \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]}))d\beta_s^i\right)^2 \\
&+ \left(\int_0^t \int_0^1 \int_{\mathbb{R}} \left[h((F_s^{[n]})^{-1}(p), \bar{X}_s^{i,[n]}, \mu_s^{[n]}, v) - h((F_s^{[n-1]})^{-1}(p), \bar{X}_s^{i,[n-1]}, \mu_s^{[n-1]}, v)\right] dW(s, p, v)\right)^2 \\
&\quad + \left(\int_0^t \int_0^1 \left[g((F_s^{[n]})^{-1}(p), \bar{X}_s^{i,[n]}, \mu_s^{[n]}) - g((F_s^{[n-1]})^{-1}(p), \bar{X}_s^{i,[n-1]}, \mu_s^{[n-1]})\right] dW^i(s, p)\right)^2. \quad (5.9)
\end{aligned}$$

Using Burkholder-Davis-Gundy's inequality to control the expectation of the stochastic integrals above, and using the fact that for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$,

$$W_1(\mu,\nu) \leq W_2(\mu,\nu),$$

we have that

$$u_t^{[n]} \leq C(1+t)\int_0^t \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2\right] ds + C(1+t)\int_0^t \mathbb{E}\left[W_2(\mu_s^{[n]},\mu_s^{[n-1]})^2\right] ds$$

$$+ C\int_0^t \mathbb{E}\left[\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp\right] ds. \quad (5.10)$$

A classical result (see e.g. Theorem 8.1 of Major (1978)) states that, if F, G are two distribution functions with associated probability measure μ and ν, respectively, then

$$\int_0^1 (F^{-1}(p) - G^{-1}(p))^2 dp = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left[(X-Y)^2\right] = W_2(\mu,\nu),$$

where the infimum is taken over all possible couplings (X,Y) of μ and ν.

This implies that

$$\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp = W_2(\mu_s^{[n]},\mu_s^{[n-1]})^2.$$

Since $X^{i,[n]}$ and $X^{i,[n-1]}$, conditionally on $\mathcal{W}$, are respectively realizations of $F^{i,[n]}$ and $F^{i,[n-1]}$, we have that, for every $s \geq 0$, almost surely,

$$W_2(\mu_s^{[n]},\mu_s^{[n-1]})^2 \leq \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2|\mathcal{W}\right].$$

Integrating with respect to W implies that

$$\mathbb{E}\left[\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp\right] \leq \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2\right].$$

Consequently, we have shown that there exists some constant $C > 0$ such that, for all $t \geq 0$,

$$u_t^{[n]} \leq C(1+t)\int_0^t u_s^{[n-1]} ds. \quad (5.11)$$

Classical computations then give

$$u_t^{[n]} \leq C^n (1+t)^n \frac{t^n}{n!}.$$

Now, introducing $v_t^{[n]} := 2^n u_t^{[n]}$, we have that

$$\sum_{n\geq 0} v_t^{[n]} < \infty.$$

Hence, using that for all $x \in \mathbb{R}, \varepsilon > 0$, $|x| \leq \max(\varepsilon, x^2/\varepsilon) \leq \varepsilon + x^2/\varepsilon$, and applying this with $\varepsilon = 1/2^n$ and $x = X_t^{i,[n+1]} - X_t^{i,[n]}$, we have

$$\sum_{n\geq 0} \mathbb{E}\left[|X_t^{i,[n+1]} - X_t^{i,[n]}|\right] \leq \sum_{n\geq 0} \frac{1}{2^n} + \sum_{n\geq 0} v_t^{[n]} < \infty.$$

As a consequence, we can define, almost surely,

$$\bar{X}_t^i := X_0^i + \sum_{n\geq 0} (X_t^{i,[n+1]} - X_t^{i,[n]}) < +\infty,$$

and we know that $\mathbb{E}\left[|\bar{X}_t^i - X_t^{i,[n]}|\right]$ vanishes as n goes to infinity, and that $X_t^{i,[n]}$ converges almost surely to $\bar{X}_t^i$.

Step 3. Let us prove that $\bar{X}^i$ has finite fourth moments. Let $w_t^{[n]} := \mathbb{E}\left[(X_t^{i,[n]})^4\right]$.

By equation (5.8), we have

$$\begin{aligned}
\frac{1}{C}\left(X_t^{i,[n]}\right)^4 \leq & \left(\int_0^t b(X_s^{i,[n-1]}, \mu_s^{[n-1]}) ds\right)^4 + \left(\int_0^t \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]}) d\beta_s^i\right)^4 \\
& + \left(\int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}, v) dM^{[n-1]}(s,x,v)\right)^4 \\
& + \left(\int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}) dM^{i,[n-1]}(s,x)\right)^4 \\
& =: A_1 + A_2 + A_3 + A_4.
\end{aligned} \tag{5.12}$$

First of all, let us note that our Lipschitz assumptions allow to consider the following control: for any $x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})$,

$$|b(x,m)| \leq |b(x,m) - b(0,\delta_0)| + |b(0,\delta_0)| \leq C(1 + |x| + W_1(\mu,\delta_0)) = C\left(1 + |x| + \int_{\mathbb{R}} |y| dm(y)\right),$$

and similar controls for the functions σ, h, g. Using this control and Jensen's inequality, we have

$$\begin{aligned}
\mathbb{E}[A_1] \leq & t^3 \int_0^t \mathbb{E}\left[b(X_s^{i,[n-1]}, \mu_s^{[n-1]})^4\right] ds \\
\leq & Ct^3 \int_0^t \left(1 + w_s^{[n-1]} + \mathbb{E}\left[\left(\int_{\mathbb{R}} |y| \mu_s^{[n-1]}(dy)\right)^4\right]\right) ds \\
\leq & Ct^3 \int_0^t \left(1 + w_s^{[n-1]}\right) ds.
\end{aligned}$$

We can obtain a similar control for the expressions A_2, A_3 and A_4 using Burkholder-Davis-Gundy's inequality noticing that the stochastic integrals involved are local martingales. We just give the details for A_2.

$$\mathbb{E}[A_2] \leq \mathbb{E}\left[\left(\int_0^t \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]})^2 ds\right)^2\right]$$

$$\leq \mathbb{E}\left[t\int_0^t \sigma(X_s^{i,[n-1]},\mu_s^{[n-1]})^4 ds\right] \leq Ct\int_0^t \left(1+w_s^{[n-1]}\right)ds,$$

where the last inequality can be obtained with the same reasoning as the one used to control $\mathbb{E}[A_1]$. With the same reasoning, we have the following controls for A_3 and A_4 :

$$\mathbb{E}[A_3]+\mathbb{E}[A_4] \leq Ct\int_0^t \left(1+w_s^{[n-1]}\right)ds.$$

Using the previous control in the inequality (5.12), we have that for all $t \geq 0$,

$$w_t^{[n+1]} \leq C(1+t^3)+C(1+t^3)\int_0^t w_s^{[n]}ds,$$

whence

$$w_t^{[n]} \leq \sum_{k=1}^n \frac{t^{k-1}}{(k-1)!}C^k(1+t^3)^k \leq C(1+t^3)e^{Ct(1+t^3)}.$$

Consequently

$$\sup_{n\in\mathbb{N}}\sup_{0\leq s\leq t}\mathbb{E}\left[(X_s^{i,[n]})^4\right] < \infty, \tag{5.13}$$

for some constant $C>0$. Then Fatou's lemma implies the result: for all $t \geq 0$,

$$\sup_{0\leq s\leq t}\mathbb{E}\left[(\bar{X}_s^i)^4\right] < \infty. \tag{5.14}$$

Step 4. Finally, we conclude the proof showing that $\bar{X}^i$ is solution to the limit equation. Roughly speaking, this step consists in letting n tend to infinity in (5.8).

We want to prove that, for all $t \geq 0$,

$$\bar{X}_t^i = G_t^i(\bar{X}^i,\mu), \tag{5.15}$$

where

$$\begin{aligned} G_t^i(\bar{X}^i,\mu) :=& \int_0^t b(\bar{X}_s^i,\mu_s)ds + \int_0^t \sigma(\bar{X}_s^i,\mu_s)d\beta_s^i \\ &+\int_0^t\int_0^1\int_{\mathbb{R}} h(F_s^{-1}(p),\bar{X}_s^i,\mu_s,v)dW(s,p,v) \\ &+\int_0^t\int_0^1 g(F_s^{-1}(p),\bar{X}_s^i,\mu_s)dW^i(s,p), \end{aligned}$$

where the functions h and g have been introduced in *Step 1*, $\mu_t := \mathcal{L}(\bar{X}_t^i|\mathcal{W})$, and F_t^{-1} is the generalized inverse of

$$F_t(x) := \mathbb{P}\left(\bar{X}_t^i \leq x \middle| \mathcal{W}\right).$$

Let us note that G_t^i has to be understood as a notation, we do not use its functional properties.

By construction, we have

$$X_t^{i,[n+1]} = G_t^i(X^{i,[n]},\mu^{[n]}). \tag{5.16}$$

We have proved in *Step 2* that $X_t^{[n+1]}$ converges to $\bar{X}_t^i$ in L^1. In other words, the LHS of (5.16) converges to the LHS of (5.15) in L^1. Now, it is sufficient to prove that the RHS converges in L^2. This will prove that the equation (5.15) holds true.

With the same computations as the ones used to obtain (5.11) (recalling that this inequality relies on (5.9) and (5.10)), we have

$$\mathbb{E}\left[\left(G_t^i(\bar{X}^i,\mu) - G_t^i(X^{i,[n]},\mu^{[n]})\right)^2\right] \le C(1+t)\int_0^t \mathbb{E}\left[\left(\bar{X}_s^i - X_s^{i,[n]}\right)^2\right] ds.$$

This proves that $G_t^i(X^{[n],i},\mu^{[n]})$ converges to $G_t^i(\bar{X}^i,\mu)$ in L^2 by dominated convergence: indeed, we know that for all $s\le t$, $X_s^{[n],i}$ converges to $\bar{X}_s^i$ almost surely thanks to *Step 2.*, and (5.13) and (5.14) give the uniform integrability.

5.3.2 Trajectorial uniqueness

We continue the proof of Theorem 5.3.1 by proving the uniqueness of the solution. For that sake, let $\hat{X}^i$ and $\check{X}^i$ be two strong solutions defined with respect to the same initial condition X_0^i and the same white noises W and W^i. Let

$$u_t := \mathbb{E}\left[(\hat{X}_t^i - \check{X}_t^i)^2\right].$$

According to the computation of the previous subsection, there exists a constant $C>0$ such that for all $t\ge 0$,

$$u_t \le C(1+t)\int_0^t u_s ds.$$

Then Grönwall's lemma implies that $u_t = 0$ for all $t\ge 0$, implying the uniqueness.

5.4 Conditional propagation of chaos

In this section, we prove the conditional propagation of chaos property of the particle systems (5.1):

Theorem 5.4.1. *Under Assumptions 5.1, 5.2 and 5.3, $(X^{N,i})_{1\le i\le N}$ converges to $(\bar{X}^i)_{i\ge1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$.*

In the above statement, we implicitly define $X^{N,i} := 0$ if $i>N$.

As the systems $(X^{N,i})_{1\le i\le N}$ $(N\in\mathbb{N}^*)$ and $(\bar{X}^i)_{i\ge1}$ are exchangeable, Theorem 5.4.1 is equivalent to

Theorem 5.4.2. *Under Assumptions 5.1, 5.2 and 5.3, the system $(\bar{X}^i)_{i\ge1}$ is exchangeable with directing measure $\mu = \mathcal{L}(\bar{X}^1|W)$, where W is as in (5.7). Moreover, the sequence of empirical measures*

$$\mu^N := N^{-1}\sum_{i=1}^N \delta_{X^{N,i}}$$

converges in law, as $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$- valued random variables, to μ.

Remark 5.4.3. *In Theorem 5.4.2, it is easy to prove that $\mathcal{L}(\bar{X}^1|W)$ is the directing measure of the system $(\bar{X}^i)_{i\geq 1}$. Indeed, it is sufficient to notice that conditionally on W, the variables $\bar{X}^i$ $(i \geq 1)$ are i.i.d. and to apply Lemma (2.12) of Aldous (1983).*

The proof of Theorem 5.4.2 is similar to the proof of Theorem 4.4.2. It consists in showing that $(\mu^N)_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))$, and that each converging subsequence converges to the same limit, using a convenient martingale problem. For this reason, in what follows we just give the proofs that substantially change compared to our previous chapter. For the other proofs, we just give the main ideas and cite precisely the corresponding statement of Chapter 4 that allows to conclude.

5.4.1 Tightness of $(\mu^N)_N$

The proof that the sequence $(\mu^N)_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))$, is almost the same as that of Proposition 4.4.3. As the systems $(X^{N,i})_{1\leq i\leq N}$ $(N \in \mathbb{N}^*)$ are exchangeable, it is equivalent to the tightness of the sequence $(X^{N,1})_N$ on $D(\mathbb{R}_+, \mathbb{R})$ (see Proposition 2.2-(ii) of Sznitman (1989)).

The tightness of $(X^{N,1})_N$ is straightforward using Aldous' criterion (see Theorem 4.5 of Jacod and Shiryaev (2003)), observing that, under our conditions, $\sup_N \mathbb{E}\left[\sup_{s\leq t} |X^{N,1}_s|\right] < \infty$ (see Lemma 5.2.2).

5.4.2 Martingale problem

To identify the structure of any possible limit of the sequence $(\mu^N)_N$, we introduce a convenient martingale problem. We have already used such a kind of martingale problem in a similar context in Section 4.4.2. Since our limit system is necessarily an infinite exchangeable system, the martingale problem is constructed such that it reflects the correlations between the particles. It is therefore stated in terms of couples of particles.

Consider a probability measure $Q \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})))$. In what follows the role of Q will be to be the law of any possible limit μ of μ^N. Our martingale problem is stated on the canonical space

$$\Omega = \mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \times D(\mathbb{R}_+, \mathbb{R})^2.$$

We endow Ω with the product of the associated Borel sigma-fields and with the probability measure defined for all $A \in \mathcal{B}(\mathcal{P}(D(\mathbb{R}_+, \mathbb{R})))$, $B \in \mathcal{B}(D(\mathbb{R}_+, \mathbb{R})^2)$ by

$$P_Q(A \times B) := \int_{\mathcal{P}_1(D(\mathbb{R}_+,\mathbb{R}))} \mathbb{1}_A(m) m \otimes m(B) Q(dm). \tag{5.17}$$

We write an atomic event $\omega \in \Omega$ as $\omega = (m, y)$ with $y = (y_t)_{t\geq 0} = (y^1_t, y^2_t)_{t\geq 0}$. We write μ and $Y = (Y^1, Y^2)$ for the random variables $\mu(\omega) = m$ and $Y(\omega) = y$. The definition (5.17) implies that Q is the distribution of μ, and that conditionally on μ, Y^1 and Y^2 are i.i.d. with distribution μ. More precisely, $\mu \otimes \mu$ is a regular conditional distribution of (Y^1, Y^2) given μ.

For $t \geq 0$ we write m_t for the t-th marginal of $m : m_t(C) = m(Y \in D^2(\mathbb{R}_+, \mathbb{R});\ y_t \in C)$; $C \in \mathcal{B}(\mathbb{R})$. We denote μ_t the r.v. $\mu_t(\omega) = m_t$ and consider the filtration $(\mathcal{G}_t)_{t\geq 0}$ given by

$$\mathcal{G}_t = \sigma(Y_s, s \leq t) \vee \sigma(\mu_s, s \leq t).$$

For all $g \in C^2_b(\mathbb{R}^2)$, define

$$Lg(y, m, x, v) := b(y^1, m)\partial_{y^1} g(y) + b(y^2, m)\partial_{y^2} g(y) \tag{5.18}$$

$$+\frac{1}{2}\sigma(y^1,m)^2\partial^2_{y^1}g(y)+\frac{1}{2}\sigma(y^2,m)^2\partial^2_{y^2}g(y)$$
$$+\frac{1}{2}f(x,m)\kappa(x,y^1,m)^2\partial^2_{y^1}g(y)+\frac{1}{2}f(x,m)\kappa(x,y^2,m)^2\partial^2_{y^2}g(y)$$
$$+\frac{1}{2}f(x,m)\sum_{i,j=1}^{2}\tilde{\Psi}(x,y^i,m,v)\tilde{\Psi}(x,y^j,m,v)\partial^2_{y^iy^j}g(y),$$

and put for all $t\geq 0$

$$M_t^g=M_t^g(\mu,Y):=g(Y_t)-g(Y_0)-\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}}Lg(Y_s,\mu_s,x,v)\nu_1(dv)\mu_s(dx)ds. \tag{5.19}$$

Definition 5.4.4. *Q is solution to the martingale problem $(\mathcal{M})$ if*

- *$Q-$almost surely, $\mu_0=\nu_0$,*
- *for all $g\in C_b^2(\mathbb{R}^2)$, $(M_t^g)_{t\geq 0}$ is a $(P_Q,(\mathcal{G}_t)_t)-$martingale.*

Let us state a Lemma that allows us to partially recover our limit equation (5.3) from the martingale problem $(\mathcal{M})$. It is the equivalent of Lemma 4.4.5 in the framework of white noises.

Lemma 5.4.5. *Grant Assumptions 5.1, 5.2 and 5.3. Let Q be a solution of $(\mathcal{M})$. Using the notation above, there exist on an extension of $(\Omega,\mathcal{F},(\mathcal{G}_t)_{t\geq 0},P_Q)$ two Brownian motions β^1,β^2 and three white noises W^1,W^2,W; defined on $(\mathbb{R}_+^2,\ \mathcal{B}(\mathbb{R}_+^2),\ dt\otimes dp)$ for W^i, $i=1,2$; and on $(\mathbb{R}_+^2\times\mathbb{R},\ \mathcal{B}(\mathbb{R}_+^2)\otimes\mathcal{B}(\mathbb{R}),\ dt\otimes dp\otimes\nu_1(dv))$ for W, such that $\beta^1,\beta^2,W^1,W^2,W$ are all independent and such that (Y_t) admits the representation*

$$\begin{aligned}
dY_t^1=&b(Y_t^1,\mu_t)dt+\sigma(Y_t^1,\mu_t)d\beta_t^1\\
&+\int_0^1\int_{\mathbb{R}}\sqrt{f(F_s^{-1}(p),\mu_t)}\tilde{\Psi}(F_s^{-1}(p),Y_t^1,\mu_t,v)dW(t,p,v)\\
&+\int_0^1\sqrt{f(F_s^{-1}(p),\mu_t)}\kappa(F_s^{-1}(p),Y_t^1,\mu_t)dW^1(t,p),\\
dY_t^2=&b(Y_t^2,\mu_t)dt+\sigma(Y_t^2,\mu_t)d\beta_t^2\\
&+\int_0^1\int_{\mathbb{R}}\sqrt{f(F_s^{-1}(p),\mu_t)}\tilde{\Psi}(F_s^{-1}(p),Y_t^2,\mu_t,v)dW(t,p,v)\\
&+\int_{\mathbb{R}}\sqrt{f(F_s^{-1}(p),\mu_t)}\kappa(F_s^{-1}(p),Y_t^2,\mu_t)dW^2(t,p),
\end{aligned}$$

where F_s is the distribution function related to μ_s, and F_s^{-1} its generalized inverse.

Proof. Theorem II.2.42 of Jacod and Shiryaev (2003) implies that Y is a continuous semimartingale with characteristics (B,C) given by

$$\begin{aligned}
B_t^i&=\int_0^t b(Y_s^i,\mu_s)ds,\quad 1\leq i\leq 2,\\
C_t^{i,i}&=\int_0^t\sigma(Y_s^i,\mu_s)^2ds+\int_0^t\int_{\mathbb{R}}f(x,\mu_s)\kappa(x,Y_s^i,\mu_s)^2\mu_s(dx)ds
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,\mu_s)\tilde{\Psi}(x,Y_s^i,\mu_s,v)^2\nu_1(dv)\mu_s(dx)ds, \quad 1\le i\le 2,\\
C_t^{1,2} &= \int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,\mu_s)\tilde{\Psi}(x,Y_s^1,\mu_s,v)\tilde{\Psi}(x,Y_s^2,\mu_s,v)\nu_1(dv)\mu_s(dx)ds.
\end{aligned}
$$

As we are interested in finding five white noises, we need to have five local martingales. This is why we introduce artificially $Y_t^i := 0$ for $3 \le i \le 5$. The rest of the proof is then an immediate consequence of Theorem III-10 and Theorem III-6 of El Karoui and Méléard (1990). More precisely, Theorem III-10 implies the existence of five orthogonal martingale measures M^i $(1 \le i \le 5)$ on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$ with intensity $dt \cdot \mu_t(dx) \cdot \nu_1(dv)$ such that

$$
\begin{aligned}
dY_t^1 =& b(Y_t^1,\mu_t)dt + \int_{\mathbb{R}}\int_{\mathbb{R}} \sigma(Y_t^1,\mu_t)dM^1(t,x,v)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\tilde{\Psi}(x,Y_t^1,\mu_t,v)dM^5(t,x,v)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\kappa(x,Y_t^1,\mu_t)dM^3(t,p,v),\\
dY_t^2 =& b(Y_t^2,\mu_t)dt + \int_{\mathbb{R}}\int_{\mathbb{R}} \sigma(Y_t^2,\mu_t)dM^2(t,x,v)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\tilde{\Psi}(x,Y_t^2,\mu_t,v)dM^5(t,x,v)\\
&+\int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\kappa(x,Y_t^2,\mu_t)dM^4(t,x,v).
\end{aligned}
$$

Then, $\beta_t^i := M^i([0,t]\times\mathbb{R}\times\mathbb{R})$ $(i = 1,2)$ are standard one-dimensional Brownian motions, and Theorem III-6 of El Karoui and Méléard (1990) allows us to write M^i $(3 \le i \le 5)$ as

$$
\begin{aligned}
M^5([0,t]\times A\times B) &= \int_0^t \mathbb{1}_A(F_s^{-1}(p))\mathbb{1}_B(v)dW(s,p,v)\\
M^3([0,t]\times A\times B) &= \int_0^t \mathbb{1}_A(F_s^{-1}(p))dW^1(s,p)\\
M^4([0,t]\times A\times B) &= \int_0^t \mathbb{1}_A(F_s^{-1}(p))dW^2(s,p),
\end{aligned}
$$

where W^1, W^2, W are white noises with respective intensities $dt\cdot dp$, $dt \cdot dp$ and $dt\cdot dp\cdot \nu_1(dv)$. □

We now prove a key result for the proof of our main results. Recall that the martingale problem $(\mathcal{M})$ is given by definition (5.4.4).

Theorem 5.4.6. *Grant Assumptions 5.1, 5.2 and 5.3. Then the law of every limit in distribution of the sequence (μ^N) is solution of the martingale problem $(\mathcal{M})$.*

Proof. Let μ be the limit in distribution of some subsequence of (μ^N) and let $Q = Q_\mu$ be its law. In the following, we still note this subsequence (μ^N). Firstly, we clearly have that $\mu^0 = \nu^0$. For $0 \le s_1 \le ... \le s_k \le s$; $\psi_1, ..., \psi_k \in C_b(\mathcal{P}(\mathbb{R}))$; $\varphi_1, ..., \varphi_k \in C_b(\mathbb{R}^2)$; $\varphi \in C_b^3(\mathbb{R}^2)$ define the following functional on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$:

$$
F(\mu) := \psi_1(\mu_{s_1})...\psi_k(\mu_{s_k})\int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu\otimes\mu(d\gamma)\varphi_1(\gamma_{s_1})\dots\varphi_k(\gamma_{s_k})\left[M_t^\varphi(\mu,\gamma) - M_s^\varphi(\mu,\gamma)\right],
$$

where $(M_t^\varphi(\mu,\gamma))_t$ is given by (5.19). To show that $(M_t^\varphi(\mu,\gamma))_{t\geq 0}$ is a $(P_{Q_\mu},(\mathcal{G}_t)_t)$ martingale, we have to show that

$$\int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))} F(m)Q_\mu(dm) = \mathbb{E}\left[F(\mu)\right] = 0,$$

where the expectation $\mathbb{E}\left[\cdot\right]$ is taken with respect to $P_{Q_\mu} = P_Q$. Note that the first equality is just the transfert formula, and hence the expectation is taken on the probability space where μ is defined.

Step 1. We show that

$$\mathbb{E}\left[F(\mu^N)\right] \to \mathbb{E}\left[F(\mu)\right] \tag{5.20}$$

as $N\to\infty$. This statement is not immediately clear since the functional F is not continuous. The main difficulty comes from the fact that μ^N converges to μ in distribution for the topology of the weak convergence, but the terms appearing in the function F require the convergence of μ_t^N to μ_t for the topology of the metric W_1.

The proof (5.20) is actually quite technical, therefore we prove it in Lemma 5.4.7 below.

Step 2. In this step we show that $\mathbb{E}\left[F(\mu)\right]$ is equal to 0. Applying F to μ^N gives

$$\begin{aligned}
F(\mu^N) = \psi_1(\mu^N_{s_1})...\psi_k(\mu^N_{s_k})\frac{1}{N^2}\sum_{i,j=1}^N \varphi_1(X^{N,i}_{s_1},X^{N,j}_{s_1})...\varphi_k(X^{N,i}_{s_k},X^{N,j}_{s_k})&\\
\Big[\varphi(X_t^{N,i},X_t^{N,j}) - \varphi(X_s^{N,i},X_s^{N,j})&\\
-\int_s^t b(X_r^{N,i},\mu_r^N)\partial_{x^1}\varphi(X_r^{N,i},X_r^{N,j})dr - \int_s^t b(X_r^{N,j},\mu_r^N)\partial_{x^2}\varphi(X_r^{N,i},X_r^{N,j})dr&\\
-\frac{1}{2}\int_s^t \sigma(X_r^{N,i},\mu_r^N)^2\partial^2_{x^1}\varphi(X_r^{N,i},X_r^{N,j})dr - \frac{1}{2}\int_s^t \sigma(X_r^{N,j},\mu_r^N)^2\partial^2_{x^2}\varphi(X_r^{N,i},X_r^{N,j})dr&\\
-\frac{1}{2}\int_s^t \frac{1}{N}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\kappa(X_r^{N,k},X_r^{N,i},\mu_r^N)^2\partial^2_{x^1}\varphi(X_r^{N,i},X_r^{N,j})dr&\\
-\frac{1}{2}\int_s^t \frac{1}{N}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\kappa(X_r^{N,k},X_r^{N,j},\mu_r^N)^2\partial^2_{x^2}\varphi(X_r^{N,i},X_r^{N,j})dr&\\
-\frac{1}{2}\int_s^t\int_{\mathbb{R}}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\frac{1}{N}\sum_{h,l=1}^2 \tilde{\Psi}(X_r^{N,k},X_r^{N,i_h},\mu_r^N,v)\tilde{\Psi}(X_r^{N,k},X_r^{N,i_l},\mu_r^N,v)\cdot&\\
\partial^2_{x^hx^l}\varphi(X_r^{N,i},X_r^{N,j})\nu_1(dv)dr\Big],&
\end{aligned} \tag{5.21}$$

with $i_1 = i$ and $i_2 = j$. Denote $\tilde{\pi}^k(dr,dz,du) := \pi(dr,dz,du) - dr\cdot dz\cdot\nu(du)$ the compensated version of π^k. For $(i,j)\in[\![1,\ldots,N]\!]^2,\ s<t$ let us define

$$M_{s,t}^{N,i,j} := \int_s^t \sigma(X_r^{N,i},\mu_r^N)\partial_{x^1}\varphi(X_r^{N,i},X_r^{N,j})d\beta_r^i + \int_s^t \sigma(X_r^{N,j},\mu_r^N)\partial_{x^2}\varphi(X_r^{N,i},X_r^{N,j})d\beta_r^j; \tag{5.22}$$

$$W_{s,t}^{N,i,j} := \sum_{k=1,k\neq i,j}^N \int_{]s,t]\times\mathbb{R}_+\times E} \mathbb{1}_{\left\{z\leq f(X^{N,k}_{r-},\mu^N_{r-})\right\}}$$

$$\Big[\varphi(X^{N,i}_{r-}+\frac{1}{\sqrt{N}}\Psi(X^{N,k}_{r-},X^{N,i}_{r-},\mu^N_{r-},u^k,u^i),\ X^{N,j}_{r-}+\frac{1}{\sqrt{N}}\Psi(X^{N,k}_{r-},X^{N,j}_{r-},\mu^N_{r-},u^k,u^j)) \\ -\varphi(X^{N,i}_{r-},X^{N,j}_{r-})\Big]\,\tilde{\pi}^k(dr,dz,du); \quad (5.23)$$

$$\Delta^{N,i,j}_{s,t} := \sum_{k=1,k\neq i,j}^{N}\int_s^t\int_E f(X^{N,k}_r,\mu^N_{r-}) \\ \Big[\varphi(X^{N,i}_{r-}+\frac{1}{\sqrt{N}}\Psi(X^{N,k}_{r-},X^{N,i}_{r-},\mu^N_{r-},u^k,u^i),\ X^{N,j}_{r-}+\frac{1}{\sqrt{N}}\Psi(X^{N,k}_{r-},X^{N,j}_{r-},\mu^N_{r-},u^k,u^j)) \\ -\varphi(X^{N,i}_{r-},X^{N,j}_{r-})\Big]\,\nu(du)dr \quad (5.24)$$

and

$$\Gamma^{N,i,j}_{s,t} = \Delta^{N,i,j}_{s,t} \\ -\sum_{l=1}^{2}\sum_{k=1,k\neq i,j}^{N}\int_s^t\int_E f(X^{N,k}_r,\mu^N_r)\frac{1}{\sqrt{N}}\Psi(X^{N,k}_{r-},X^{N,i_l}_{r-},\mu^N_{r-},u^k,u^{i_l})\partial_{x^l}\varphi(X^{N,i}_r,X^{N,j}_r)\nu(du)dr \\ -\sum_{h,l=1}^{2}\int_s^t\int_E\frac{1}{N}\sum_{k=1,k\neq i,j}^{N} f(X^{N,k}_r,\mu^N_r)\Psi(X^{N,k}_r,X^{N,i_h}_r,\mu^N_r,u^k,u^{i_h})\Psi(X^{N,k}_r,X^{N,i_l}_r,\mu^N_r,u^k,u^{i_l}) \\ \partial^2_{x^hx^l}\varphi(X^{N,i}_r,X^{N,j}_r)\nu(du)dr, \quad (5.25)$$

with again $i_1=i$ and $i_2=j$.

Applying Ito's formula, we have

$$\varphi(X^{N,i}_t,X^{N,j}_t)=\varphi(X^{N,i}_s,X^{N,j}_s)+M^{N,i,j}_{s,t}+W^{N,i,j}_{s,t}+\Delta^{N,i,j}_{s,t} \\ +\int_s^t b(X^{N,i}_r,\mu^N_r)\partial_{x^1}\varphi(X^{N,i}_r,X^{N,j}_r)dr+\int_s^t b(X^{N,j}_r,\mu^N_r)\partial_{x^2}\varphi(X^{N,i}_r,X^{N,j}_r)dr \\ +\frac{1}{2}\int_s^t\sigma(X^{N,i}_r,\mu^N_r)^2\partial^2_{x^1}\varphi(X^{N,i}_r,X^{N,j}_r)dr+\frac{1}{2}\int_s^t\sigma(X^{N,j}_r,\mu^N_r)^2\partial^2_{x^2}\varphi(X^{N,i}_r,X^{N,j}_r)dr. \quad (5.26)$$

Note that thanks to Assumption (5.3) i), the term in the second line of (5.25) is zero. Again, if $i\neq j$, using the definitions (5.5) and (5.6)

$$\int_E\Psi(X^{N,k}_r,X^{N,i}_r,\mu^N_r,u^k,u^i)\Psi(X^{N,k}_r,X^{N,j}_r,\mu^N_r,u^k,u^j)\nu(du) \\ =\int_{\mathbb{R}}\tilde{\Psi}(X^{N,k}_r,X^{N,i}_r,\mu^N_r,v)\tilde{\Psi}(X^{N,r},X^{N,j},\mu^N_r,v)\nu_1(dv),$$

and

$$\int_E\Psi(X^{N,k}_r,X^{N,i}_r,\mu^N_r,u^k,u^i)^2\nu(du)\ =\ \kappa(X^{N,k}_r,X^{N,i}_r,\mu^N_r)^2+\int_{\mathbb{R}}\tilde{\Psi}(X^{N,k}_r,X^{N,i}_r,\mu^N_r,v)^2\nu_1(dv).$$

This implies that the last line of (5.25) is equal to the sum of three last lines of (5.21), up to an error term which is of order $O(t/N)$ if we include the terms $k = i, j$ in (5.25).

As a consequence, plugging (5.26) in (5.21), using the definitions (5.23), (5.22), (5.25) and the previous remark we obtain

$$F(\mu^N) = \psi_1(\mu^N_{s_1})...\psi_k(\mu^N_{s_k})\frac{1}{N^2}\sum_{i,j=1}^{N}\varphi_1(X^{N,i}_{s_1}, X^{N,j}_{s_1})...\varphi_k(X^{N,i}_{s_k}, X^{N,j}_{s_k})\left[M^{N,i,j}_{s,t} + W^{N,i,j}_{s,t} + \Gamma^{N,i,j}_{s,t}\right].$$

Using (5.26) we see that $(M^{N,i,j}_{s,t} + W^{N,i,j}_{s,t}),\quad t \geq s$, is a martingale with respect to the filtration $(\mathcal{F}^{X^N}_t)_{t\geq 0}$ with $\mathcal{F}^{X^N}_t := \sigma(X^{N,i}_u; X^{N,j}_u;\ s \leq u \leq t)$ on the space where X^N is defined. Hence, using that φ_{s_k} and ψ_{s_k} are bounded, and that μ^N is $(\mathcal{F}^{X^N}_t)_{t\geq 0}$ adapted,

$$\begin{aligned}
&\mathbb{E}\left[F(\mu^N)\right] = \mathbb{E}\left[\mathbb{E}\left[F(\mu^N)|\mathcal{F}^{X^N}_s\right]\right] = \\
&\qquad \mathbb{E}\left[\psi_1(\mu^N_{s_1})...\psi_k(\mu^N_{s_k})\frac{1}{N^2}\sum_{i,j=1}^{N}\varphi_1(X^{N,i}_{s_1}, X^{N,j}_{s_1})...\varphi_k(X^{N,i}_{s_k}, X^{N,j}_{s_k})\mathbb{E}\left[\Gamma^{N,i,j}_{s,t}|\mathcal{F}^{X,\mu}_s\right]\right] \\
&\qquad\qquad \leq \frac{C}{N^2}\sum_{i,j=1}^{N}\mathbb{E}\left[|\Gamma^{N,i,j}_{s,t}|\right], \qquad (5.27)
\end{aligned}$$

implying that

$$|\mathbb{E}\left[F(\mu^N)\right]| \leq C\mathbb{E}\left[|\Gamma^{N,1,2}_{s,t}| + \frac{|\Gamma^{N,1,1}_{s,t}|}{N}\right].$$

Taylor-Lagrange's inequality gives for all $i \neq j$,

$$\begin{aligned}
&\mathbb{E}\left[|\Gamma^{N,i,j}_{s,t}|\right] \leq \\
&C\frac{1}{N\sqrt{N}}\sum_{k=1,k\neq i,j}^{N}\sum_{n=0}^{3}\int_s^t\int_E \mathbb{E}\left[\Psi(X^{N,k}_r, X^{N,i}_r, \mu^N_r, u^k, u^i)^n\Psi(X^{N,k}_r, X^{N,j}_r, \mu^N_r, u^k, u^j)^{3-n}\right]\nu(du)dr \\
&\leq C\frac{1}{N\sqrt{N}}\sum_{k=1,k\neq i,j}^{N}\int_s^t\int_E \mathbb{E}\left[|\Psi(X^{N,k}_r, X^{N,i}_r, \mu^N_r, u^k, u^i)|^3 + |\Psi(X^{N,k}_r, X^{N,j}_r, \mu^N_r, u^k, u^j)|^3\right]\nu(du)dr \\
&\qquad \leq C\frac{1}{\sqrt{N}},
\end{aligned}$$

and a similar result holds for $\Gamma^{N,1,1}_{s,t}$. Consequently,

$$|\mathbb{E}\left[F(\mu^N)\right]| \leq CN^{-1/2},$$

implying together with (5.20) that

$$\mathbb{E}\left[F(\mu)\right] = \lim_N \mathbb{E}\left[F(\mu^N)\right] = 0.$$

□

Now we prove (5.20) with the following lemma.

Lemma 5.4.7. *Grant Assumptions 5.1, 5.2 and 5.3. With the notation introduced in the proof of Theorem 5.4.6, we have*

$$\mathbb{E}\left[F(\mu^N)\right] \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left[F(\mu)\right].$$

Proof. Let us recall that μ^N denotes the empirical measure of $(X^{N,i})_{1\le i\le N}$ and that μ is the limit in distribution of (a subsequence of) μ^N.

Step 1. We first show that almost surely, μ is supported by continuous trajectories. For that sake, we start showing that $P^N := \mathbb{E}\left[\mu^N\right] = \mathcal{L}(X^{N,1})$ is $C-$tight. This follows from Prop VI. 3.26 in Jacod and Shiryaev (2003), observing that

$$\lim_{N\to\infty} \mathbb{E}\left[\sup_{s\le T} |\Delta X_s^{N,1}|^3\right] = 0,$$

which follows from our conditions on ψ. Indeed, writing $\psi^*(u^1,u^2) := \sup_{x,y,m} \psi(x,y,m,u^1,u^2)$, we can stochastically upper bound

$$\sup_{s\le T} |\Delta X_s^{N,1}|^3 \le \sup_{k\le K} |\psi^*(U^{k,1},U^{k,2})|^3/N^{3/2},$$

where $K\sim Poiss(NT\|f\|_\infty)$ is Poisson distributed with parameter $NT\|f\|_\infty$, and where $(U^{k,1},U^{k,2})_k$ is an i.i.d. sequence of $\nu_1\otimes\nu_1-$distributed random variables, independent of K. The conclusion then follows from the fact that due to our Assumption 5.3, $\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right] < \infty$ such that we can upper bound

$$\begin{aligned}\mathbb{E}\left[\sup_{k\le K} |\psi^*(U^{k,1},U^{k,2})|^3/N^{3/2}\right] &\le \mathbb{E}\left[\frac{1}{N^{3/2}}\sum_{k=1}^K |\psi^*(U^{k,1},U^{k,2})|^3\right]\\ &\le \frac{\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right]}{N^{3/2}}\mathbb{E}\left[K\right] = \frac{\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right]}{N^{3/2}}NT\|f\|_\infty \to 0\end{aligned}$$

as $N\to 0$.

As a consequence of the above arguments, we know that $\mathbb{E}\left[\mu(\cdot)\right]$ is supported by continuous trajectories. In particular, almost surely, μ is also supported by continuous trajectories. Indeed, $\mu(C(\mathbb{R}_+,\mathbb{R}))$ is a r.v. taking values in $[0,1]$, and its expectation equals one. Thus $\mu(C(\mathbb{R}_+,\mathbb{R}))$ equals one a.s.

We now turn to the heart of this proof and show that $\mathbb{E}\left[F(\mu^N)\right] \to \mathbb{E}\left[F(\mu\right]$. The latter expression contains terms like

$$\int_s^t b(Y_r^1,\mu_r)\partial_{x^1}\varphi(Y_r^1,Y_r^2)dr$$

for some bounded smooth function φ. However, by our assumptions, the continuity of $m\mapsto b(x,m)$ is expressed with respect to the Wasserstein 1−distance. Yet, we only have information on the convergence of μ_r^N to μ_r for the topology of the weak convergence.

In what follows we make use of Skorohod's representation theorem and realize all random measures μ^N and μ on an appropriate probability space such that we have almost sure convergence of these realizations (we do not change notation), that is, we know that almost surely,

$$\mu^N \to \mu$$

as $N \to \infty$. (Recall that we have already chosen a subsequence in the beginning of the proof of Theorem 5.4.6). Since μ is almost surely supported by continuous trajectories, we also know that almost surely, $\mu^N_t \to \mu_t$ weakly for all t (this is a consequence of Theorem 12.5.(i) of Billingsley (1999)).

Step 2. In a first time, let us prove that, a.s., for all r, μ^N_r converges to μ_r for the metric W_1. Thus we need to show additionally that almost surely, for all $t \geq 0$, $\int |x| d\mu^N_t(x) \to \int |x| d\mu_t(x)$.

To prove this last fact, it will be helpful to consider rather the convergence of the triplets $(\mu^N, X^{N,1}, \mu^N(|x|))$. Since the sequence of laws of these triplets is tight as well (the tightness of $(\mu^N)_N$ and $(X^{N,1})_N$ have been stated in Section 5.4.1, and the tightness of $(\mu^N(|x|)_N)$ is classical from Aldous' criterion since $\mu^N_t(|x|) = N^{-1} \sum_{k=1}^N |X^{N,k}_t|$), we may assume that, after having chosen another subsequence and then a convenient realization of this subsequence, we dispose of a sequence of random triplets such that almost surely, as $N \to \infty$,

$$(\mu^N, X^{N,1}, \mu^N(|x|)) \to (\mu, Y, A),$$

where $A = (A_t)_t$ is some process having càdlàg trajectories. In addition, it can be proven that the sequence $(\mu^N(|x|))_N$ is $C-$tight (for similar reasons as $(X^{N,1})_N$), hence A has continuous trajectories.

Taking a bounded and continuous function $\Phi : D(\mathbb{R}_+, \mathbb{R}) \to \mathbb{R}$, we observe that, as $N \to \infty$,

$$\mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})} \Phi d\mu\right] \leftarrow \mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})} \Phi d\mu^N\right] = \mathbb{E}\left[\Phi(X^{N,1})\right] \to \mathbb{E}\left[\Phi(Y)\right],$$

such that $\mathbb{E}[\mu] = \mathcal{L}(Y)$.

Notice that from the above follows that Y is necessarily a continuous process, since $\mathbb{E}[\mu]$ is supported by continuous trajectories. Notice also that for the moment we do not know if $A = \mu(|x|)$.

Using that $\sup_N \mathbb{E}\left[\sup_{t\leq T} |X^{N,1}_t|^2\right] < \infty$ (see our a priori estimates Lemma 5.2.2), we deduce that the sequence $(\sup_{t\leq T} |X^{N,1}_t|^{3/2})_N$ is uniformly integrable. Therefore, $\mathbb{E}\left[\sup_{t\leq T} |X^{N,1}_t|^{3/2}\right] \to \mathbb{E}\left[\sup_{t\leq T} |Y_t|^{3/2}\right] < \infty$. In particular, we also have that

$$\mathbb{E}\left[\sup_{t\leq T} \mu_t(|x|^{3/2})\right] < \infty \quad \text{ and thus } \quad \sup_{t\leq T} \mu_t(|x|^{3/2}) < \infty \text{ almost surely,}$$

for all T, since

$$\mathbb{E}\left[\sup_{t\leq T} \mu_t(|x|^{3/2})\right] = \mathbb{E}\left[\sup_{t\leq T} \int_{D(\mathbb{R}_+,\mathbb{R})} |\gamma_t|^{3/2} \mu(d\gamma)\right] \leq \mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})} \sup_{t\leq T} |\gamma_t|^{3/2} \mu(d\gamma)\right]$$
$$= \mathbb{E}\left[\sup_{t\leq T} |Y_t|^{3/2}\right] < \infty.$$

We know that, a.s., μ^N converges weakly to μ and $\mu(C(\mathbb{R}_+, \mathbb{R})) = 1$. Let us fix some $\omega \in \Omega$ for which the two previous properties hold. In the following, we omit this ω in the notation. Let $\varepsilon > 0$, $t \leq T$ and choose M such that $\int |x| \wedge M d\mu_t \geq \int |x| d\mu_t - \varepsilon$. Then, as $N \to \infty$, almost surely,

$$\int |x| d\mu^N_t \geq \int |x| \wedge M d\mu^N_t \to \int |x| \wedge M d\mu_t.$$

Thus

$$\liminf_N \int |x| \ d\mu_t^N \geq \int |x| d\mu_t - \varepsilon,$$

such that

$$\liminf_N \int |x| \ d\mu_t^N \geq \int |x| d\mu_t. \tag{5.28}$$

Fatou's lemma implies that

$$\mathbb{E}\left[\liminf_N \int |x| \ d\mu_t^N\right] \leq \liminf_N \mathbb{E}\left[\int |x| \ d\mu_t^N\right] = \liminf_N \mathbb{E}\left[|X_t^{N,1}|\right] = \mathbb{E}\left[|Y_t|\right] = \mathbb{E}\left[\int |x| d\mu_t\right].$$

Together with (5.28) this implies that, almost surely,

$$\liminf_N \int |x| \ d\mu_t^N = \int |x| d\mu_t.$$

Finally, since $\int |x| d\mu^N \to A$ and since A is continuous, for all t,

$$\liminf_N \int |x| \ d\mu_t^N = \limsup_N \int |x| \ d\mu_t^N = \int |x| d\mu_t.$$

This implies that almost surely, for all $t \geq 0$, $\int |x| d\mu_t^N(x) \to \int |x| d\mu_t(x) = A_t < \infty$. In particular, almost surely, for all $t \geq 0$,

$$W_1(\mu_t^N, \mu_t) \to 0$$

(see e.g. Theorem 6.9 of Villani (2008)).

Step 3. Now we prove that $\mathbb{E}\left[F(\mu^N)\right]$ converges to $\mathbb{E}\left[F(\mu)\right]$, where we recall that

$$F(\mu) = \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k}) \int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu \otimes \mu(d\gamma) \varphi_1(\gamma_{s_1}) \ldots \varphi_k(\gamma_{s_k}) \left[\varphi(\gamma_t) - \varphi(\gamma_s) - \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv) \mu_r(dx) dr\right],$$

where $\psi_i \in C_b(\mathcal{P}(\mathbb{R})), \varphi_i \in C_b(\mathbb{R}^2)$ $(1 \leq i \leq k)$ and $\varphi \in C_b^3(\mathbb{R}^2)$. Let us recall some facts: by the boundedness of the functions ψ_i $(1 \leq i \leq k)$ and our boundedness Assumption 5.3, it is sufficient to prove the two following convergence:

$$\mathbb{E}\left[|\psi_1(\mu_{s_1}^N) \cdot \ldots \cdot \psi_k(\mu_{s_k}^N) - \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k})|\right] \underset{N\to\infty}{\longrightarrow} 0, \tag{5.29}$$

$$\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right] \underset{N\to\infty}{\longrightarrow} 0, \tag{5.30}$$

with

$$G(\mu) := \int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu \otimes \mu(d\gamma) \varphi_1(\gamma_{s_1}) \ldots \varphi_k(\gamma_{s_k}) \left[\varphi(\gamma_t) - \varphi(\gamma_s) - \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv) \mu_r(dx) dr\right].$$

Indeed, since the functions ψ_i ($1 \le i \le k$) and G are bounded, we have

$$\mathbb{E}\left[|F(\mu^N) - F(\mu)|\right] \le C\mathbb{E}\left[|\psi_1(\mu^N_{s_1}) \cdot \ldots \cdot \psi_k(\mu^N_{s_k}) - \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k})|\right] + C\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right].$$

The convergence (5.29) follows from dominated convergence and the fact that the function

$$m \in \mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \mapsto \psi_1(m_{s_1})...\psi_k(m_{s_k}) \in \mathbb{R}$$

is bounded and continuous at μ, since μ is supported by continuous trajectories.

To prove the convergence (5.30), let us recall that we have already shown that

1. $\sup_N \sup_{0 \le s \le t} \mathbb{E}\left[\mu^N_s(|x|^{3/2})\right] < \infty$,
2. $\sup_{0 \le s \le t} \mathbb{E}\left[\mu_t(|x|^{3/2})\right] < \infty$,
3. $\mu(C(\mathbb{R}_+, \mathbb{R})) = 1$ $a.s.$
4. a.s. $\forall r$, μ^N_r converges to μ_r for the metric W_1,
5. for all $x, x' \in \mathbb{R}, y, y' \in \mathbb{R}^2, m, m' \in \mathcal{P}_1(\mathbb{R}), v \in \mathbb{R}$,

   $$|L\varphi(y, m, x, v) - L\varphi(y', m', x', v)| \le C(v)(||y - y'||_1 + |x - x'| + W_1(m, m')),$$

   such that $\int_{\mathbb{R}} C(v)\nu_1(dv) < \infty$,
6. $$\int_{\mathbb{R}} \sup_{x,y,m} L\varphi(y, m, x, v)\nu_1(dv) < \infty.$$

In order to simplify the presentation, let us assume that the function G is of the form

$$G(\mu) = \int_{D^2} \mu \otimes \mu(d\gamma) \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr.$$

Now, let us show that $\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right]$ vanishes as N goes to infinity.

$$\begin{aligned}|G(\mu) - G(\mu^N)| \le &\left|G(\mu) - \int_{D^2} \mu^N \otimes \mu^N(d\gamma)\left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr\right)\right| \\ &+ \left|\int_{D^2} \mu^N \otimes \mu^N(d\gamma)\left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr\right)\right. \\ &\quad \left. - \int_{D^2} \mu^N \otimes \mu^N(d\gamma)\left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu^N_r(dx)dr\right)\right| \\ &+ \left|G(\mu^N) - \int_{D^2} \mu^N \otimes \mu^N(d\gamma)\left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu^N_r(dx)dr\right)\right| \\ &=: A_1 + A_2 + A_3.\end{aligned}$$

We first show that A_1 vanishes a.s. (this implies that $\mathbb{E}[A_1]$ vanishes by dominated convergence). A_1 is of the form

$$A_1 = \left| \int_{D^2} \mu \otimes \mu(d\gamma) H(\gamma) - \int_{D^2} \mu^N \otimes \mu^N(d\gamma) H(\gamma) \right|,$$

with

$$H : \gamma \in D^2 \mapsto \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv) \mu_r(dx) dr \in \mathbb{R}.$$

We just have to prove that H is continuous and bounded. The boundedness is obvious, so let us verify the continuity.

Let $(\gamma^n)_n$ converge to γ in $D(\mathbb{R}_+, \mathbb{R})^2$. We have

$$\begin{aligned} |H(\gamma) - H(\gamma^n)| \leq & \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} |H(\gamma_r, \mu_r, x, v) - H(\gamma^n_r, \mu_r, x, v)| \nu_1(dv) \mu_r(dx) dr \\ \leq & \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} C(v) ||\gamma_r - \gamma^n_r||_1 \nu_1(dv) \mu_r(dx) dr \\ \leq & C \int_s^t ||\gamma_r - \gamma^n_r||_1 dr, \end{aligned}$$

which vanishes by dominated convergence: the integrand vanishes at every continuity point r of γ (whence for a.e. r), and, for n big enough, $\sup_{r \leq t} ||\gamma^n_r||_1 \leq 2 \sup_{r \leq t} ||\gamma_r||_1$.

Now we show that $\mathbb{E}[A_2]$ vanishes. We have

$$\begin{aligned} A_2 \leq & \int_{D^2} \mu^N \otimes \mu^N(d\gamma) \\ & \left(\int_s^t \left| \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv) \mu_r(dx) - \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv) \mu^N_r(dx) \right| dr \right). \end{aligned}$$

Since the function $x \in \mathbb{R} \mapsto \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v) \nu_1(dv)$ is Lipschitz continuous (with Lipschitz constant independent of γ_r and μ_r), we have, by Kantorovich-Rubinstein duality (see e.g. Remark 6.5 of Villani (2008)),

$$A_2 \leq C \int_{D^2} \mu^N \otimes \mu^N(d\gamma) \int_s^t W_1(\mu^N_r, \mu_r) dr = C \int_s^t W_1(\mu^N_r, \mu_r) dr.$$

Hence

$$\mathbb{E}[A_2] \leq C \int_s^t \mathbb{E}\left[W_1(\mu^N_r, \mu_r) \right] dr,$$

which vanishes by dominated convergence: the integrand vanishes thanks to *Step 2*, and the uniform integrability follows from the fact that

$$\sup_N \int_s^t \mathbb{E}\left[W_1(\mu^N_r, \mu_r)^{3/2} \right] dr \leq C(t-s) \sup_N \sup_{0 \leq s \leq t} \mathbb{E}\left[\mu^N_s(|x|)^{3/2} \right] + C(t-s) \sup_{0 \leq s \leq t} \mathbb{E}\left[\mu_s(|x|)^{3/2} \right].$$

We finally show that $\mathbb{E}[A_3]$ vanishes.

$$A_3 \leq \int_{D^2} \mu^N \otimes \mu^N(d\gamma) \left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} \left| L\varphi(\gamma_r, \mu^N_r, x, v) - L\varphi(\gamma_r, \mu_r, x, v) \right| \nu_1(dv) \mu^N_r(dx) dr \right)$$

$$\leq \int_{D^2} \mu^N \otimes \mu^N(d\gamma) \left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} C(v) W_1(\mu_r^N, \mu_r) \nu_1(dv) \mu_r^N(dx) dr \right).$$

Then,

$$\mathbb{E}\left[A_3\right] \leq C \int_s^t \mathbb{E}\left[W_1(\mu_r^n, \mu_r)\right] dr,$$

which vanishes for the same reasons as in the previous step where we have shown that $\mathbb{E}\left[A_2\right]$ vanishes. □

Finally, we give the

Proof of Theorem 5.4.2. The beginning of the proof is similar to the proof of Theorem 4.4.2. It mainly consists in applying Lemma 5.4.5 and Theorem 5.4.6.

Let μ be the limit in distribution of some converging subsequence of (μ^N) (that we still note (μ^N)). Then, by Proposition (7.20) of Aldous, μ is the directing measure of an exchangeable system $(\bar{Y}^i)_{i\geq 1}$, and $(X^{N,i})_{1\leq i\leq N}$ converges in distribution to $(\bar{Y}^i)_{i\geq 1}$.

According to Theorem 5.4.6 and Lemma 5.4.5 for every $i \neq j$, there exist, on an extension, Brownian motions $\beta^{i,j,1}, \beta^{i,j,2}$ and white noises $W^{i,j,1}, W^{i,j,2}, W^{i,j}$ with respective intensities $dt\cdot dp$, $dt \cdot dp$ and $dt \cdot dp \cdot \nu_1(dv)$ all independent such that

$$\begin{aligned}
d\bar{Y}_t^i =& b(\bar{Y}_t^i, \mu_t)dt + \sigma(\bar{Y}_t^i, \mu_t) d\beta_t^{i,j,1} \\
&+ \int_0^1 \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)} \tilde{\Psi}(F_s^{-1}(p), \bar{Y}_t^i, \mu_t, v) dW^{i,j}(t,p,v) \\
&+ \int_0^1 \sqrt{f(F_s^{-1}(p), \mu_t)} \kappa(F_s^{-1}(p), \bar{Y}_t^i, \mu_t) dW^{i,j,1}(t,p), \\
d\bar{Y}_t^j =& b(\bar{Y}_t^j, \mu_t)dt + \sigma(\bar{Y}_t^j, \mu_t) d\beta_t^{i,j,2} \\
&+ \int_0^p \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)} \tilde{\Psi}(F_s^{-1}(p), \bar{Y}_t^j, \mu_t, v) dW^{i,j}(t,p,v) \\
&+ \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)} \kappa(F_s^{-1}(p), \bar{Y}_t^j, \mu_t) dW^{i,j,2}(t,p),
\end{aligned}$$

where F_s is the distribution function related to μ_s, and F_s^{-1} its generalized inverse. We can construct this extension in a global way such that it works for all couples (i,j) simultaneously.

As the system $(\bar{Y}^i)_{i\geq 1}$ is exchangeable, we know that $W^{i,j,1} = W^i$, $W^{i,j,2} =: W^j$, $W^{i,j} =: W$, $\beta^{i,j,1} =: \beta^i$ and $\beta^{i,j,2} =: \beta^j$.

It remains to identify the structure of μ_t appearing above as the one prescribed in (5.4) as conditional law. This is what we are going to do now. To do so, we introduce the following auxiliary system

$$\begin{aligned}
dZ_t^{N,i} =& b(\tilde{Z}_t^{N,i}, \mu_t^{Z,N})dt + \sigma(Z_t^{N,i}, \mu_t^{Z,N}) d\beta_t^i \\
&+ \int_0^1 \int_{\mathbb{R}} \sqrt{f((F_s^{Z,N})^{-1}(p), \mu_t^{Z,N})} \tilde{\Psi}((F_s^{Z,N})^{-1}(p), Z_t^{N,i}, \mu_t^{Z,N}, v) dW(t,p,v) \\
&+ \int_0^1 \sqrt{f((F_s^{Z,N})^{-1}(p), \mu_t^{Z,N})} \kappa((F_s^{Z,N})^{-1}(p), Z_t^{N,i}, \mu_t^{Z,N}) dW^i(t,p),
\end{aligned}$$

where $\mu^{Z,N} := N^{-1}\sum_{i=1}^N \delta_{Z^{N,i}}$, $F^{Z,N}$ is the distribution function related to $\mu^{Z,N}$. In addition, in the rest of the proof, let $\mu^X := \mathcal{L}(\bar{X}^1|W)$ and μ^Y be the directing measure of $(\bar{Y}^i)_{i\geq 1}$ that was denoted by μ so far in this proof.

With similar computations as in Theorem 5.3.1, one can prove that, for all $t \geq 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leq C(1+t)\int_0^t \mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right] ds + C(1+t)\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{Z,N})^2\right] ds.$$

Besides, introducing $\mu^{X,N} := N^{-1}\sum_{i=1}^N \delta_{\bar{X}^i}$, we have

$$\begin{aligned}\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{Z,N})^2\right] \leq & 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[W_1(\mu_s^{X,N}, \mu_s^{Z,N})^2\right] \\ \leq & 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[|Z_s^{N,i} - \bar{X}_s^i|\right]^2 \\ \leq & 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right].\end{aligned}$$

As a consequence, we obtain the following inequality, for all $t \geq 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leq C(1+t)\int_0^t \mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right] ds + C_t\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] ds,$$

with C_t a constant depending on t. Grönwall's lemma now implies that, for all $t \geq 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leq C_t\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] ds. \tag{5.31}$$

Now, let us prove that the expression above vanishes by dominated convergence. Note that $\mu_s^{X,N}$ converges weakly to μ_s^X a.s. by Glivenko-Cantelli's theorem (applying the theorem conditionally on $\sigma(W)$) recalling that the variables $\bar{X}_s^j$ are conditionally i.i.d. given $\sigma(W)$. And $\mu_s^{X,N}(|x|)$ converges to $\mu_s^X(|x|)$ a.s. as a consequence of the strong law of large numbers applied conditionally on $\sigma(W)$ (once again because of the conditional independence property). Hence the characterization (i) of Definition 6.8 and Theorem 6.9 of Villani (2008) implies that $W_1(\mu_s^{X,N}, \mu_s^X)$ vanishes a.s. as N goes to infinity for every $s \geq 0$.

We also have, by Jensen's inequality and the definition of W^1, for all $t \geq 0$,

$$\begin{aligned}\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^4\right] \leq C\mathbb{E}\left[\mu_s^X(x^4)\right] + C\mathbb{E}\left[\mu_s^{X,N}(x^4)\right] = C\mathbb{E}\left[(\bar{X}_s^1)^4\right] + \mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N (\bar{X}_s^j)^4\right] \\ \leq C\mathbb{E}\left[(\bar{X}_s^1)^4\right].\end{aligned}$$

Consequently

$$\sup_{0\leq s\leq t} \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^4\right] \leq C \sup_{0\leq s\leq t} \mathbb{E}\left[(\bar{X}_s^1)^4\right] < \infty,$$

recalling (5.14).

So we just have just proven that, the sequence of variables $W_1(\mu_s^X, \mu_s^{X,N})^2$ vanishes a.s. and for every $s \geq 0$, and that this sequence is uniformly integrable on $\Omega \times [0,t]$ for any $t \geq 0$. This implies, by dominated convergence that

$$\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] ds \underset{N\to\infty}{\longrightarrow} 0.$$

Recalling (5.31), we have that

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \underset{N\to\infty}{\longrightarrow} 0. \tag{5.32}$$

With the same reasoning, we can prove that for all $t \geq 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{Y}_t^i)^2\right] \underset{N\to\infty}{\longrightarrow} 0. \tag{5.33}$$

Finally, using (5.32) and (5.33), we have that, for all $t \geq 0$,

$$\mathbb{E}\left[(\bar{Y}_t^i - \bar{X}_t^i)^2\right] = 0.$$

This proves that the systems $(\bar{X}^i)_{i\geq 1}$ and $(\bar{Y}^i)_{i\geq 1}$ are equal. Hence, recalling that $\mu =: \mu^Y$ is a limit of μ^N and also the directing measure of $(\bar{Y}^i)_{i\geq 1}$, it is also the directing measure of $(\bar{X}^i)_{i\geq 1}$, which is $\mu^X := \mathcal{L}(\bar{X}^1|W)$. □

5.5 Model of interacting populations

The aim of this section is to generalize the previous model, considering n populations instead of one. Within each population, the particles interact as in the previous model, and, in addition, there are interactions at the level of the populations. The chaoticity properties of this kind of model have been studied by Graham (2008). Two examples of such systems are given in Figure 5.1.

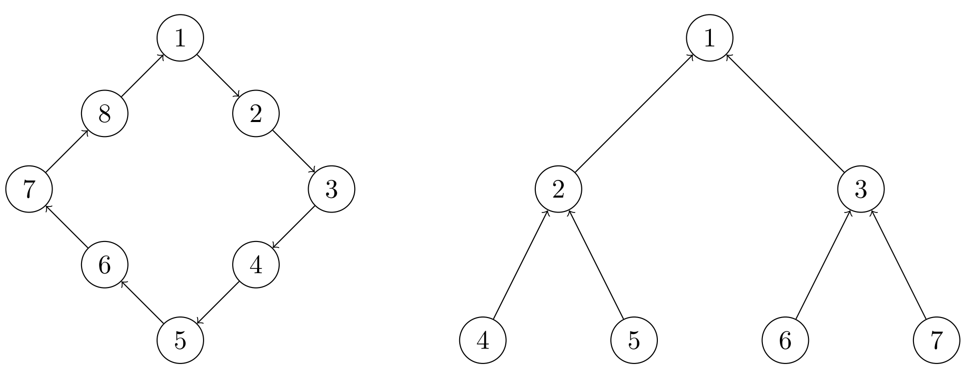

Figure 5.1: Two examples of interacting populations.

If we consider a number of N particles, we note, for each $1 \leq k \leq n$, N_k the number of particles of the $k-$th population. In particular $N = N_1 + ... + N_n$. We assume that, for all $1 \leq k \leq n$, N_k/N converges to some positive number, such that each population survives in the limit system.

For all $1 \leq k \leq n$, let $I(k) \subseteq [\![1, n]\!]$ be the set of populations that are "inputs" of the population k, that is, such that particles within these populations have a direct influence on those in population k. The dynamic of the $N-$particle system $(X^{N,k,i})_{\substack{1\leq k\leq n\\ 1\leq i\leq N_k}}$ is governed by the following equations.

$$dX^{N,k,i} = b^k\left(X_t^{N,k,i}, \mu_t^{N,k}\right)dt + \sigma^k\left(X_t^{N,k,i}, \mu_t^{N,k}\right)d\beta_t^{k,i}$$

$$+\sum_{l\in I(k)}\frac{1}{\sqrt{N_l}}\sum_{\substack{j=1\\(l,j)\neq(k,i)}}^{N_l}\int_{\mathbb{R}_+\times E^n}\Psi^{lk}(X^{N,l,j}_{t-},X^{N,k,i}_{t-},\mu^{N,l}_{t-},\mu^{N,k}_{t-},u^{l,j},u^{k,i})\mathbb{1}_{\left\{z\leq f^l(X^{N,l,j}_{t-},\mu^{N,l}_{t-})\right\}}d\pi^{l,j}(t,z,u).$$

In the above equation,

$$\mu^{N,k}_t=N_k^{-1}\sum_{j=1}^{N_k}\delta_{X^{N,k,j}_t},$$

$\pi^{l,j}$ $(1\leq l\leq n, j\geq 1)$ are independent Poisson measures of intensity $dtdz\nu(du)$, where ν is a probability measure on $(\mathbb{R}^{\mathbb{N}^*})^n$ which is of the form

$$\nu=(\nu^{1,1})^{\otimes\mathbb{N}^*}\otimes(\nu^{2,1})^{\otimes\mathbb{N}^*}\otimes...\otimes(\nu^{n,1})^{\otimes\mathbb{N}^*}.$$

The associated limit system $(\bar{X}^{k,i})_{\substack{1\leq k\leq n\\ i\geq 1}}$ is given by

$$\begin{aligned}d\bar{X}^{k,i}=&b^k\left(\bar{X}^{k,i}_t,\bar{\mu}^k_t\right)dt+\sigma^k\left(\bar{X}^{k,i}_t,\bar{\mu}^k_t\right)d\beta^{k,i}_t\\&+\sum_{l\in I(k)}\int_{\mathbb{R}}\int_{\mathbb{R}}\sqrt{f^l(x,\bar{\mu}^l_t)}\tilde{\Psi}^{lk}(x,\bar{X}^{k,i}_t,\mu^l_t,\mu^k_t,v)dM^l(t,x,v)\\&+\sum_{l\in I(k)}\int_{\mathbb{R}}\sqrt{f^l(x,\bar{\mu}^l_t)}\kappa^{lk}(x,\bar{X}^{k,i}_t,\mu^l_t,\mu^k_t)dM^{l,k,i}(t,x),\end{aligned}$$

with

$$\tilde{\Psi}^{lk}(x,y,m_1,m_2,v):=\int_{\mathbb{R}}\Psi^{lk}(x,y,m_1,m_2,v,w)d\nu^{k,1}(w),$$

$$\begin{aligned}&\kappa^{lk}(x,y,m_1,m_2)^2:=\int_{E^n}\Psi^{lk}(x,y,m_1,m_2,u^{l,1},u^{k,2})^2d\nu(u)-\int_{\mathbb{R}}\tilde{\Psi}^{lk}(x,y,m,v)^2d\nu^{l,1}(v)\\=&\int_{E^n}\Psi^{lk}(x,y,m_1,m_2,u^{l,1},u^{k,2})^2d\nu(u)-\int_{E^n}\Psi^{lk}(x,y,m,u^{l,1},u^{k,2})\Psi^{lk}(x,y,m,u^{l,1},u^{k,3})d\nu(u),\end{aligned}$$

and

$$M^{l,k,i}_t(A)=\int_0^t\mathbb{1}_A((F^l_s)^{-1}(p))dW^{l,k,i}(s,p)\text{ and }M^l_t(A\times B)=\int_0^t\mathbb{1}_A((F^l_s)^{-1}(p))\mathbb{1}_B(v)dW^l(s,p,v).$$

In the above formulas, $\mu^k_t:=\mathcal{L}\left(\bar{X}^{k,1}_t|\sigma\left(\bigcup_{l\in I(k)}\mathcal{W}^l\right)\right)$ and $(F^l_s)^{-1}$ is the generalized inverse of the function $F^l_s(x):=P(\bar{X}^{l,1}_s\leq x)$. Finally, $W^{l,k,i}$ and W^l $(1\leq l\leq n, k\in I(l), i\geq 1)$ are independent white noises of respective intensities $dsdp$ and $dsdp\nu^{l,1}(dv)$.

Both previous systems are "multi-exchangeable" in the sense that each population is "internally exchangeable" as in Corollary (3.9) of Aldous (1983). With the same reasoning as in the proof of Theorem 5.3.1, we can prove the existence of unique strong solutions $(X^{N,k,i})_{\substack{1\leq k\leq n\\1\leq i\leq N_k}}$ and $(\bar{X}^{k,i})_{\substack{1\leq k\leq n\\ i\geq 1}}$ as well as the convergence of the $N-$particle system to the limit system:

Theorem 5.5.1. *The following convergence in distribution in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n$ holds true:*

$$\left(\mu^{N,1},\mu^{N,2},...,\mu^{N,n}\right) \longrightarrow \left(\bar{\mu}^1,\bar{\mu}^2,...,\bar{\mu}^n\right),$$

as $N\to\infty$.

Before giving a sketch of the proof of Theorem 5.5.1, we quickly state the following result.

Proposition 5.5.2. *Let $r\leq n$ and $1\leq k_1<...<k_r\leq n$. If the sets $I(k_i)$ $(1\leq i\leq r)$ are disjoint, then the random variables μ^{k_i} $(1\leq i\leq r)$ are independent.*

Proof. For any $1\leq k\leq n$, the system $(\bar{X}^{k,i})_{i\geq 1}$ is conditionally i.i.d. given $\sigma\left(\bigcup_{l\in I(k)}\mathcal{W}^l\right)$. So, by Lemma (2.12) of Aldous (1983), μ^k is $\sigma\left(\bigcup_{l\in I(k)}\mathcal{W}^l\right)$ −measurable. □

Remark 5.5.3. *In the two examples of Figure 5.1, all the variables μ^k $(1\leq k\leq n)$ are independent.*

Coming back to Theorem 5.5.1, its proof is similar to the proof of Theorem 5.4.2. The main argument relies on a generalization of the martingale problem discussed in Section 5.4.2. Let us formulate it. Consider

$$\Omega' := \mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n\times\left(D(\mathbb{R}_+,\mathbb{R})^2\right)^n,$$

and write any atomic event $\omega'\in\Omega'$ as

$$\omega' = \left(\mu^1,\mu^2,...,\mu^n,Y^{1,1},Y^{1,2},Y^{2,1},Y^{2,2},...,Y^{n,1},Y^{n,2}\right) = (\mu,Y),$$

$\mu=(\mu^1,\ldots,\mu^n), Y=(Y^1,\ldots,Y^n)$.

For $Q\in\mathcal{P}(\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n)$, consider the law P' on Ω' defined by

$$P'(A\times B_1\times...\times B_n) = \int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n}\mathbb{1}_A(m)m^1\otimes m^1(B_1)...m^n\otimes m^n(B_n)Q(dm),$$

with A a Borel set of $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n$ and $B_1,...,B_n$ Borel sets of $D(\mathbb{R}_+,\mathbb{R})^2$.

Then, we say that Q is solution to our martingale problem if, for all $g\in C^2_b((\mathbb{R}^2)^n)$,

$$g(Y_t)-g(Y_0)-\int_0^t\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}Lg(Y_s,\mu_s,x,v)\mu^1_s\otimes...\otimes\mu^n_s(dx)\nu^{1,1}\otimes...\otimes\nu^{n,1}(dv)$$

is a martingale, where

$$\begin{aligned}Lg(y,m,x,u) = &\sum_{k=1}^n\sum_{i=1}^2 b^k(y^{k,i},m^k)\partial_{y^{k,i}}g(y)+\frac{1}{2}\sum_{k=1}^n\sum_{i=1}^2\sigma^k(y^{k,i},m^k)^2\partial^2_{y^{k,i}}g(y)\\ &+\frac{1}{2}\sum_{k=1}^n\sum_{l\in I(k)}\sum_{i=1}^2 f^l(x^l,m^l)\kappa^{lk}(x^l,y^{k,i},m^l,m^k)^2\partial^2_{y^{k,i}}g(y)\\ &+\frac{1}{2}\sum_{k_1,k_2=1}^n\sum_{i_1,i_2=1}^2\sum_{l\in I(k_1)\cap I(k_2)} f^l(x^l,m^l)\tilde{\Psi}^{lk_1}(x^l,y^{k_1,i_1},m^l,m^{k_1},u^l)\cdot\\ &\qquad\tilde{\Psi}^{lk_2}(x^l,y^{k_2,i_2},m^l,m^{k_2},u^l)\partial^2_{y^{k_1,i_1}y^{k_2,i_2}}g(y).\end{aligned}$$

Sketch of proof of Theorem 5.5.1. To prove the convergence in distribution of $(\mu^{N,1}, ..., \mu^{N,n})_N$, we begin by proving its tightness. Following the same reasoning as in the section 5.4.1, we can prove that, for each $1 \leq k \leq n$, the sequence $(\mu^{N,k})_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))$. Hence, the sequence $(\mu^{N,1}, ..., \mu^{N,n})_N$ on $\mathcal{P}(D(\mathbb{R}_+, \mathbb{R}))^n$.

Then a generalization of Lemma 5.4.5 allows to prove that the distribution of $(\bar{\mu}^1, ..., \bar{\mu}^n)$ is the unique solution of the martingale problem defined above.

Finally, we can conclude the proof showing that the law of any limit of a converging subsequence of $(\mu^{N,1}, ..., \mu^{N,n})_N$ is solution to the martingale problem using similar computations as the one in the proof of Theorem 5.4.6. □

Chapter 6

Well-posedness and propagation of chaos for McKean-Vlasov systems with locally Lipschitz coefficients in linear regime

This chapter is based on Erny (2021).

The aim of this chapter is to prove the strong well-posedness and a propagation of chaos property for Mckean-Vlasov equations. Contrary to the other models studied previously, in this chapter the strength of the interactions of the particle systems is linear: the interaction strength of a $N-$particle system is of order N^{-1}.

In this framework, the weak well-posedness and the propagation of chaos are classical for the McKean-Vlasov equations without jump term, even without assuming that the coefficients are Lipschitz continuous. Gärtner (1988) treats both questions in this frame. We can also mention more recent work on the well-posedness of McKean-Vlasov equations without jump terms as Mishura and Veretennikov (2020) and Chaudru de Raynal (2020), with different assumptions on the smoothness of the coefficients, and Lacker (2018) that investigates the well-posedness and the propagation of chaos.

Here we consider McKean-Vlasov equations with jumps. The questions about the strong well-posedness and the propagation of chaos have also been studied in this framework under globally Lipschitz assumptions on the coefficients: see Graham (1992) for the well-posedness and Andreis et al. (2018) for the propagation of chaos. Note that in Section 4 of Andreis et al. (2018), these questions are treated in a multi-dimensional case, where the drift coefficient is of the form $-\nabla b_1(x)+b_2(x,m)$, where b_1 is C^1 and convex, and b_2, as well as the jump coefficient and the volatility coefficient, are globally Lipschitz.

The novelty of our results is to work on McKean-Vlasov equations with jumps and with generic locally Lipschitz assumptions (see Assumption 6.1 for a precise and complete statement of the hypothesis).

The first main result is the strong well-posedness, under this locally Lipschitz assumption, of the following McKean-Vlasov equation

$$dX_t = b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \int_{\mathbb{R}_+ \times E} \Phi(X_{t-}, \mu_{t-}, u)\mathbb{1}_{\{z \le f(X_{t-}, \mu_{t-})\}} d\pi(t, z, u),$$

where μ_t is the distribution of X_t, W a Brownian motion, π a Poisson measure and E some measurable space (see the beginning of Section 6.1 for details on the notation). To prove the well-posedness when the coefficients are locally Lipschitz continuous, we adapt the computations of the proofs in the globally Lipschitz continuous case. We use a truncation argument to handle the dependency of the local Lipschitz constant w.r.t. to the variables. On the contrary of the globally Lipschitz case, Grönwall's lemma does not allow to conclude immediately. In the locally Lipschitz case, we have to use a generalization of this lemma: Osgood's lemma (see Lemma C.2.1). The uniqueness of solution of the McKean-Vlasov equation follows rather quickly from Osgood's lemma and the truncation argument, but other difficulties emerge in the proof of the existence of solution. We construct a weak solution of the equation using a Picard iteration scheme, but the fact that the coefficients are only locally Lipschitz continuous does not allow to prove that this scheme converges in an L^1-sense. Instead, we prove that a subsequence converges in distribution to some limit that is shown to be a solution of the equation. Some technical difficulties emerge in this part of the proof for two reasons. The first one is that the Picard scheme is not shown to converge but only to have a converging subsequence. This implies that we need to control the difference between two consecutive steps of the scheme. The second one is that we only prove a convergence in distribution, thus it is not straightforward that the limit of the Picard scheme is solution to the equation. It is shown studying its semimartingale characteristics.

The second main result is a propagation of chaos property of McKean-Vlasov particle systems under the same locally Lipschitz assumptions. More precisely, it is the convergence of the following $N-$particle system

$$\begin{aligned} dX_t^{N,i} =& b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i)\mathbb{1}_{\left\{z \le f(X_{t-}^{N,i}, \mu_{t-}^N)\right\}} d\pi^i(t, z, v) \\ &+ \frac{1}{N}\sum_{j=1}^N \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Theta(X_{t-}^{N,j}, X_{t-}^{N,i}, \mu_{t-}^N, v^j, v^i)\mathbb{1}_{\left\{z \le f(X_{t-}^{N,j}, \mu_{t-}^N)\right\}} d\pi^j(t, z, v), \end{aligned}$$

where $\mu_t^N := N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$, to the infinite system

$$\begin{aligned} d\bar{X}_t^i =& b(\bar{X}_t^i, \bar{\mu}_t)dt + \sigma(\bar{X}_t^i, \bar{\mu}_t)dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(\bar{X}_{t-}^i, \bar{\mu}_{t-}, v^i)\mathbb{1}_{\left\{z \le f(\bar{X}_{t-}^i, \bar{\mu}_{t-})\right\}} d\pi^i(t, z, v) \\ &+ \int_{\mathbb{R}} \int_{F^{\mathbb{N}^*}} \Theta(x, \bar{X}_t^i, \bar{\mu}_t, v^1, v^2) f(x, \bar{\mu}_t) d\nu(v) d\bar{\mu}_t(x), \end{aligned}$$

where $\bar{\mu}_t := \mathcal{L}(\bar{X}_t)$, as N goes to infinity. The W^i ($i \ge 1$) are independent standard Brownian motions, the π^i ($i \ge 1$) are independent Poisson measures, F is some measurable space and $\mathbb{N}^*$ denotes the set of the positive integers (see Section 6.2 for details on the notation). The proof of this second main result relies on a similar reasoning as the one used to prove the uniqueness of the McKean-Vlasov equation: a truncation argument and Osgood's lemma.

Let us note that this propagation of chaos property has already been proven under different hypothesis. Indeed, the $N-$particle system and the limit system above are the same as in Andreis et al. (2018). Note also that in this model, for each $N \in \mathbb{N}^*$, the particles $X_t^{N,i}$ $(1 \leq i \leq N)$ do not only interact through the empirical measure μ_t^N, but also through the simultaneous jumps term.

Let us mention the other conditions that we state in Assumption 6.1. We assume that some boundedness conditions on the coefficients hold true. To the best of our knowledge, in order to prove a priori estimates for solutions of Mckean-Vlasov equations, one needs to assume that the coefficients are bounded w.r.t. the measure variable. In this chapter, we need the coefficients to be bounded w.r.t. to both variables, because for the truncation arguments mentioned above, we need a priori estimates on the exponential moments of the solutions of the McKean-Vlasov equation. This is also the reason why we need conditions on exponential moments.

In Section 6.1, we state and prove our first main result: the well-posedness of the McKean-Vlasov equation (6.1) with locally Lipschitz coefficients. Section 6.2 is devoted to our second main result, the propagation of chaos in the same framework.

6.1 Well-posedness of McKean-Vlasov equations

This section is dedicated to prove the well-posedness of the following McKean-Vlasov equation.

$$dX_t = b(X_t, \mu_t)dt + \sigma(X_t, \mu_t)dW_t + \int_{\mathbb{R}_+ \times E} \Phi(X_{t-}, \mu_{t-}, u)\mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}} d\pi(t, z, u), \tag{6.1}$$

with $\mu_t = \mathcal{L}(X_t)$, W a standard one-dimensional Brownian motion, π a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}_+ \times E$ having intensity $dt \cdot dz \cdot d\rho(u)$, where $(E, \mathcal{E}, \rho)$ is a $\sigma-$finite measure space. The assumptions on the coefficients are specified in Assumption 6.1 below. Let us note here that f is assumed to be non-negative.

Assumption 6.1.

1. *Locally Lipschitz conditions: there exists* $a > 0$, *such that for all* $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$,

$$|b(x_1, m_1) - b(x_2, m_2)| + \int_E \int_{\mathbb{R}_+} |\Phi(x_1, m_1, u)\mathbb{1}_{\{z \leq f(x_1, m_1)\}} - \Phi(x_2, m_2, u)\mathbb{1}_{\{z \leq f(x_2, m_2)\}}| dz d\rho(u)$$
$$\leq L\left(1 + |x_1| + |x_2| + \int_{\mathbb{R}} e^{a|x|} dm_1(x) + \int_{\mathbb{R}} e^{a|x|} dm_2(x)\right)(|x_1 - x_2| + W_1(m_1, m_2)).$$

2. *Globally Lipschitz condition for* σ : *for all* $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$,

$$|\sigma(x_1, m_1) - \sigma(x_2, m_2)| \leq L(|x_1 - x_2| + W_1(m_1, m_2)).$$

3. *Boundedness conditions: the functions* b, σ *and* f *are bounded (uniformly w.r.t. all the variables), and for the same constant* $a > 0$ *as in Item 1.*,

$$\sup_{x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_E e^{a|\Phi(x,m,u)|} d\rho(u) < \infty.$$

4. *Initial condition: for the same* $a > 0$ *as in Items 1. and 3.*,

$$\mathbb{E}\left[e^{a|X_0|}\right] < \infty.$$

Remark 6.1.1. *If we consider equation (6.1) without the jump term (that is $\Phi \equiv 0$), then, we can adapt the proof of Theorem 6.1.4 to the case where σ is also locally Lipschitz continuous. More precisely, we can replace the two first Items of Assumption 6.1 by: for all $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$,*

$$\begin{aligned}&|b(x_1, m_1) - b(x_2, m_2)| + |\sigma(x_1, m_1) - \sigma(x_2, m_2)| \\ &\leq L \left(1 + \sqrt{|x_1|} + \sqrt{|x_2|} + \sqrt{\int_{\mathbb{R}} e^{a|x|} dm_1(x)} + \sqrt{\int_{\mathbb{R}} e^{a|x|} dm_2(x)}\right) (|x_1 - x_2| + W_1(m_1, m_2)).\end{aligned}$$

See Remark 6.1.7 for more details on the adaptation of the proof.

Note that Item 3 of Assumption 6.1 implies that, for all $n \in \mathbb{N}$,

$$\sup_{x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})} \int_E |\Phi(x, m, u)|^n \, d\rho(u) < \infty.$$

Remark 6.1.2. *Under the boundedness conditions of Assumption 6.1, a sufficient condition to obtain the locally Lipschitz condition of the jump term in Assumption 6.1 is given by the following direct conditions on the functions f and Φ : for all $x_1, x_2 \in \mathbb{R}, m_1, m_2 \in \mathcal{P}_1(\mathbb{R})$,*

$$\begin{aligned}&|f(x_1, m_1) - f(x_2, m_2)| + \int_E |\Phi(x_1, m_1, u) - \Phi(x_2, m_2, u)| d\rho(u) \\ &\qquad \leq L \left(1 + |x_1| + |x_2| + \int_{\mathbb{R}} e^{a|x|} dm_1(x) + \int_{\mathbb{R}} e^{a|x|} dm_2(x)\right) (|x_1 - x_2| + W_1(m_1, m_2)).\end{aligned}$$

Example 6.1.3. *A natural form of the coefficient of a McKean-Vlasov equation is given by the so-called "true McKean-Vlasov" case. For simplicity, we only give the form for b, but a similar form can be considered for the other coefficient.*

$$b(x, m) = \int_{\mathbb{R}} \tilde{b}(x, y) dm(y),$$

with $\tilde{b} : \mathbb{R}^2 \to \mathbb{R}$.

For b to satisfy the conditions of Assumption 6.1 in this example, it is sufficient to assume that $\tilde{b}$ is bounded and that: for all $x, x', y, y' \in \mathbb{R}$,

$$|\tilde{b}(x, y) - \tilde{b}(x', y')| \leq C(1 + |x| + |x'|)(|x - x'| + |y - y'|).$$

Indeed, for any $x, x' \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$\begin{aligned}|b(x, m) - b(x', m')| \leq& |b(x, m) - b(x', m)| + |b(x', m) - b(x', m')| \\ \leq& \int_{\mathbb{R}} |\tilde{b}(x, y) - \tilde{b}(x', y)| dm(y) + \left| \int_{\mathbb{R}} \tilde{b}(x', y) dm(y) - \int_{\mathbb{R}} \tilde{b}(x', y) dm'(y) \right| \\ \leq& C(1 + |x| + |x'|)|x - x'| + C(1 + 2|x'|) W_1(m, m'),\end{aligned}$$

where the second quantity of the last line has been obtained using Kantorovich-Rubinstein duality (see Remark 6.5 of Villani (2008)) and the fact that, for a fixed x', the function $y \mapsto \tilde{b}(x', y)$ is Lipschitz continuous with Lipschitz constant $C(1 + 2|x'|)$.

Theorem 6.1.4. *Under Assumption 6.1, there exists a unique strong solution of (6.1).*

The rest of this section is dedicated to prove Theorem 6.1.4.

6.1.1 A priori estimates for equation (6.1)

In this section, we prove the following a priori estimates for the solutions of the stochastic differential equation (6.1).

Lemma 6.1.5. *Grant the boundedness conditions and the initial condition of Assumption 6.1. Any solution $(X_t)_{t\geq 0}$ of (6.1) satisfies for all $t > 0$,*

(i) $\sup_{0\leq s\leq t} \mathbb{E}\left[e^{a|X_s|}\right] < \infty$, *with $a > 0$ the same as constant as in Assumption 6.1,*

(ii) $\mathbb{E}\left[\sup_{0\leq s\leq t} |X_s|\right] < \infty$.

Proof. Let us prove (i). It is sufficient to prove

$$\sup_{0\leq s\leq t} \mathbb{E}\left[e^{aX_s}\right] < \infty, \tag{6.2}$$

and

$$\sup_{0\leq s\leq t} \mathbb{E}\left[e^{-aX_s}\right] < \infty. \tag{6.3}$$

To prove (6.2), let us apply Ito's formula

$$\begin{aligned} e^{aX_t} =& e^{aX_0} + a\int_0^t e^{aX_s} b(X_s,\mu_s)ds + a\int_0^t e^{aX_s}\sigma(X_s,\mu_s)dW_s + \frac{a^2}{2}\int_0^t e^{aX_s}\sigma(X_s,\mu_s)^2 ds \\ &+ \int_{[0,t]\times\mathbb{R}_+\times E}\left[e^{a(X_{s-}+\Phi(X_{s-},\mu_{s-},u))} - e^{aX_{s-}}\right]\mathbb{1}_{\{z\leq f(X_{s-},\mu_{s-})\}} d\pi(s,z,u). \end{aligned}$$

Introducing, for $M > 0$, the stopping time $\tau_M := \inf\{t > 0 : |X_t| > M\}$, we have

$$\begin{aligned} \mathbb{E}\left[e^{aX_{t\wedge\tau_M}}\right] \leq& \mathbb{E}\left[e^{aX_0}\right] + a||b||_\infty \int_0^t \mathbb{E}\left[e^{aX_{s\wedge\tau_M}}\right] ds + \frac{1}{2}a^2||\sigma||_\infty^2 \int_0^t \mathbb{E}\left[e^{aX_{s\wedge\tau_M}}\right] ds \\ &+ \int_0^t\int_E e^{aX_{s\wedge\tau_M}}\left[e^{\Phi(X_{s\wedge\tau_M},\mu_{s\wedge\tau_M},u)} - 1\right] f(X_{s\wedge\tau_M},\mu_{s\wedge\tau_M}) d\rho(u) ds. \end{aligned}$$

Then, introducing $u_t^M := \mathbb{E}\left[e^{aX_{t\wedge\tau_M}}\right]$, and using the boundedness condition of Φ from Assumption 6.1, we obtain, for all $t > 0$,

$$u_t^M \leq \mathbb{E}\left[e^{aX_0}\right] + K\int_0^t u_s^M ds,$$

with

$$K := a||b||_\infty + \frac{1}{2}a^2||\sigma||_\infty^2 + ||f||_\infty \sup_{x\in\mathbb{R}, m\in\mathcal{P}_1(\mathbb{R})} \int_E e^{a|\Phi(x,m,u)|} d\rho(u) < \infty.$$

Consequently, Grönwall's lemma implies

$$\sup_{0\leq s\leq t} u_s^M \leq \mathbb{E}\left[e^{aX_0}\right] e^{Kt}.$$

As the bound above does not depend on M, it implies that τ_M goes to infinity almost surely as M goes to infinity. Fatou's lemma then implies

$$\sup_{0\le s\le t} \mathbb{E}\left[e^{aX_t}\right] \le \mathbb{E}\left[e^{aX_0}\right] e^{Kt}.$$

This proves (6.2), and with the same reasoning we can prove

$$\sup_{0\le s\le t} \mathbb{E}\left[e^{-aX_t}\right] \le \mathbb{E}\left[e^{-aX_0}\right] e^{Kt}.$$

which proves (6.3), whence the point (i) of the lemma.

To prove the point (ii), let us use the following bound, which is a direct consequence of the form of the equation (6.1),

$$\mathbb{E}\left[\sup_{0\le s\le t}|X_s|\right] \le \mathbb{E}\left[|X_0|\right] + ||b||_\infty t + \mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \sigma(X_t,\mu_r)dW_r\right|\right] + ||f||_\infty \int_0^t \mathbb{E}\left[\int_E |\Phi(X_s,\mu_s,u)|d\rho(u)\right] ds.$$

Then the result follows from Burkholder-Davis-Gundy's inequality and the boundedness conditions of σ and Φ from Assumption 6.1. □

6.1.2 Pathwise uniqueness for equation (6.1)

Proposition 6.1.6. *Grant Assumption 6.1. The pathwise uniqueness property holds true for (6.1).*

Proof. Let $(\hat{X}_t)_{t\ge 0}$ and $(\check{X}_t)_{t\ge 0}$ be two solutions of (6.1) defined w.r.t. the same initial condition X_0, the same Brownian motion W and the same Poisson measure π. The proof consists in showing that the function

$$u(t) := \mathbb{E}\left[\sup_{0\le s\le t}|\hat{X}_s - \check{X}_s|\right]$$

is zero. This choice of function u is inspired of the proof of Theorem 2.1 of Graham (1992). It allows to treat equations with both a jump term and a Brownian term.

We know that, for all $t\ge 0$, $u_t < \infty$ by Lemma 6.1.5.(ii).

Writing $\hat{\mu}_t := \mathcal{L}(\hat{X}_t)$ and $\check{\mu}_t := \mathcal{L}(\check{X}_t)$, we have

$$\left|\hat{X}_s - \check{X}_s\right| \le \int_0^s |b(\hat{X}_r,\hat{\mu}_r) - b(\check{X}_r,\check{\mu}_r)|dr + \left|\int_0^s (\sigma(\hat{X}_r,\hat{\mu}_r) - \sigma(\check{X}_r,\check{\mu}_r))dW_r\right| + \int_{[0,s]\times\mathbb{R}_+\times\mathbb{R}} \left|\Phi(\hat{X}_{r-},\hat{\mu}_{r-},u)\mathbb{1}_{\left\{z\le f(\hat{X}_{r-},\hat{\mu}_{r-})\right\}} - \Phi(\check{X}_{r-},\check{\mu}_{r-},u)\mathbb{1}_{\left\{z\le f(\check{X}_{r-},\check{\mu}_{r-})\right\}}\right| d\pi(r,z,u).$$

This implies that

$$u(t) \le L\mathbb{E}\left[\left(\int_0^t \left(|\hat{X}_s - \check{X}_s| + W_1(\hat{\mu}_s,\check{\mu}_s)\right)^2 ds\right)^{1/2}\right]$$

$$+ 2L \int_0^t \mathbb{E}\left[\left(1 + |\hat{X}_s| + |\check{X}_s| + \int_{\mathbb{R}} e^{a|x|} d\hat{\mu}_s(x) + \int_{\mathbb{R}} e^{a|x|} d\check{\mu}_s(x)\right)\left(|\hat{X}_s - \check{X}_s| + W_1(\hat{\mu}_s, \check{\mu}_s)\right)\right] ds \tag{6.4}$$

where we have used Burkholder-Davis-Gundy's inequality to deal with the Brownian term that corresponds to the term at the first line above. The term at the first line corresponds to the controls of the drift term and the jump term.

By Lemma 6.1.5.(i), we have for all $t \geq 0$,

$$\sup_{0 \leq s \leq t} \int_{\mathbb{R}} e^{a|x|} d\hat{\mu}_s(x) + \sup_{0 \leq s \leq t} \int_{\mathbb{R}} e^{a|x|} d\check{\mu}_s(x) \leq C_t < \infty.$$

And, from the definition of W_1, we have the following bound

$$W_1(\hat{\mu}_s, \check{\mu}_s) \leq \mathbb{E}\left[\left|\hat{X}_s - \check{X}_s\right|\right] \leq u(s).$$

Then, (6.4) and Lemma 6.1.5 imply that, for all $0 \leq t \leq T$,

$$\begin{aligned}
u(t) \leq & \int_0^t \mathbb{E}\left[\left(1 + |\hat{X}_s| + |\check{X}_s| + C_T\right)\left(|\hat{X}_s - \check{X}_s| + u(s)\right)\right] ds \\
& + L\mathbb{E}\left[\left(\int_0^t \left(|\hat{X}_s - \check{X}_s| + u(s)\right)^2 ds\right)^{1/2}\right] \\
\leq & C_T \int_0^t \mathbb{E}\left[\left(1 + |\hat{X}_s| + |\check{X}_s|\right)\left(|\hat{X}_s - \check{X}_s| + u(s)\right)\right] ds + 2L\sqrt{t}u(t) \\
\leq & C_T \int_0^t \mathbb{E}\left[\left(1 + |\hat{X}_s| + |\check{X}_s|\right)\left(|\hat{X}_s - \check{X}_s|\right)\right] ds + C_T \int_0^t u(s) ds + 2L\sqrt{t}u(t),
\end{aligned}$$

where we have bounded the second integral of the RHS of the first inequality above by t times the supremum of the integrand. Note that the value of C_T changes from line to line.

Now, to end the proof, we have to control a term of the type $(1 + |x| + |y|)|x - y|$. To do so, we use a truncation argument based on the following inequality: for all $x, y \in \mathbb{R}, R > 0$,

$$(1 + |x| + |y|)|x - y| \leq (1 + 2R)|x - y| + (1 + |x| + |y|)|x - y| \left(\mathbb{1}_{\{|x| > R\}} + \mathbb{1}_{\{|y| > R\}}\right).$$

Let $R : s \mapsto R_s > 0$ be the truncation function whose values will be chosen later. By Lemma 6.1.5, for any $0 \leq s \leq T$,

$$\mathbb{E}\left[\left(1 + |\hat{X}_s| + |\check{X}_s|\right)\left|\hat{X}_s - \check{X}_s\right|\right] \leq (1 + 2R_s)u(s) + C_T\sqrt{\mathbb{P}\left(|\hat{X}_s| > R_s\right)} + C_T\sqrt{\mathbb{P}\left(|\check{X}_s| > R_s\right)}.$$

The exponential moments proven in Lemma 6.1.5.(i) are used to control the two last term above. Indeed, by Markov's inequality

$$\mathbb{P}\left(|\hat{X}_s| > R_s\right) + \mathbb{P}\left(|\check{X}_s| > R_s\right) \leq C_T e^{-aR_s}.$$

Consequently, defining $r_s := aR_s/2$, for any $0 \leq t \leq T$,

$$u(t) \leq C_T \int_0^t \left[(1 + r_s)u(s) + e^{-r_s}\right] ds + 2L\sqrt{t}u(t).$$

Now, let $T = 1/(16L^2)$ such that $2L\sqrt{T} \leq 1/2$. Then, we can rewrite the above inequality as, for all $t \in [0, T]$,

$$u(t) \leq C_T \int_0^t \left[(1 + r_s)u(s) + e^{-r_s}\right] ds.$$

Let us prove by contradiction that, for all $t \leq T$, $u(t) = 0$. To do so, let $t_0 := \inf\{t > 0 : u(t) > 0\}$ and assume that $t_0 < T$. Notice that, as u is non-decreasing, this implies that, for all $t \in [0, t_0[$, $u(t) = 0$. In particular, for all $t \in [t_0, T]$,

$$u(t) \leq C_T \int_{t_0}^t \left[(1 + r_s)u(s) + e^{-r_s}\right] ds.$$

Besides $u(t)$ is finite and bounded (see Lemma 6.1.5.(ii)) on $[0, T]$, say by a constant $D > 1$. Let $v(t) := u(t)/(De^2) < e^{-2}$. Obviously v satisfies the same inequality as u above. Now we define $r_s := -\ln v(s)$, so that, for all $t_0 < t \leq T$,

$$v(t) \leq C_T \int_{t_0}^t (2 - \ln v(s))v(s)ds \leq -2C_T \int_{t_0}^t v(s) \ln v(s)ds,$$

where we have used that $2 - \ln v(s) \leq -2 \ln v(s)$ since $-\ln v(s) \geq 2$.

In particular, for any $c \in]0, e^{-2}[$, for all $t_0 \leq t \leq T$,

$$v(t) \leq c - 2C_T \int_{t_0}^t v(s) \ln v(s)ds.$$

Then, introducing $M : x \in]0, e^{-2}[\mapsto \int_x^{e^{-2}} \frac{1}{-s \ln s} ds$, we may apply Osgood's lemma (see Lemma C.2.1) with $\gamma \equiv 2C_T$ and $\mu(v) = (-\ln v)v$ to obtain that

$$-M(v(T)) + M(c) \leq \int_{t_0}^T 2C_T ds = 2C_T(T - t_0)$$

or equivalently,

$$M(c) \leq M(v(T)) + 2C_T T.$$

Recalling that we assumed $v(T) > 0$ such that the right hand side of the above equality is finite, if we let c tend to 0, we obtain

$$M(0) = \int_0^{e^{-2}} \frac{1}{-s \ln s} ds \leq \int_{v(T)}^{e^{-2}} \frac{1}{-s \ln s} ds + 2C_T T < \infty,$$

which is absurd since $M(0) = \infty$.

A consequence of the above considerations is that for all $t \in [0, T]$, $u(t) = 0$. Recalling the definition of u, we have proven that the processes $\hat{X}$ and $\check{X}$ are equal on $[0, T]$.

We can repeat this argument on the interval $[T, 2T]$ and iterate up to any finite time interval $[0, T_0]$ since $T = 1/(16L^2)$ does only depend on the coefficients of the system but not on the initial condition. This proves the pathwise-uniqueness property for the McKean-Vlasov equation (6.1). □

Let us complete our previous Remark 6.1.1.

Remark 6.1.7. *The adaptation suggested in Remark 6.1.1 is the following: in the proof of Proposition 6.1.6 above, one has to replace the distance* $\mathbb{E}\left[\sup_{s \leq t} |\hat{X}_s - \check{X}_s|\right]$ *by* $\mathbb{E}\left[(\hat{X}_t - \check{X}_t)^2\right]$, *and one has to do similar changes in the proof of Proposition 6.1.8 below.*

6.1.3 Existence of a weak solution of equation (6.1)

The aim of this section is to construct a weak solution of the McKean-Vlasov equation (6.1), using a Picard iteration. However, because of our locally Lipschitz conditions, we cannot prove it directly. Instead, we prove that a subsequence converges in distribution by tightness. That is why, in a first time, we only construct a weak solution.

Proposition 6.1.8. *Grant Assumption 6.1. There exists a weak solution of* (6.1) *on* $[0,T]$, *with* $T = 1/(16L^2)$.

Proof. As in the proof of the pathwise uniqueness of Section 6.1.2, we work on a time interval $[0,T]$ where $T > 0$ is a number whose value can be fixed at $1/(16L^2)$.

Step 1. In this first step, we introduce the iteration scheme, and state its basic properties at (6.5). Let $X_t^{[0]} := X_0$, and define the process $X^{[n+1]}$ from $X^{[n]}$ and $\mu_t^{[n]} := \mathcal{L}(X_t^{[n]})$ by

$$\begin{aligned} X_t^{[n+1]} :=& X_0 + \int_0^t b(X_s^{[n]}, \mu_s^{[n]})ds + \int_0^t \sigma(X_s^{[n]}, \mu_s^{[n]})dW_s \\ &+ \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \Phi(X_{s-}^{[n]}, \mu_{s-}^{[n]}, u)\mathbb{1}_{\left\{z\le f(X_{s-}^{[n]},\mu_{s-}^{[n]})\right\}} d\pi(s,z,u) \end{aligned}$$

Note that, thanks to the boundedness conditions of Assumption 6.1, using the same computations as in the proof of Lemma 6.1.5, we can prove that, for all $t \geq 0$,

$$\sup_{n\in\mathbb{N}} \sup_{0\le s\le t} \mathbb{E}\left[e^{a|X_s^{[n]}|}\right] < \infty \text{ and } \sup_{n\in\mathbb{N}} \mathbb{E}\left[\sup_{0\le s\le t} |X_s^{[n]}|\right] < \infty. \tag{6.5}$$

Step 2. Now let us show that $(X^{[n]}, X^{[n+1]})_n$ has a converging subsequence in distribution in $D([0,T],\mathbb{R}^2)$, by showing that it satisfies Aldous' tightness criterion:

- (a) for all $\varepsilon > 0$, $\lim_{\delta\downarrow 0} \limsup_{N\to\infty} \sup_{(S,S')\in A_{\delta,T}} P(|X_{S'}^{[n]} - X_S^{[n]}| + |X_{S'}^{[n+1]} - X_S^{[n]}| > \varepsilon) = 0$, where $A_{\delta,T}$ is the set of all pairs of stopping times (S,S') such that $0 \le S \le S' \le S+\delta \le T$ a.s.,
- (b) $\lim_{K\uparrow\infty} \sup_n \mathbb{P}(\sup_{t\in[0,T]} |X_t^{[n]}| + |X_t^{[n+1]}| \ge K) = 0$.

Assertion (b) is a straightforward consequence of (6.5) and Markov's inequality. To check assertion (a), notice that, for any $(S,S') \in A_{\delta,T}$, by BDG inequality,

$$\mathbb{E}\left[\left|X_{S'}^{[n+1]} - X_S^{[n+1]}\right|\right] \le ||b||_\infty\delta + ||\sigma||_\infty\sqrt{\delta} + \delta||f||_\infty \sup_{0\le s\le T} \mathbb{E}\left[\int_E |\Phi(X_s^{[n]}, \mu_s^{[n]}, u)|d\rho(u)\right]. \tag{6.6}$$

Then, by tightness, there exists a subsequence of $(X^{[n]}, X^{[n+1]})_n$ that converges in distribution to some (X,Y) in $D([0,T],\mathbb{R}^2)$. In the rest of the proof, we work on this subsequence without writing it explicitly for the sake of notation.

Step 3. In this step, we show that $X = Y$ almost surely. Note that, since we work on a subsequence, this is not obvious. It is for this part of the proof that we need to restrict our processes to a time interval of the form $[0,T]$. We show in Lemma C.3.1 that it is sufficient to prove that, for a subsequence,

$$\mathbb{E}\left[\sup_{0\le s\le T}\left|X_s^{[n+1]} - X_s^{[n]}\right|\right] \underset{n\to\infty}{\longrightarrow} 0. \tag{6.7}$$

Indeed, (6.7) implies that, for another subsequence, $\sup_{s\le T}|X_s^{[n+1]}-X_s^{[n]}|$ converges to zero almost surely. Then, we can apply Skorohod representation theorem (see Theorem 6.7 of Billingsley (1999)) to the following sequence

$$\left(X^{[n]},X^{[n+1]}\right)_n$$

that converges in distribution in $D([0,T],\mathbb{R}^2)$ to (X,Y). Thus we can consider, for $n\in\mathbb{N}$, random variables $(\tilde{X}^{[n]},\tilde{X}^{[n+1]})$ (resp. $(\tilde{X},\tilde{Y})$) having the same distribution as $(X^{[n]},X^{[n+1]})$ (resp. (X,Y)) for which the previous convergence is almost sure. In particular, we also know that $\sup_{s\le T}|\tilde{X}_s^{[n+1]}-\tilde{X}_s^{[n]}|$ vanishes almost surely. Hence, by Lemma C.3.1, the representing r.v. $(\tilde{X}^{[n]},\tilde{X}^{[n+1]})$ converges a.s. to the representing r.v. $(\tilde{X},\tilde{X})$ in $D([0,T],\mathbb{R}^2)$. As a consequence $\tilde{X}=\tilde{Y}$ almost surely, and so $X=Y$ almost surely.

Now let us prove (6.7). Let

$$u^{[n]}(t):=\mathbb{E}\left[\sup_{0\le s\le t}\left|X_s^{[n+1]}-X_s^{[n]}\right|\right].$$

By (6.5),

$$\sup_{n\in\mathbb{N}}u^{[n]}(t)<\infty.$$

Let us fix some $n\in\mathbb{N}^*$ and consider a truncation function $r_t^{[n]}>0$ whose values will be fixed later. The same truncation argument used in Section 6.1.2 allows to prove that, for all $0\le k\le n-1, t\le T$,

$$\begin{aligned}u^{[k+1]}(t)\le& C_T\int_0^t\left[(1+r_s^{[n]})u^{[k]}(s)+e^{-r_s^{[n]}}\right]ds+2L\sqrt{T}u^{[k]}(t)\\ \le& C_T\int_0^t\left[(1+r_s^{[n]})u^{[k]}(s)+e^{-r_s^{[n]}}\right]ds+\frac{1}{2}u^{[k]}(t).\end{aligned}$$

where $C_T>0$ does not depend on n thanks to (6.5). The second inequality above comes from the fact that we fix the value of $T>0$ such that $L\sqrt{T}<1/4$.

Now, introducing $S_n(t):=\sum_{k=0}^n u_t^{[k]}$ and summing the above inequality from $k=0$ to $k=n-1$, we have, for all $t\le T$,

$$S_n(t)\le C_T+C_T\int_0^t\left[(1+r_s^{[n]})S_n(s)ds+ne^{-r_s^{[n]}}]\right)ds+\frac{1}{2}S_n(t),$$

where we have used that $u_t^{[0]}\le C_T$ and $S_{n-1}(t)\le S_n(t)$. This implies

$$S_n(t)\le C_T+C_T\int_0^t\left[(1+r_s^{[n]})S_n(s)ds+ne^{-r_s^{[n]}}\right]ds.$$

Let $D_T:=\max(\sup_{k\ge0}\sup_{s\le T}|u_s^{[k]}|,C_T,1)<\infty$, and introduce

$$R_n(t):=\frac{S_n(t)}{(n+1)D_Te^2}\le e^{-2}.$$

Consequently, for all $t \leq T$,

$$R_n(t) \leq \frac{1}{n+1} + C_T \int_0^t \left[(1 + r_s^{[n]}) R_n(s) + e^{-r_s^{[n]}} \right] ds.$$

Finally we choose $r_t^{[n]} := -\ln R_n(t) \geq 2$ and obtain for all $t \leq T$,

$$R_n(t) \leq \frac{1}{n+1} + C_T \int_0^t \left(2 - \ln R_n(s)\right) R_n(s) ds \leq \frac{1}{n+1} - C_T \int_0^t R_n(s) \ln R_n(s) ds.$$

As before we apply Osgood's lemma. Let $M(x) := \int_x^{e^{-2}} \frac{1}{-s \ln s} ds = \ln(-\ln x) - \ln 2$. Then

$$-M(R_n(T)) + M(1/(n+1)) \leq C_T T$$

or equivalently

$$R_n(T) \leq (n+1)^{-e^{-C_T T}} \text{ such that } S_n(T) \leq C_T n^{1-e^{-C_T T}}.$$

Lemma C.4.1 then implies that there exists a subsequence of $(u_T^{[n]})_n$ that converges to 0 as n goes to infinity. This proves (6.7).

Step 4. Let us prove that a subsequence of $(\mu^{[n]})_n$ converges to some limit $\mu : t \mapsto \mu_t$ in the following sense

$$\sup_{0 \leq t \leq T} W_1(\mu_t^{[n]}, \mu_t) \underset{n \to \infty}{\longrightarrow} 0,$$

where $\mu_t := \mathcal{L}(X_t)$ for a.e. $t \leq T$.

We prove this point by proving that the sequence of functions $\mu^{[n]} : t \mapsto \mu_t^{[n]} = \mathcal{L}(X_t^{[n]}) \in \mathcal{P}_1(\mathbb{R})$ is relatively compact, using Arzelà-Ascoli's theorem.

To begin with, the definition of W_1 and the same computation as the one used to obtain (6.6) allows to prove that, for all $s, t \leq T$, for all $n \in \mathbb{N}$,

$$W_1(\mu_t^{[n]}, \mu_s^{[n]}) \leq \mathbb{E}\left[\left| X_t^{[n]} - X_s^{[n]} \right| \right] \leq C \left(|t-s| + \sqrt{|t-s|} \right), \tag{6.8}$$

for a constant $C > 0$ independent of n.

This implies that the sequence $\mu^{[n]} : t \mapsto \mu_t^{[n]}$ is equicontinuous. In addition, by (6.5) we know that, for every $t \leq T$, the set $(\mu_t^{[n]})_n$ is tight, and whence relatively compact (in the topology of the weak convergence, but not in $\mathcal{P}_1(\mathbb{R})$ a priori). Indeed, for any $\varepsilon > 0$, considering $M_\varepsilon := \sup_n \mathbb{E}\left[|X_t^{[n]}| \right] / \varepsilon$, we have, for all n,

$$\mu_t^{[n]}(\mathbb{R} \backslash [-M_\varepsilon, M_\varepsilon]) = \mathbb{P}\left(|X_t^{[n]}| > M_\varepsilon \right) \leq \frac{1}{M_\varepsilon} \mathbb{E}\left[|X_t^{[n]}| \right] \leq \varepsilon.$$

In particular, for every $t \leq T$, we can consider a subsequence of $(\mu_t^{[n]})_n$ that converges weakly. To prove that this convergence holds for the metric W_1, we rely on the characterization (iii) of W_1 given in Definition 6.8, and Theorem 6.9 of Villani (2008). According to this result, the convergence of the same subsequence of $(\mu_t^{[n]})_n$ for W_1 follows from (6.5), Markov's inequality, Cauchy-Schwarz's inequality and the fact that,

$$\mathbb{E}\left[|X_t^{[n]}| \mathbb{1}_{\left\{ |X_t^{[n]}| > R| \right\}} \right] \leq \frac{1}{R} \sup_{k \in \mathbb{N}} \mathbb{E}\left[|X_t^{[k]}| \right] \mathbb{E}\left[(X_t^{[k]})^2 \right]^{1/2} \underset{R \to \infty}{\longrightarrow} 0.$$

We can then conclude that, for all $t \le T$, the sequence $(\mu_t^{[n]})_n$ is also relatively compact on $\mathcal{P}_1(\mathbb{R})$.

Then, thanks to (6.8), Arzelà-Ascoli's theorem implies that the sequence $(\mu^{[n]})_n$ is relatively compact. As a consequence, there exists a subsequence of $(\mu^{[n]})_n$ (as previously, we do not write this subsequence explicitly in the notation) that converges to some $\mu : t \mapsto \mu_t \in \mathcal{P}_1(\mathbb{R})$ in the following sense

$$\sup_{0 \le t \le T} W_1(\mu_t^{[n]}, \mu_t) \underset{n\to\infty}{\longrightarrow} 0.$$

The last thing to show in this step is that $\mu_t = \mathcal{L}(X_t)$ for a.e. $t \le T$. By construction, μ_t is the limit of $\mu_t^{[n]} := \mathcal{L}(X_t^{[n]})$. Recalling that $X^{[n]}$ converges to X in distribution in Skorohod topology, we know that for all continuity point t of $s \mapsto \mathcal{L}(X_s)$, $\mu_t = \mathcal{L}(X_t)$.

Step 5. Recall that, for a subsequence, $(X^{[n]}, X^{[n+1]})_n$ converges to (X, X) in distribution in $D([0,T], \mathbb{R}^2)$, and $(\mu^{[n]})_n$ (which is a sequence of deterministic and continuous functions from $\mathbb{R}_+$ to $\mathcal{P}_1(\mathbb{R})$) converges uniformly to μ on $[0,T]$. The aim of this step is to prove that $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$ converges to (X, X, μ) in distribution in $D([0,T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$. We consider $\mu^{[n]}$ in the previous distribution even though it is deterministic, because the important point in the convergence we want to prove is that $\mu^{[n]}$ must converge w.r.t. the same sequence of time-changes as the one of $(X^{[n]}, X^{[n+1]})$ to be able to apply Lemma C.3.2 almost surely. In particular, it is important to have convergence in the topology of $D([0,T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$ rather than in the weaker topology $D([0,T], \mathbb{R}^2) \times \mathcal{P}_1(\mathbb{R})$.

As μ is continuous, we have that, for any sequence of time-changes $(\lambda_n)_n$ the convergence

$$\sup_{0 \le t \le T} W_1(\mu_t^{[n]}, \mu_{\lambda_n(t)}) \underset{n\to\infty}{\longrightarrow} 0. \tag{6.9}$$

By Skorohod's representation theorem, we can assume that some representative r.v. $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]})$ of $(X^{[n]}, X^{[n+1]})$ converges a.s. to representative r.v. $(\tilde{X}, \tilde{X})$ of (X, X) in $D([0,T], \mathbb{R}^2)$. This implies that, almost surely, there exists a sequence of time-changes $(\lambda_n)_n$ such that

$$\sup_{0 \le t \le T} \left| \tilde{X}_t^{[n]} - \tilde{X}_{\lambda_n(t)} \right| \text{ and } \sup_{0 \le t \le T} \left| \tilde{X}_t^{[n+1]} - \tilde{X}_{\lambda_n(t)} \right|$$

vanish as n goes to infinity. So, by (6.9), almost surely, there exists a sequence of time-changes $(\lambda_n)_n$ such that

$$\sup_{0 \le t \le T} d\left[\left(\tilde{X}_t^{[n]}, \tilde{X}_t^{[n+1]}, \mu^{[n]}]_t \right), \left(\tilde{X}_{\lambda_n(t)}, \tilde{X}_{\lambda_n}(t), \mu_{\lambda_n(t)} \right) \right] \underset{n\to\infty}{\longrightarrow} 0,$$

with $d[(x, y, m), (x', y', m')] = |x - x'| + |y - y'| + W_1(m, m')$. In particular, we know that the sequence $(\tilde{X}^{[n]}, \tilde{X}^{[n+1]}, \mu^{[n]})_n$ converges to $(\tilde{X}, \tilde{X}, \mu)$ almost surely in $D([0,T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$. This implies that $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$ converges to (X, X, μ) in distribution in $D([0,T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$.

Step 6. This step concludes the proof, showing that X is solution to (6.1). In order to prove that X is solution to (6.1), we use the fact that, using the notation of Definitions II.2.6 and II.2.16 of Jacod and Shiryaev (2003), $X^{[n+1]}$ is a semimartingale with characteristics $(B^{[n+1]}, C^{[n+1]}, \nu^{[n+1]})$ given by

$$B_t^{[n+1]} = \int_0^t b(X_s^{[n]}, \mu_s^{[n]}) ds,$$

$$\begin{aligned} C_t^{[n+1]} =& \int_0^t \sigma(X_s^{[n]}, \mu_s^{[n]})^2 ds, \\ \nu^{[n+1]}(dt, dx) =& f(X_t^{[n]}, \mu_t^{[n]}) dt \int_E \delta_{\Phi(X_t^{[n]}, \mu_t^{[n]}, u)}(dx) d\rho(u). \end{aligned}$$

Let us note that, above, we have chosen as truncation function $h = 0$, hence the modified second characteristics $\tilde{C}^{[n+1]}$ is the same as $C^{[n+1]}$.

Recall that, in *Step 5*, we have shown that, for a subsequence, $(X^{[n]}, X^{[n+1]}, \mu^{[n]})_n$ converges in distribution in $D([0,T], \mathbb{R}^2 \times \mathcal{P}_1(\mathbb{R}))$ to (X, X, μ). Using once again Skorohod's representation theorem, we can consider representative r.v. for which the previous convergence is almost sure. Whence, by Lemma C.3.2 and Remark C.3.3, for all $g \in C_b(\mathbb{R})$, the following convergences hold almost surely for the representative r.v. and hence in distribution:

$$\begin{aligned} &\left(X^{[n+1]}, B^{[n+1]}, C^{[n+1]}\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \left(X, \int_0^\cdot b(X_s, \mu_s) ds, \int_0^\cdot \sigma(X_s, \mu_s)^2 ds\right), \\ &\left(X^{[n+1]}, \int_{[0,\cdot] \times \mathbb{R}} g(x) \nu^{[n+1]}(ds, dx)\right) \xrightarrow[n \to +\infty]{\mathcal{L}} \left(X, \int_0^\cdot \int_E g(\Phi(X_s, \mu_s, u)) f(X_s, \mu_s) d\rho(u) ds\right), \end{aligned}$$

where the convergences hold respectively in the spaces $D(\mathbb{R}_+, \mathbb{R}^3)$ and $D(\mathbb{R}_+, \mathbb{R}^2)$.

Then, Theorem IX.2.4 of Jacod and Shiryaev (2003) implies that X is a semimartingale with characteristics (B, C, ν) given by

$$\begin{aligned} B_t =& \int_0^t b(X_s, \mu_s) ds, \\ C_t =& \int_0^t \sigma(X_s, \mu_s)^2 ds, \\ \nu(dt, dx) =& f(X_t, \mu_t) dt \int_E \delta_{\Phi(X_t, \mu_t, u)}(dx) d\rho(u). \end{aligned}$$

Then, we can use the canonical representation of X (see Theorem II.2.34 of Jacod and Shiryaev (2003)): $X = X_0 + B + M^c + Id * \mu^X$, where M^c is a continuous locale martingale, $\mu^X = \sum_s \mathbb{1}_{\{\Delta Y_s \neq 0\}} \delta_{(s, X_s)}$ is the jump measure of X (let us recall that we chose the truncation function $h = 0$) and $(Id * \mu^X)_t := \int_0^t \int_{\mathbb{R}} x d\mu^X(s, x)$. By definition of the characteristics, $\langle M^c \rangle_t = C_t$. Whence, by Theorem II.7.1 of Ikeda and Watanabe (1989), there exists a Brownian motion W such that

$$M_t^c = \int_0^t \sigma(X_s, \mu_s) dW_s. \tag{6.10}$$

In addition, we know that ν is the compensator of μ^X. We rely on Theorem II.7.4 of Ikeda and Watanabe (1989). Using the notation therein, we introduce $Z = \mathbb{R}_+ \times E$, $m(dz, du) = dz \rho(du)$ and

$$\theta(t, z, u) := \Phi(X_{t-}, \mu_{t-}, u) \mathbb{1}_{\{z \leq f(X_{t-}, \mu_{t-})\}}.$$

According to Theorem II.7.4 of Ikeda and Watanabe (1989), there exists a Poisson measure π on $\mathbb{R}_+ \times \mathbb{R}_+ \times E$ having intensity $dt \cdot dz \cdot d\rho(u)$ such that, for all $A \in \mathcal{B}(\mathbb{R})$,

$$\mu^X([0,t] \times A) = \int_0^t \int_0^\infty \int_E \mathbb{1}_{\{\theta(s,z,u) \in A\}} d\pi(s, z, u).$$

This implies that

$$(Id * \mu^X)_t = \int_{[0,t]\times\mathbb{R}_+\times E} \Phi(X_{s-}, \mu_{s-}, u)\mathbb{1}_{\{z\leq f(X_{s-},\mu_{s-})\}} d\pi(s,z,u). \tag{6.11}$$

Finally, recalling that $X = X_0 + B + M^c + Id * \mu^X$, (6.10) and (6.11), we have just shown that X is a weak solution to (6.1) on $[0,T]$. □

6.1.4 Proof of Theorem 6.1.4

In Section 6.1.2, we have proven the (global) pathwise uniqueness of solutions of (6.1), and, in Section 6.1.3, the existence of a weak solution of (6.1) on $[0,T]$, with $T = 1/(16L^2)$.

Then, generalizations of Yamada-Watanabe results allow to construct a strong solution on $[0,T]$: it is a consequence of Theorem 1.5 and Lemma 2.10 of Kurtz (2014) (see the discussion before Lemma 2.10 or Example 2.14 for more details).

More precisely, given a Brownian motion W, a Poisson random measure π and an initial condition X_0, there exists a strong solution $(X_t)_{0\leq t\leq T}$ defined w.r.t. these W, π, X_0. Then, one can construct a strong solution $(X_t)_{T\leq t\leq 2T}$ on $[T, 2T]$ defined w.r.t. the Brownian motion $(W_{T+t} - W_T)_{t\geq 0}$, the Poisson measure π_T defined by

$$\pi_T(A\times B) = \pi(\{T + x : x\in A\}\times B),$$

and the initial condition X_T. Iterating this reasoning, we can construct a strong solution of (6.1) on $[0, kT]$ for any $k\in\mathbb{N}^*$, with $T = 1/(16L^2) > 0$. Hence, there exists a (global) strong solution of (6.1). This proves Theorem 6.1.4.

6.2 Propagation of chaos

In this section, we prove a propagation of chaos for McKean-Vlasov systems: Theorem 6.2.3. This property has been proven in the globally Lipschitz case in Proposition 3.1 of Andreis et al. (2018). Let us introduce the $N-$particle system $(X^{N,i})_{1\leq i\leq N}$

$$\begin{aligned} dX_t^{N,i} =& b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+\times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i)\mathbb{1}_{\left\{z\leq f(X_{t-}^{N,i}, \mu_{t-}^N)\right\}} d\pi^i(t,z,v) \\ &+ \frac{1}{N}\sum_{j=1}^N \int_{\mathbb{R}_+\times F^{\mathbb{N}^*}} \Theta(X_{t-}^{N,j}, X_{t-}^{N,i}, \mu_{t-}^N, v^j, v^i)\mathbb{1}_{\left\{z\leq f(X_{t-}^{N,j}, \mu_{t-}^N)\right\}} d\pi^j(t,z,v), \end{aligned} \tag{6.12}$$

with $\mu^N := N^{-1}\sum_{j=1}^N \delta_{X^{N,j}}$, W^i $(i\geq 1)$ independent standard one-dimensional Brownian motions, and π^i $(i\geq 1)$ independent Poisson measures on $\mathbb{R}_+^2\times F^{\mathbb{N}^*}$ with intensity $dt\cdot dz\cdot d\nu(v)$, where F is a measurable space, and ν is a $\sigma-$finite symmetric measure on $F^{\mathbb{N}^*}$ (i.e. ν is invariant under finite permutations).

In the following, we assume that b, σ and f satisfy the same conditions as in Assumption 6.1, and that Ψ satisfies the same as Φ for some constant $a > 0$, with $E = F^{\mathbb{N}^*}$ and $\rho = \nu$. We also assume that Θ satisfies similar conditions: for all $x_1, x_1', x_2, x_2'\in\mathbb{R}$, $m_1, m_2\in\mathcal{P}_1(\mathbb{R})$,

$$\int_{F^{\mathbb{N}^*}}\int_{\mathbb{R}_+}|\Theta(x_1,x_1',m_1,v^1,v^2)\mathbb{1}_{\{z\leq f(x_1,m_1)\}}-\Theta(x_2,x_2',m_2,v^1,v^2)\mathbb{1}_{\{z\leq f(x_2,m_2)\}}|dzd\nu(v)$$
$$\leq L\left(1+|x_1|+|x_1'|+|x_2|+|x_2'|+\int_{\mathbb{R}}e^{a|x|}dm_1(x)+\int_{\mathbb{R}}e^{a|x|}dm_2(x)\right)$$
$$\left(|x_1-x_2|+|x_1'-x_2'|+W_1(m_1,m_2)\right),$$

and

$$\sup_{x,x'\in\mathbb{R},m\in\mathcal{P}_1(\mathbb{R})}\int_{F^{\mathbb{N}^*}}e^{a|\Theta(x,x',m,v^1,v^2)|}d\nu(v)<\infty.$$

In addition, we assume that each $X_0^{N,i}$ $(i\geq 1,N\geq 1)$ satisfies the initial condition of Assumption 6.1, and that, for every $N\in\mathbb{N}^*$, the system $(X_0^{N,i})_{1\leq i\leq N}$ is i.i.d.

We prove that these $N-$particles systems converge as N goes to infinity to the following limit system.

$$\begin{aligned}d\bar{X}_t^i=&b(\bar{X}_t^i,\bar{\mu}_t)dt+\sigma(\bar{X}_t^i,\bar{\mu}_t)dW_t^i+\int_{\mathbb{R}_+\times F^{\mathbb{N}^*}}\Psi(\bar{X}_{t-}^i,\bar{\mu}_{t-},v^i)\mathbb{1}_{\left\{z\leq f(\bar{X}_{t-}^i,\bar{\mu}_{t-})\right\}}d\pi^i(t,z,v)\\&+\int_{\mathbb{R}}\int_{F^{\mathbb{N}^*}}\Theta(x,\bar{X}_t^i,\bar{\mu}_t,v^1,v^2)f(x,\bar{\mu}_t)d\nu(v)d\bar{\mu}_t(x),\end{aligned}\tag{6.13}$$

where $\bar{\mu}_t=\mathcal{L}(\bar{X}_t)$. We assume that the variables $\bar{X}_0^i$ $(i\geq 1)$ are i.i.d. and satisfy the initial condition of Assumption 6.1.

Let us remark that the (strong) well-posedness of equation (6.13) is a consequence of Theorem 6.1.4 for the same σ, $\Psi=\Phi$ and for the drift

$$b(x,m)+\int_{\mathbb{R}}\int_{F^{\mathbb{N}^*}}\Theta(y,x,m,v^1,v^2)f(y,m)d\nu(v)dm(y).$$

One can also prove the (strong) well-posedness of equation (6.12) using a similar reasoning as the one used in the proof of Theorem 6.1.4. The only difference is for the *Step 4* of the proof of Proposition 6.1.8, since, for (6.12) the measure μ^N is not deterministic. Instead of proving that the sequence of measures $(\mu^{[n]})_n$ constructed in the Picard scheme is relatively compact by Arzelà-Ascoli's theorem, we rely exclusively on the following lemma.

Lemma 6.2.1. *Let $N\in\mathbb{N}^*,T>0$, and $(x^k)_{1\leq k\leq N}$ and $(x_n^k)_{1\leq k\leq N}$ $(n\in\mathbb{N})$ be càdlàg functions. Define*

$$\mu_n(t):=N^{-1}\sum_{k=1}^N\delta_{x_n^k(t)}\text{ and }\mu(t):=\sum_{k=1}^N\delta_{x^k(t)}.$$

Let λ_n $(n\in\mathbb{N})$ be continuous, increasing functions satisfying $\lambda_n(0)=0$, $\lambda_n(T)=T$, and that, for any $1\leq k\leq N$,

$$\sup_{0\leq t\leq T}\left|x_n^k(t)-x^k(\lambda_n(t))\right|\text{ and }\sup_{0\leq t\leq T}|t-\lambda_n(t)|$$

vanish as n goes to infinity. Then,

$$\sup_{0\leq t\leq T}W_1\left(\mu_n(t),\mu(\lambda_n(t))\right)\underset{n\to\infty}{\longrightarrow}0.$$

Proof. For all $t \leq T$,

$$W_1(\mu_n(t), \mu(\lambda_n(t))) \leq \frac{1}{N} \sum_{k=1}^{N} \left| x_n^k(t) - x^k(\lambda_n(t)) \right|,$$

which proves the result. □

Remark 6.2.2. *This lemma allows to prove that, if $(x_n, y_n)_n$ converges to (x, y) in $D([0,T], (\mathbb{R}^N)^2)$, then, the sequence $(x_n, y_n, \mu_n)_n$ converges to (x, y, μ) in $D([0,T], (\mathbb{R}^N)^2 \times \mathcal{P}_1(\mathbb{R}))$.*

In the following, we assume that $(X^{N,i})_{1\leq i\leq N}$ and $(\bar{X}^i)_{i\geq 1}$ are strong solutions of respectively (6.12) and (6.13) defined w.r.t. the same Brownian motions W^i $(i \geq 1)$ and the same Poisson measures π^i $(i \geq 1)$ such that all the systems are defined on the same space. In addition, we assume that the following condition holds true

$$\varepsilon_0^N := \mathbb{E}\left[\left| X_0^{N,1} - \bar{X}_0^1 \right| \right] \underset{N\to\infty}{\longrightarrow} 0. \tag{6.14}$$

Now let us state the main result of this section: the propagation of chaos of the $N-$particle systems, that is, the convergence of the systems $(X^{N,i})_{1\leq i\leq N}$ to the i.i.d. system $(\bar{X}^i)_{i\geq 1}$ as N goes to infinity. We comment the convergence speed in Remark 6.2.5.

Theorem 6.2.3. *Recall (6.14) and the assumptions given at the beginning of Section 6.2. We have, for all $T_0 > 0$,*

$$\mathbb{E}\left[\sup_{0\leq t\leq T_0} \left| X_t^{N,1} - \bar{X}_t^1 \right| \right] \underset{N\to\infty}{\longrightarrow} 0.$$

Consequently, for all $k \geq 1$, the following weak convergence holds true:

$$\mathcal{L}(X^{N,1}, X^{N,2}, ..., X^{N,k}) \underset{N\to\infty}{\longrightarrow} \mathcal{L}(\bar{X}^1) \otimes \mathcal{L}(\bar{X}^2) \otimes ... \otimes \mathcal{L}(\bar{X}^k),$$

in the product topology of the topology of the uniform convergence on every compact set.

Remark 6.2.4. *We just state the result of Theorem 6.2.3 for the first coordinate because both systems $(X^{N,i})_{1\leq i\leq N}$ and $(\bar{X}^i)_{i\geq 1}$ are exchangeable.*

Remark 6.2.5. *In the proof of Theorem 6.2.3, we obtain a convergence speed for*

$$\mathbb{E}\left[\sup_{0\leq t\leq T_0} \left| X_t^{N,1} - \bar{X}_t^1 \right| \right] \underset{N\to\infty}{\longrightarrow} 0$$

that depends on T_0. Indeed, if $T := 1/(16L^2)$, the formula (6.16) below gives a convergence speed for

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left| X_t^{N,1} - \bar{X}_t^1 \right| \right]$$

of the form

$$S_0^N := C^1 \left(\varepsilon_0^N + N^{-1/2} \right)^{C^2},$$

for some positive constants C^1, C^2, *where* ε_0^N *is given at* (6.14)*. And, for all* $k \in \mathbb{N}^*$, *the convergence speed* S_k^N *of*

$$\mathbb{E}\left[\sup_{kT \le t \le (k+1)T} \left|X_t^{N,1} - \bar{X}_t^1\right|\right]$$

can be obtained inductively by

$$S_k^N = C^1\left(S_{k-1}^N + N^{-1/2}\right)^{C^2}$$

for the same constants C^1, C^2 *for all* k.

Before proving Theorem 6.2.3, let us state a lemma about some a priori estimates of the process $(X_t^{N,1})_{t\ge 0}$.

Lemma 6.2.6. *For every* $N \in \mathbb{N}^*$, *let* $(X^{N,i})_{1\le i\le N}$ *be the solution of* (6.12)*. For any* $t \ge 0$,

(i) $\sup_{N\in\mathbb{N}^*}\sup_{0\le s\le t} \mathbb{E}\left[e^{a|X_s^{N,1}|}\right] < \infty$ *and* $\sup_{0\le s\le t} \mathbb{E}\left[e^{a|\bar{X}_s^1|}\right] < \infty$,

(ii) $\sup_{N\in\mathbb{N}^*} \mathbb{E}\left[\sup_{0\le s\le t} |X_s^{N,1}|\right] < \infty$ *and* $\mathbb{E}\left[\sup_{0\le s\le t} |\bar{X}_s^1|\right] < \infty$.

Sketch of proof of Lemma 6.2.6. We only give the main steps of the proof of the first part of the point (i) since the main part of the proof relies on the same computations as in the proof of Lemma 6.1.5. The only new property of the lemma is the fact that the bounds has to be uniform in N. This property is easy to prove for the point (ii) of the lemma in this model because of the normalization in N^{-1} of the jump term Θ of (6.12). For the point (i), as in the proof of Lemma 6.1.5, we apply Ito's formula to both functions $x \mapsto e^{ax}$ and $x \mapsto e^{-ax}$, and then, one has to note that, by exchangeability,

$$\begin{aligned}
\mathbb{E}\left[e^{aX_t^{N,1}}\right] \le& \mathbb{E}\left[e^{X_0^{N,1}}\right] + a||b||_\infty \int_0^t \mathbb{E}\left[e^{aX_s^{N,1}}\right] ds + \frac{1}{2}a^2||\sigma||_\infty^2 \int_0^t \mathbb{E}\left[e^{aX_s^{N,1}}\right] ds \\
&+ ||f||_\infty \int_0^t \mathbb{E}\left[e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} \left(e^{a\Psi(X_s^{N,1},\mu_s^N,v^1)} - 1\right) d\nu(v)\right] ds \\
&+ (N-1)||f||_\infty \int_0^t \mathbb{E}\left[e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} \left(e^{aN^{-1}\Theta(X_s^{N,2},X_s^{N,1},\mu_s^N,v^1,v^2)} - 1\right) d\nu(v)\right] ds.
\end{aligned}$$

So, the only term that does depend on N is the one of the last line that is bounded by

$$\begin{aligned}
||f||_\infty \int_0^t \mathbb{E}&\left[e^{aX_s^{N,1}} \int_{F^{\mathbb{N}^*}} N\left(e^{aN^{-1}|\Theta(X_s^{N,2},X_s^{N,1},\mu_s^N,v^1,v^2)|} - 1\right) d\nu(v)\right] ds \\
&\le ||f||_\infty \int_0^t \mathbb{E}\left[e^{aX_s^{N,1}}\right] ds \sup_{x,x'\in\mathbb{R}, m\in\mathcal{P}_1(\mathbb{R})} \int_{F^{\mathbb{N}^*}} e^{a|\Theta(x,x',m,v^1,v^2)|} d\nu(v),
\end{aligned}$$

where we have used the fact that, for all $x \ge 0, n \in \mathbb{N}^*$, $n(e^{x/n} - 1) \le e^x$.

Finally, Grönwall's lemma allows to obtain a bound of $\mathbb{E}\left[e^{aX_s^{N,1}}\right]$ that is uniform in N. With the same reasoning, we can obtain the same bound for $\mathbb{E}\left[e^{-aX_s^{N,1}}\right]$. This allows to have a bound for $\mathbb{E}\left[e^{a|X_s^{N,1}|}\right]$ uniform in N. □

Remark 6.2.7. *For the previous proof to be formal, we should have introduce stopping times as in the proof of Lemma 6.1.5, do the computations above for the stopped processes, and then apply Fatou's lemma to prove that the control obtained for the stopped processes still hold for the real process.*

Proof of Theorem 6.2.3. We follow the ideas of Proposition 3.1 of Andreis et al. (2018). Instead of introducing an auxiliary system, we rewrite artificially the equation (6.12) as

$$\begin{aligned} dX_t^{N,i} =& b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)dW_t^i + \int_{\mathbb{R}_+ \times F^{\mathbb{N}^*}} \Psi(X_{t-}^{N,i}, \mu_{t-}^N, v^i)\mathbb{1}_{\left\{z \le f(X_{t-}^{N,i}, \mu_{t-}^N)\right\}} d\pi^i(t,z,v) \\ &+ \int_{\mathbb{R}} \int_{F^{\mathbb{N}^*}} \Theta(x, X_t^{N,i}, \mu_t^N, v^1, v^2) f(x, \mu_t^N) d\nu(v) d\mu_t^N(x) + dG_t^N, \end{aligned}$$

where

$$\begin{aligned} G_t^N = \frac{1}{N}\sum_{j=1}^N \Bigg[& \int_{[0,t]\times\mathbb{R}_+\times F^{\mathbb{N}^*}} \Theta(X_{s-}^{N,j}, X_{s-}^{N,i}, \mu_{s-}^N, v^j, v^i)\mathbb{1}_{\left\{z\le f(X_{s-}^{N,j},\mu_{s-}^N)\right\}} d\pi^j(s,z,v) \\ & - \int_0^t \int_{F^{\mathbb{N}^*}} \Theta(X_s^{N,j}, X_s^{N,i}, \mu_s^N, v^1, v^2) f(X_s^{N,j}, \mu_s^N) d\nu(v) ds \Bigg]. \end{aligned}$$

Let us define

$$u^N(t) := \mathbb{E}\left[\sup_{0\le s\le t} \left|X_s^{N,1} - \bar{X}_s^1\right|\right].$$

Let us note $\bar{\mu}_t^N := N^{-1}\sum_{j=1}^N \delta_{\bar{X}^j}$. By triangle inequality and since for all $m, m' \in \mathcal{P}_1(\mathbb{R})$, $W_1(m,m') \le W_2(m,m')$, we have for all $t \ge 0$,

$$W_1(\mu_t^N, \bar{\mu}_t) \le W_1(\mu_t^N, \bar{\mu}_t^N) + W_1(\bar{\mu}_t^N, \bar{\mu}_t) \le \frac{1}{N}\sum_{j=1}^N \left|X_t^{N,j} - \bar{X}_t^j\right| + W_2(\bar{\mu}_t^N, \bar{\mu}_t).$$

Besides, using Theorem 1 of Fournier and Guillin (2015) with $d = 1$, $p = 2$ and any $q > 4$ (and using Lemma 6.2.6), we have

$$\mathbb{E}\left[W_2(\bar{\mu}_t^N, \bar{\mu}_t)\right] \le C\mathbb{E}\left[|\bar{X}_t|^q\right]^{p/q} N^{-1/2} \le C\left(1 + \mathbb{E}\left[e^{a|\bar{X}_t|}\right]\right)^{p/q} N^{-1/2} \le C_t N^{-1/2}.$$

As a consequence, and thanks to Lemma 6.2.6, for all $t \ge 0$,

$$\mathbb{E}\left[\sup_{0\le s\le t} W_1(\mu_s^N, \mu_s)\right] \le u^N(t) + C_t N^{-1/2}.$$

Whence the same truncation arguments as the ones used in Sections 6.1.2 and 6.1.3 allows to prove that, for all $t \le T$,

$$u^N(t) \le u^N(0) + C_T \int_0^t \left[(1 + r_s^N)u^N(s) + e^{-r_s^N}\right] ds + 2L\sqrt{T}u^N(t) + \mathbb{E}\left[\sup_{0\le s\le T} |G_s^N|\right] + C_T N^{-1/2},$$

where C_T does not depend on N thanks to Lemma 6.2.6.

Now, fixing $T = 1/(16L^2)$ (such that $2L\sqrt{T} \le 1/2$), we obtain

$$u^N(t) \le 2u^N(0) + C_T \int_0^t \left[(1 + r_s^N)u^N(s) + e^{-r_s^N}\right] ds + 2\mathbb{E}\left[\sup_{0\le s\le T} |G_s^N|\right] + C_t N^{-1/2}. \qquad (6.15)$$

To control the term G_s^N, we use Burkhölder-Davis-Gundy's inequality,

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|G_t^N\right|^2\right] \le \frac{C}{N^2}\sum_{j=1}^N \mathbb{E}\left[\int_0^t \int_{F^{\mathbb{N}^*}} \Theta(X_s^{N,j}, X_s^{N,i}, \mu_s^N, v^j, v^i)^2 f(X_s^{N,j}, \mu_s^N) d\nu(v) ds\right] \le \frac{C_T}{N},$$

where we have used the boundedness conditions of the functions f and Θ, and the fact that the Poisson measures π^j $(1 \le j \le N)$ are independent.

Hence, by Cauchy-Schwarz's inequality,

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|G_t^N\right|\right] \le C_T N^{-1/2}.$$

Then we can rewrite (6.15) as

$$u^N(t) \le 2u^N(0) + C_T \int_0^t \left[(1 + r_s^N)u^N(s) + e^{-r_s^N}\right] ds + C_T N^{-1/2}.$$

Now, let $D_T := \max(C_T, \sup_N \sup_{s\le T} u^N(s), 1)$ which is finite by Lemma 6.2.6.(ii), and define

$$v^N(t) := \frac{u^N(t)}{D_T e^2} \le e^{-2}.$$

Choosing $r_t^N := -\ln v^N(t)$, we have for all $0 \le t \le T$,

$$\begin{aligned} v^N(t) \le& 2v^N(0) + C_T \int_0^t (2 - \ln v^N(s)) v^N(s) ds + \frac{C_T}{\sqrt{N}} \\ \le& 2v^N(0) - 2C_T \int_0^t v^N(s) \ln v^N(s) ds + \frac{C_T}{\sqrt{N}}, \end{aligned}$$

and Osgood's lemma allows to conclude that

$$-M(v(T)) + M(2v^N(0) + C_T N^{-1/2}) \le 2C_T T,$$

where $M(x) = \int_x^{e^{-2}} \frac{1}{-s\ln s} ds = \ln(-\ln x) - \ln 2$. This implies that

$$\ln(-\ln(2v^N(0) + C_T^1 N^{-1/2})) - \ln(-\ln v^N(t)) \le C_T^2,$$

for some constants $C_T^1, C_T^2 > 0$, where we distinguish C_T^1 and C_T^2 for clarity. This implies that, for all $0 \le t \le T$,

$$v^N(t) \le \left(2v^N(0) + C_T^1 N^{-1/2}\right)^{\exp(-C_T^2)}. \qquad (6.16)$$

Hence

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|X_t^{N,1} - \bar{X}_t^1\right|\right] \underset{N\to\infty}{\longrightarrow} 0,$$

for a $T > 0$ sufficiently small that does not depend on the initial conditions (recalling that we have taken $T = 1/(16L^2)$). Then, iterating this reasoning on $[T, 2T]$, we can prove

$$\mathbb{E}\left[\sup_{T\leq t\leq 2T}\left|X_t^{N,1} - \bar{X}_t^1\right|\right] \underset{N\to\infty}{\longrightarrow} 0,$$

noticing that the "initial conditions" on $[T, 2T]$ satisfy the same condition as (6.14):

$$\varepsilon_T^N := \mathbb{E}\left[\left|X_T^{N,1} - \bar{X}_T^1\right|\right] \underset{N\to\infty}{\longrightarrow} 0.$$

Finally, by induction, we can prove that for all $k \in \mathbb{N}^*$,

$$\mathbb{E}\left[\sup_{0\leq t\leq kT}\left|X_t^{N,1} - \bar{X}_t^1\right|\right] \underset{N\to\infty}{\longrightarrow} 0,$$

which proves the result. □

Chapitre 7

Conclusion (en français)

7.1 Bilan

Dans cette thèse nous avons étudié une question nouvelle : des limites de grandes échelles en régime diffusif (i.e. quand la force d'interaction d'un système à N particules est de l'ordre de $N^{-1/2}$) pour des systèmes de processus de Hawkes et des systèmes McKean-Vlasov. En effet, dans la littérature, les questions de limites de grande échelle ont surtout été traitées en régime linéaire (avec des interactions en N^{-1}), mais peu de travaux existent pour le régime diffusif.

Notons aussi que, bien que les modèles étudiés semblaient similaires, ils nécessitaient des techniques de preuves très différentes. Passons brièvement en revue ces techniques, qu'on peut catégoriser en trois parties : les techniques du cadre markovien des Chapitres 2 et 3, les problèmes martingales dans les Chapitres 4 et 5, et enfin des techniques analytiques plus directes dans le Chapitre 6. Les deux premières techniques ont un point commun : la notion de générateur.

7.1.1 Techniques markoviennes

Dans la première partie de cette thèse, nous avons étudié des limites de grande échelle de processus de Hawkes. Le premier modèle (Chapitre 2) était classique, car nos systèmes à $N-$particules étaient caractérisés par un processus X^N de dimension un, solution de l'équation différentielle stochastique suivante :

$$dX_t^N = -\alpha X_t^N dt + \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbb{1}_{\{z \le f(X_{t-}^N)\}} d\pi^j(t,z,u).$$

Il est important de remarquer que nous avons pu nous ramener à une telle équation parce que nous avons considéré des noyaux de convolutions exponentiels. En général, l'intensité des processus de Hawkes n'est ni un processus de Markov, ni une semi-martingale.

Cette structure particulière nous a permis d'utiliser des outils puissants, notamment ceux liés aux semi-groupes et aux générateurs infinitésimaux. Notons A^N (resp. P^N) le générateur (resp. le semi-groupe) de X^N, et $\bar{A}$ (resp. $\bar{P}$) celui de sa limite $\bar{X}$. La convergence en loi de X^N était une conséquence plus ou moins directe de celle de son semi-groupe, qui était elle-même une conséquence de la convergence de son générateur via la formule suivante

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds. \tag{7.1}$$

7.1.2 Lien entre les Chapitres 2 et 4

Nous avons déjà évoqué le fait que l'approche du Chapitre 2 ne fonctionne plus dans le cadre du Chapitre 4. Dans cette section, nous en expliquons la raison.

Tout d'abord, il n'est pas clair que les processus limites $\bar{X}^i$ du modèle du Chapitre 4 sont markoviens. Rappelons que chaque $\bar{X}^i$ est solution de

$$\begin{aligned} d\bar{X}_t^i =& -\alpha \bar{X}_t^i dt - \int_{\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{t-}^i \mathbf{1}_{\{z \le f(\bar{X}_{t-}^i)\}} \pi^i(dt, dz, du) \\ &+ \sigma \sqrt{\mathbb{E}\left[f\left(\bar{X}_t^i\right) \middle| \mathcal{W}\right]} dW_t. \end{aligned}$$

Rappelons aussi que nous avons introduit un système auxiliaire $(\tilde{X}^{N,i})_{1\le i\le N}$ qui lui est markovien, et que nous avons montré la convergence de ce système auxiliaire vers le système limite dans un sens L^1 avec une vitesse de convergence en $N^{-1/2}$. Donc, nous pourrions espérer obtenir une convergence avec une vitesse entre les générateurs des systèmes initiaux $(X^{N,i})_{1\le i\le N}$ et ceux des systèmes auxiliaires $(\tilde{X}^{N,i})_{1\le i\le N}$. Le générateur du système X^N s'écrit

$$A^N g(x) = -\alpha \sum_{i=1}^N x^i \partial_i g(x) + \sum_{i=1}^N f(x^i) \int_{\mathbb{R}} \left[g\left(x - x^i e_i + \frac{u}{\sqrt{N}} \sum_{j\neq i} e^j \right) - g(x) \right] d\nu(u),$$

avec e_i le vecteur unité de la $i-$ième coordonnée. Dans la suite de cette section, pour simplifier, nous supposerons que la somme ci-dessus porte sur tous les $1 \le j \le N$ (i.e. sans exclure i). Et le générateur de $\tilde{X}^N$ s'écrit

$$\tilde{A}^N g(x) = -\alpha \sum_{i=1}^N x^i \partial_i g(x) + \sum_{i=1}^N f(x^i) \left[g\left(x - x^i e_i\right) - g(x) \right] + \frac{\sigma^2}{2N} \sum_{i=1}^N f(x_i) \sum_{j,k=1}^N \partial_{j,k}^2 g(x).$$

En rappelant que ν est une loi centrée, nous pourrions écrire

$$\begin{aligned} \left| A^N g(x) - \tilde{A}^N g(x) \right| \le \sum_{i=1}^N f(x^i) \int_{\mathbb{R}} \Bigg[& g\left(x - x^i e_i + \frac{u}{\sqrt{N}} \sum_{j\neq i} e^j \right) - g(x - x^i e_i) \\ & - \frac{u}{\sqrt{N}} \sum_{j=1}^N \partial_j g(x - x^i e_i) - \frac{u^2}{2N} \sum_{k,j=1}^N \partial_{j,k}^2 g(x) \Bigg] d\nu(u). \end{aligned}$$

Le problème, c'est que pour appliquer l'inégalité de Taylor-Lagrange, il faudrait que le terme $\partial_{j,k}^2 g(x)$ ci-dessus soit remplacé par $\partial_{j,k}^2 g(x - x^i e_i)$. Et même si c'était le cas, cela signifierait que ce terme devrait apparaître dans l'expression des générateurs. Et il ne semble pas clair ce que ce terme $\partial_{j,k}^2 g(x - x^i e_i)$ représenterait en pratique pour les processus concernés.

Remarquons que, si le terme de saut de réinitialisation n'était pas présent dans ce modèle, alors la différence des générateurs s'écrirait

$$\left|A^N g(x) - \tilde{A}^N g(x)\right| \leq \sum_{i=1}^{N} f(x^i) \int_{\mathbb{R}} \left[g\left(x + \frac{u}{\sqrt{N}} \sum_{j\neq i} e^j\right) - g(x) \right.$$
$$\left. - \frac{u}{\sqrt{N}} \sum_{j=1}^{N} \partial_j g(x) - \frac{u^2}{2N} \sum_{k,j=1}^{N} \partial^2_{j,k} g(x) \right] d\nu(u).$$

L'inégalité de Taylor-Lagrange donnerait alors que, pour tout $k \geq 1$, pour toute fonction test $g \in C^3_b(\mathbb{R}^k)$ (que l'on peut interpréter artificiellement comme $g \in C^3_b(\mathbb{R}^N)$ pour $N \geq k$),

$$\left|A^N g(x) - \bar{A} g(x)\right| \leq C k^3 N^{-1/2}.$$

Nous pourrions ensuite essayer d'utiliser la formule (7.1) pour obtenir une vitesse de convergence des semi-groupes pour de tels fonctions tests $g \in C^3_b(\mathbb{R}^k)$ à k fixé quand N tend vers l'infini. Mais nous serions confrontés au problème suivant : même si la fonction test g n'est défini que sur $\mathbb{R}^k$, la fonction $x \mapsto \tilde{P}^N_s g(x)$ sera défini sur $\mathbb{R}^N$, et donc l'expression que l'on obtiendrait exploserait quand N tendrait vers l'infini.

Terminons cette section en soulignant le fait que, bien que le terme de saut de réinitialisation ne soit pas présent dans le modèle du Chapitre 5, il y aurait une difficulté en plus dans ce cadre. En effet, pour utiliser la formule (7.1), il est important de pouvoir contrôler la régularité du flot stochastique du système auxiliaire. Dans le modèle du Chapitre 5, nous manipulons des systèmes McKean-Vlasov généraux, et donc les coefficients des équations différentielles stochastiques dépendent des mesures empiriques des systèmes. Cette dépendance pourrait engendrer des difficultés techniques dans le contrôle du flot stochastique.

7.1.3 Problème martingale

Dans les Chapitres 4 et 5, les techniques de preuves précédentes ne s'appliquaient plus. Toutefois les arguments principaux des preuves de ces chapitres reposaient aussi sur la notion de générateur, via des problèmes martingales. Les problèmes martingales que nous avons étudiés sont d'un nouveau type.

En effet, habituellement, on dit que la loi d'un processus X est solution d'un problème martingale caractérisé par un générateur A (A est donc un opérateur sur un espace de fonctions) si pour toute fonction g suffisament régulière, le processus

$$g(X_t) - g(X_0) - \int_0^t A g(X_s) ds \tag{7.2}$$

est une martingale locale. Remarquons que dans les cadres classiques, il existe des résultats généraux permettant d'établir l'équivalence entre une équation différentielle stochastique et son problème martingale associé (équivalence dans le sens où l'équation et le problème martingale ont exactement les mêmes solutions), et qu'il est possible de reconstruire l'équation associée à un problème martingale en examinant la forme du générateur A.

Dans nos modèles il y avait plusieurs difficultés :

- La première était que notre cadre n'était pas classique à cause de la mesure aléatoire présente dans les équations différentielles stochastiques. Donc il a fallu démontrer l'équivalence entre l'équation et le problème martingale utilisé.

- Une autre difficulté venait du fait que nous ne manipulions une équation, mais un système d'équations dont les solutions n'étaient pas indépendantes, mais seulement condionnellement indépendantes.

- Dans la formule (7.2), la générateur A dépend aussi d'une mesure aléatoire en plus du processus.

- La dernière difficulté venait du fait que nous nous intéressions à la convergence de mesures aléatoires, et non de processus.

C'est pour ces raisons que nous avons dû formuler le problème martingale dont la solution est la loi d'une mesure aléatoire et non celle d'un processus. Pour cela nous avons considéré un espace canonique où la variable canonique était un triplet composé de la mesure aléatoire limite et deux coordonnées distinctes du système d'équations limite. La loi du triplet a été choisie de telle sorte que les deux coordonnées était un mélange d'i.i.d. dirigé par la mesure aléatoire. En particulier cette loi était caractérisée par la loi de la mesure aléatoire, et donc le problème martingale aussi. Notons finalement que nous avions besoin de considérer deux coordonnées pour retrouver les corrélations du système d'équations, et donc pour identifier le bruit commun du système. Une solution classique pour gérer le bruit commun dans les jeux en champ moyen consiste à ajouter le bruit commun dans le problème martingale. Mais cette solution ne fonctionnerait pas dans notre cadre car le bruit commun n'est pas présent dans les systèmes initiaux, mais seulement dans le système limite.

Notons que dans le Chapitre 5, il y avait une autre difficulté technique importante. Contrairement au Chapitre 4, où la convergence faible de mesures $(\mu_t^N)_N$ implique clairement la convergence de $(\mu_t^N(f))_N$, cette même convergence n'implique pas en général la convergence de $(\mu_t^N)_N$ au sens de la distance W_1. C'est pourquoi nous avons eu besoin du Lemme 5.4.7.

7.1.4 Techniques analytiques

Finalement, dans le Chapitre 6, nous avons travaillé en normalisation linéaire sur des équations McKean-Vlasov avec des sauts. Contrairement aux cadres classiques étudiés dans la littérature, nous avons travaillé avec des coefficients qui n'étaient pas globalement lipschitziens. Plus précisément, nos coefficients étaient localement lipschitziens, et leurs constantes de Lipschitz étaient sous-linéaires par rapport à la variable d'espace et sous-exponentielles par rapport à la variable de mesure. Nous avons démontré que ces équations étaient bien posées et une propriété de propagation du chaos.

Quand on travaille avec des coefficients lipschitziens, le lemme de Grönwall est un argument essentiel pour prouver ces résultats. Dans notre cadre, nous ne pouvions pas appliquer ce lemme, donc nous avons utilisé une de ces généralisation : le lemme d'Osgood. Les techniques classiques utilisées conjointement avec ce lemme nous ont permis de démontrer la propagation du chaos de nos systèmes ainsi que l'unicité trajectorielle de nos équations.

Toutefois, nous avons rencontré des difficultés supplémentaires pour démontrer l'existence de solutions. Pour construire une solution de nos équations, nous sommes partis d'une technique classique : un schéma d'itération à la Banach-Picard. Mais, le fait que nous considérions des coefficients seulement localement lipschitziens nous a empêché de montrer directement la convergence de ce schéma dans un sens L^1. Dans le cadre classique où les coefficients sont globalement lipschitziens, on montre que la suite construite dans le schéma d'itération est la somme d'une série téléscopique convergeante de la forme :

$$X_t := X_0 + \sum_{n=0}^{+\infty} \mathbb{E}\left[|X_t^{[n+1]} - X_t^{[n]}| \right],$$

avec $X^{[n]}$ le processus construit au $n-$ième pas d'itération du schéma. Cette convergence est vraie à la fois p.s. et dans L^1.

Dans notre cadre, nous avons dû démontrer deux propriétés pour arriver à une conclusion plus faible (car la série n'était a priori pas convergente) :

- Nous n'avons pas montré la convergence de la suite $X^{[n]}$, mais uniquement la convergence (en loi) d'une sous-suite par un argument de tension.
- Pour montrer que la limite de cette sous-suite était bien solution, nous avons démontré que si la suite $(X^{[\varphi(n)]})_n$ convergeait vers un processus X (où φ est une extractrice), alors la suite $(X^{[\varphi(n)]}, X^{[\varphi(n)+1]})_n$ convergeait aussi vers (X, X). C'est pour cette raison que nous avons introduis un lemme général sur les séries.

Ces deux propriétés nous ont permis de construire une solution de notre équation McKean-Vlasov. Mais, comme la solution construite n'était qu'une limite en loi du schéma d'itération, ce n'était qu'une solution au sens faible. Il a donc fallu utiliser une généralisation des résultats de Yamada et Watanabe pour obtenir une solution forte.

7.2 Perspectives

Dans cette section, nous discutons de perspectives de recherche sur des sujets liés aux résultats de cette thèse.

7.2.1 Processus de Hawkes avec avec un noyau général

Une question naturelle que l'on pourrait se poser est : est-il possible de généraliser les résultats sur les processus de Hawkes des Chapitres 2 et 3 aux cas où le noyau de convolution n'est ni sous forme exponentielle ni sous forme Erlang.

L'intérêt des noyaux exponentiels réside dans le fait que l'intensité du processus de Hawkes associé satisfait une équation différentielle stochastique classique d'Ito (pour les noyaux d'Erlang, l'intensité est une coordonnée d'une solution d'une équation multi-dimensionnelle). En particulier, une telle intensité est à la fois un processus de Markov et une semi-martingale, ce qui n'est pas le cas quand on considère un noyau de convolution général.

Rappelons brièvement le modèle du Chapitre 2. Ce chapitre était consacré à l'étude des systèmes de processus de Hawkes $(Z^{N,i})_{1\leq i\leq N}$ $(N \in \mathbb{N}^*)$ où, pour chaque $N \in \mathbb{N}^*$, tous les processus ponctuels $Z^{N,i}$ $(1 \leq i \leq N)$ avaient la même intensité

$$\lambda_t^N := f\left(X_{t-}^N\right),$$

avec

$$X_t^N = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u e^{-\alpha(t-s)} \mathbb{1}_{\left\{z\leq f(X_{s-}^N)\right\}} d\pi^j(s,z,u).$$

Rappelons que la convergence des systèmes de processus de Hawkes était une conséquence de la convergence des processus X^N. Pour démontrer cette convergence, nous avions exploité une formule de type Trotter-Kato pour nous ramener à la convergence des générateurs infinitésimaux des processus X^N.

Si le processus X^N n'est pas un processus de Markov, le schéma de preuve utilisé au Chapitre 2 ne peut plus être utilisé. Nous pouvons aussi remarquer que le Théorème $IX.4.15$ de Jacod and Shiryaev (2003) ne permet plus de montrer que la convergence de X^N implique celle des processus ponctuels $Z^{N,1}$, puisque le processus $(X^N, Z^{N,1})$ n'est plus une semi-martingale. Mais notre Théorème B.0.1 peut encore s'appliquer dans ce cadre. D'après ce théorème, pour montrer la convergence en loi sur l'espace de Skorokhod de processus ponctuels, il suffit de montrer la convergence de leurs intensités stochastiques en loi sur l'espace de Skorokhod, sous la condition que l'intensité limite est indépendante de la mesure de Poisson amincie.

Il est donc naturel de s'intéresser à la question suivante : étant donné une fonction générique h, est-ce que le processus $X^{N,h}$ défini par

$$X_t^{N,h} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} uh(t-s)\mathbb{1}_{\left\{z\leq f(X_{s-}^{N,h})\right\}} d\pi^j(s,z,u) \tag{7.3}$$

converge vers le processus limite $\bar{X}^h$ solution de

$$\bar{X}_t^h = \sigma \int_0^t h(t-s)\sqrt{f(\bar{X}_s^h)}dW_s, \tag{7.4}$$

où les π^j $(j \geq 1)$ sont des mesures de Poisson indépendantes sur $\mathbb{R}_+^2 \times \mathbb{R}$ d'intensité $dt \cdot dz \cdot d\mu(u)$ avec μ une loi centrée sur $\mathbb{R}$ de variance σ^2 et W un mouvement brownien standard de dimension un. Les processus $\bar{X}^h$ de ce type (i.e. les solutions d'équations différentielles stochastiques de convolution avec un terme brownien) s'appellent des processus de Volterra.

La question de la convergence de $X^{N,h}$ vers $\bar{X}^h$ est délicate dans le cas général. À notre connaissance, le résultat qui s'approche le plus de cette convergence est le Théorème 7.2 de Abi Jaber et al. (2019) qui implique entre autre que, si la fonction h est localement L^p et localement lipschitzienne, alors la suite $(X^N)_N$ est tendue (dans la topologie L^p_{loc}) et n'importe quelle limite (au sens L^p_{loc}) de sous-suite de $X^{N,h}$ est solution de (7.4). Remarquons que la seule chose qui manque pour conclure sur la convergence de $X^{N,h}$ est l'unicité des solutions de (7.4).

Dans Abi Jaber et al. (2019), les auteurs travaillent avec la topologie L^p_{loc} plutôt qu'avec la topologie de Skorokhod, car ils considèrent des noyaux de convolution h qui explosent en zéro. Avec de tels noyaux, les processus $X^{N,h}$ et $\bar{X}^h$ ne sont pas càdlàg. Ceci est une difficulté supplémentaire.

Nous pourrions envisager un cadre plus simple avec une fonction de convolution générale mais qui n'explose pas. Dans ce cadre, nous pourrions chercher à démontrer une vraie convergence et donner une vitesse de convergence. Pour cela une idée pourrait être de considérer une fonction h continue, et de l'approcher par une suite de fonctions polynomiales h_n. L'intérêt de se ramener à des fonctions polynomiales, c'est que nous avons déjà démontré que, pour chaque $n \in \mathbb{N}$, la suite $(X^{N,h_n})_N$ converge vers $\bar{X}^{h_n}$ (voir Théorème 3.2.2) et même obtenu une vitesse de convergence explicite (voir Théorème 3.2.3). Nous pourrions ensuite trouver une vitesse de convergence (et donc prouver la convergence) de $X^{N,h}$ vers $\bar{X}^h$ en écrivant pour n'importe quelle fonction test g suffisament régulière

$$\left|\mathbb{E}\left[g\left(\bar{X}_t^h\right)\right] - \mathbb{E}\left[g\left(X_t^{N,h}\right)\right]\right| \leq \left|\mathbb{E}\left[g\left(\bar{X}_t^h\right)\right] - \mathbb{E}\left[g\left(\bar{X}_t^{h_n}\right)\right]\right| \tag{7.5}$$

$$+\left|\mathbb{E}\left[g\left(\bar{X}_t^{h_n}\right)\right]-\mathbb{E}\left[g\left(X_t^{N,h_n}\right)\right]\right| \tag{7.6}$$
$$+\left|\mathbb{E}\left[g\left(X_t^{N,h_n}\right)\right]-\mathbb{E}\left[g\left(X_t^{N,h}\right)\right]\right|. \tag{7.7}$$

En supposant que la suite de fonctions $(h_n)_n$ converge vers h au sens d'une certaine distance (par exemple, si h est continue, $(h_n)_n$ est une suite de polynome, et si nous nous intéressons à la convergence sur $[0,T]$, nous pouvons prendre la norme infinie), nous savons que le terme (7.5) converge vers zéro quand n tend vers l'infini (indépendamment de N) : si la fonction g est lipschitzienne, il suffit de contrôler (7.5) par

$$\left|\mathbb{E}\left[g\left(\bar{X}_t^{h}\right)\right]-\mathbb{E}\left[g\left(\bar{X}_t^{h_n}\right)\right]\right| \leq C\mathbb{E}\left[\left|X_t^h-X_t^{h_n}\right|\right] \leq C_t d(h,h_n).$$

Le terme (7.7) est plus dur à contrôler, car il dépend aussi de N. Nous pouvons facilement obtenir le contrôle suivant (par le lemme de Grönwall)

$$\left|\mathbb{E}\left[g\left(X_t^{N,h_n}\right)\right]-\mathbb{E}\left[g\left(X_t^{N,h}\right)\right]\right| \leq C_t N^{1/2}e^{N^{1/2}}d(h,h_n).$$

Finalement, le Théorème 3.2.3 nous permet de contrôler le terme (7.6) :

$$\left|\mathbb{E}\left[g\left(\bar{X}_t^{h_n}\right)\right]-\mathbb{E}\left[g\left(X_t^{N,h_n}\right)\right]\right| \leq C_t M(h_n)N^{-1/2},$$

où $M(h_n)$ est une constante qui dépend de h_n. Nous pouvons alors "contrôler" la convergence de $X^{N,h}$ vers $\bar{X}^h$ de la manière suivante

$$\left|\mathbb{E}\left[g\left(\bar{X}_t^{h}\right)\right]-\mathbb{E}\left[g\left(X_t^{N,h}\right)\right]\right| \leq C_t\left[d(h,h_n)\left(1+N^{1/2}e^{N^{1/2}}\right)+M(h_n)N^{-1/2}\right].$$

Dans l'expression ci-dessus, il y a deux paramètres n et N. Le paramètre N est un paramètre du modèle, et le n est un paramètre que l'on peut contrôler. Nous pouvons remarquer que chaque terme de la somme dans les crochets ci-dessus tend vers zéro quand l'un des deux paramètres tend vers l'infini (et que l'autre est fixé), mais explose quand l'autre paramètre tend vers l'infini. Donc pour obtenir la convergence de $X^{N,h}$ vers $\bar{X}^h$ il faut trouver un compromis entre la vitesse de convergence de ces deux paramètres (i.e. choisir $n=v(N)$ pour une fonction v bien choisie) pour que l'expression au-dessus tende vers zéro. Cela donnerait même une vitesse de convergence.

Toutefois, il y a plusieurs problèmes qui empêchent de conclure :

- Il faut calculer la constante $M(h_n)$ explicitement. Cette constante fera naturellement apparaître la norme infinie locale des dérivées de h_n. Il faudra donc réussir à garantir un contrôle de ces dérivées, en choisissant un schéma d'approximation fonctionnelle convenable.

- Dans le contrôle de la convergence de $X^{N,h}$ vers $\bar{X}^h$, le terme en $e^{\sqrt{N}}$ est problématique. En effet, ce terme va forcer le paramètre n à tendre vers l'infini beaucoup plus vite que le paramètre N pour compenser cette exponentielle. Ce qui peut être problématique pour trouver un compromis entre les deux paramètres.

7.2.2 Milieu aléatoire

Dans les Chapitres 2, 3, 4 et 5, nous avons étudié des limites de grande échelle de systèmes de particules en interaction en régime diffusif (i.e. la force des interactions d'un système à N particules est de l'ordre de $N^{-1/2}$). Pour avoir une convergence dans ce cadre, nous avons déjà expliqué qu'il fallait centrer les termes d'interactions pour se ramener à un résultat du type "théorème central limite". C'est pour cela que dans les équations de nos modèles, le terme de saut avait la forme suivante

$$S_t^N = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \mathbb{1}_{\{z\le\lambda_s^N\}} d\pi^j(s,z,u), \tag{7.8}$$

avec π^j $(j \ge 1)$ des mesures de Poisson indépendantes sur $\mathbb{R}_+^2 \times \mathbb{R}$ d'intensité $dt \cdot dz \cdot d\mu(u)$ avec μ une loi centrée sur $\mathbb{R}$.

Nous pouvons définir ce processus S^N d'une manière équivalente à (7.8) : pour chaque $1 \le i \le N$, soit $T_k^{N,i}$ $(k \ge 1)$ les instants de sauts de $Z^{N,i}$ (i.e. ces points forment un processus ponctuel d'intensité λ^N), et pour chaque $i, k \ge 1$, soit $U_k^{N,i}$ une variable aléatoire centrée de loi μ, indépendantes des variables $T_l^{N,j}$ $(l \ge 1, j \ge 1)$. Alors on peut écrire

$$S_t^N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \sum_{k\ge1} U_k^{N,i} \mathbb{1}_{\{T_k^{N,i}\le t\}}. \tag{7.9}$$

Autrement dit, à chaque fois que l'un des processus ponctuels crée un saut à un instant $T_k^{N,i}$, une nouvelle variable aléatoire $U_k^{N,i}$ est créée pour faire une amplitude de saut $N^{-1/2}U_k^{N,i}$.

Une alternative à ce modèle serait de considérer un modèle similaire en milieu aléatoire. Au lieu de créer une nouvelle variable $U_k^{N,i}$ à chaque saut, on choisirait au départ des variables U^i $(i \ge 1)$ i.i.d. de loi μ, de telle sorte qu'à chaque fois que l'un des processus $Z^{N,i}$ saute, on lui associe toujours la variable U^i. Nous pourrions alors définir un terme de saut de manière similaire à (7.9) :

$$S_t^N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} U^i \sum_{k\ge1} \mathbb{1}_{\{T_k^{N,i}\le t\}}. \tag{7.10}$$

Ce modèle est relativement facile à définir : il faut d'abord se donner ces variables U^i $(i \ge 1)$ i.i.d. puis construire les processus que l'on veut étudier conditionnellement à ces variables. Mais il est plus difficile à étudier que les modèles de cette thèse. Ce type de modèle a été étudié par Guionnet (1997) à la différence près que dans ce dernier article, le terme d'interaction renormalisé est la dérive et qu'il n'y a pas de saut. Pour se rapprocher de (7.10), donnons une version plus simple du modèle étudié dans Guionnet (1997) :

$$dX_t^N = b(X_t^N)dt + dW_t + \frac{1}{\sqrt{N}} \sum_{j=1}^{N} U^j dt,$$

avec W un mouvement brownien standard de dimension un, et des variables i.i.d. centrées U^i $(i \ge 1)$ indépendantes de W.

Dans Guionnet (1997), un autre processus Y^N est introduit, comme solution de

$$dY_t^N = b(Y_t^N)dt + dW_t.$$

Les preuves de cet article reposent sur le fait que ce processus Y^N est plus simple à étudier que le processus X^N, et que le théorème de Girsanov donne la densité de Radon-Nikodym explicite entre les lois de ces deux processus. Une autre difficulté de cet article est que les processus ne sont pas de dimension un, mais de dimension N.

L'idée importante à retenir est que le théorème de Girsanov permet de se ramener à un processus (ou un système) plus simple en "enlevant" le terme d'interaction. Dans notre cadre de travail, une version du théorème de Girsanov à saut ne permettrait pas de retirer le terme d'interaction (qui est le terme de saut), mais cela permettrait de modifier l'intensité des sauts. En particulier, nous pourrions nous ramener à une intensité constante, et donc le terme de saut ne serait plus un terme d'interaction (puisque l'interaction vient du fait que l'intensité stochastique est la même pour chacun des processus ponctuels du système à N particules).

Appendix A

Extended generators

In this appendix, we define precisely the notion of generators we use in Chapter 2 and we prove that classical results still hold with this notion. In the general theory of semigroups, one defines the generators on some Banach space. In the frame of semigroups related to Markov processes, one generally considers $\left(C_b(\mathbb{R}^d), ||\bullet||_\infty\right)$. In this context, the generator A of a semigroup $(P_t)_t$ is defined on the set of functions $\mathcal{D}(A) = \left\{g \in C_b(\mathbb{R}^d) : \exists h \in C_b(\mathbb{R}^d), \left|\left|\frac{1}{t}(P_t g - g) - h\right|\right|_\infty \underset{t\to 0}{\longrightarrow} 0\right\}$. Then one denotes the previous function h as Ag. If A is the generator of a diffusion, we can only guarantee that $\mathcal{D}(A)$ contains the functions that have a compact support. But the main result of the appendix is to prove Proposition A.0.3: under suitable assumptions on the processes $\bar{X}$ and X^N having semigroups $\bar{P}, P^N$ and generators $\bar{A}, A^N$,

$$\left(\bar{P}_t - P_t^N\right) g(x) = \int_0^t P^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) ds.$$

In this formula, we need to apply the generators of the processes $\left(X_t^N\right)_t$ and $\left(\bar{X}_t\right)_t$ to functions of the type $\bar{P}_s g$, and we cannot guarantee that $\bar{P}_s g$ has compact support even if we assume g to be in $C_c^\infty(\mathbb{R}^d)$.

That is why we consider extended generators (see for instance Meyn and Tweedie (1993) or Davis (1993)) defined by the point-wise convergence on $\mathbb{R}^d$ instead of the uniform convergence that allows us to define the generator on $C_b^n(\mathbb{R}^d)$ for suitable $n \in \mathbb{N}^*$ and to prove that some properties of the classical theory of semigroups still hold for this larger class of functions.

Definition A.0.1. *Let $(X_t)_t$ be a Markov process on $\mathbb{R}^d$. We define $P_t g(x) = \mathbb{E}_x\left[g\left(X_t\right)\right]$ for all functions g such that the previous expression is well-defined and finite for $x \in \mathbb{R}^d$. Then we define $\mathcal{D}'(A)$ to be the set of functions $g \in C_b(\mathbb{R}^d)$ such that for each $x \in \mathbb{R}$, $\frac{1}{t}(P_t g(x) - g(x))$ converge to some limit that we note $Ag(x)$ and such that:*

- *for all $t \geq 0$, $\int_0^t |Ag(X_s)|ds$ is almost surely finite and $(\Omega, \mathbb{P}_x)-$integrable,*
- *$g(X_t) - g(X_0) - \int_0^t Ag(X_s)ds$ is a $\mathbb{P}_x-$martingale for all x.*

We note $\mathcal{D}'(A)$ the domain of the extended generator to avoid confusions with $\mathcal{D}(A)$ which is reserved for the domain of A for the uniform convergence.

Now we generalize a classical result for generators defined with respect to the uniform convergence to extended generators.

Lemma A.0.2. *Let $(X_t)_t$ be a Markov process with semigroup $(P_t)_t$ and extended generator A.*

(1) Let $g \in \mathcal{D}'(A)$ and $x \in \mathbb{R}^d$ such that for all $t \geq 0$, $\mathbb{E}_x\left[\sup_{0\leq s\leq t} |P_sAg(X_t)|\right]$ is finite. Then the function $t \mapsto P_tg(x)$ is right differentiable at every $t \geq 0$, and we have

$$\frac{d^+}{dt}(P_tg(x)) = P_tAg(x).$$

In addition, if $P_tg \in \mathcal{D}'(A)$, then $AP_tg(x) = P_tAg(x)$.

(2) Let $g \in \mathcal{D}'(A)$ and $x \in \mathbb{R}^d$ such that there exists some non-negative function $M : \mathbb{R}^d \to \mathbb{R}_+$ such that for all $t \geq 0$, $\sup_{0\leq s\leq t} \mathbb{E}_x[M(X_s)]$ is finite and such that for all $0 \leq t \leq 1$ and $y \in \mathbb{R}$, we have $|P_tAg(y) - Ag(y)| \leq CM(y)\varepsilon(t)$ for some constant C that is allowed to depend on g, where $\varepsilon(t)$ vanishes when t goes to 0. Then the function $t \mapsto P_tg(x)$ is left differentiable at every $t > 0$, and we have

$$\frac{d^-}{dt}(P_tg(x)) = P_tAg(x).$$

Proof. For the point (1), we know that for all $h > 0$, we have:

$$\left|\frac{1}{h}(P_{t+h}g(x) - P_tg(x)) - P_tAg(x)\right| \leq \mathbb{E}_x\left[\left|\frac{1}{h}(P_hg(X_t) - g(X_t)) - Ag(X_t)\right|\right].$$

As the expression appearing within the expectation above vanishes almost surely when h goes to 0 (since $g \in \mathcal{D}'(A)$), and as we can bound it by $\sup_{0\leq s\leq t} |P_sAg(X_t)| + |Ag(X_t)|$ (using the fact that $P_hg(y) - g(y) = \int_0^h P_sAg(y)ds$ since we take $g \in \mathcal{D}'(A)$), we know that this expectation vanishes as h goes to 0 by dominated convergence. This means exactly that $\frac{d^+}{dt}(P_tg(x))$ exists and is $P_tAg(x)$.

If we suppose in addition that $P_tg \in \mathcal{D}'(A)$, then $AP_tg(x)$ is the limit of $h^{-1}(P_{t+h}g(x) - P_tg(x))$, which is $\frac{d^+}{dt}P_tg(x) = P_tAg(x)$.

Now we prove the point (2) of the lemma. Let h be some positive number. We know that

$$\left|\frac{1}{-h}(P_{t-h}g(x) - P_tg(x)) - P_tAg(x)\right|$$

is upper bounded by

$$\begin{aligned}&\mathbb{E}_x\left[\left|\frac{1}{h}(P_hg(X_{t-h}) - g(X_{t-h})) - Ag(X_{t-h})\right|\right] + \mathbb{E}_x[|Ag(X_{t-h}) - P_hAg(X_{t-h})|]\\ &\leq \mathbb{E}_x\left[\sup_{0\leq s\leq h}|Ag(X_{t-h}) - P_sAg(X_{t-h})|\right] + \mathbb{E}_x[|Ag(X_{t-h}) - P_hAg(X_{t-h})|].\end{aligned}$$

Then we just have to show that $\mathbb{E}_x\left[\sup_{0\leq s\leq h}|Ag(X_{t-h}) - P_sAg(X_{t-h})|\right]$ vanishes when h goes to 0. But this follows from the fact that it is upper bounded by $C\left(\sup_{0\leq s\leq h}\varepsilon(s)\right)\left(\sup_{0\leq r\leq t}\mathbb{E}_x[M(X_r)]\right)$. □

The goal of the next proposition is to obtain a control of the difference between the semigroups of two Markov processes, provided we dispose already of a control of the distance between the two generators. This proposition is an adaptation of Lemma 1.6.2 from Ethier and Kurtz (2005) to the notion of extended generators defined by the point-wise convergence.

Proposition A.0.3. *Let $\left(Y_t^N\right)_{t\in\mathbb{R}_+}$ and $\left(\bar{Y}_t\right)_{t\in\mathbb{R}_+}$ be $\mathbb{R}^d-$valued Markov processes whose semigroups and (extended) generators are respectively P^N, A^N and $\bar{P}, \bar{A}$. We assume that there exist $n_1, n_2, \alpha \in \mathbb{N}^*$ such that:*

(i) for all $x\in\mathbb{R}^d$ and $T>0$, $\sup\limits_{0\leq t\leq T}\mathbb{E}_x\left[\left|\left|\bar{Y}_t\right|\right|_{n_1}^{2n_1}\right]\leq C_T\left(1+||x||_{n_1}^{2n_1}\right)$ and $\sup\limits_{0\leq t\leq T}\mathbb{E}_x^N\left[\left|\left|Y_t^N\right|\right|_{n_1}^{2n_1}\right]\leq C_T\left(1+||x||_{n_1}^{2n_1}\right)$.

(ii) for all $T>0$, $\mathbb{E}\left[\sup\limits_{0\leq t\leq T}\left|\left|Y_t^N\right|\right|_{n_1}^{n_1}\right]<+\infty$.

(iii) for all $0\leq s,t\leq T$ and $x\in\mathbb{R}^d$,

$$\mathbb{E}_x\left[\left|\left|\bar{Y}_t-\bar{Y}_s\right|\right|_{n_2}^{n_2}\right]\leq C_T\left(1+||x||_{n_2}^{n_2}\right)\varepsilon(|t-s|) \quad \text{and} \quad \mathbb{E}_x^N\left[\left|\left|Y_t^N-Y_s^N\right|\right|_{n_2}^{n_2}\right]\leq C_T\left(1+||x||_{n_2}^{n_2}\right)\varepsilon(|t-s|).$$

(iv) for all $g\in C_b^\alpha\left(\mathbb{R}^d\right)$, $\bar{P}_tg\in C_b^\alpha\left(\mathbb{R}^d\right)$, and for all $T>0$, $\left|\left|\bar{P}_tg\right|\right|_{\alpha,\infty}\leq Q_T\,||g||_{\alpha,\infty}$.

(v) for all $g\in C_b^\alpha\left(\mathbb{R}^d\right), 0\leq i\leq\alpha-1, 1\leq j\leq d$ and $x\in\mathbb{R}^d$, $s\mapsto\frac{\partial^i\bar{P}_sg}{\partial x_j^i}(x)$ is continuous.

(vi) $C_b^\alpha\left(\mathbb{R}^d\right)\subseteq\mathcal{D}'(\bar{A})\cap\mathcal{D}'(A^N)$. For all $g\in C_b^\alpha\left(\mathbb{R}^d\right)$ and $x\in\mathbb{R}^d$, $\left|\bar{A}g(x)\right|\leq C\,||g||_{\alpha,\infty}\left(1+||x||_{n_1}^{n_1}\right)$ and $\left|A^Ng(x)\right|\leq C\,||g||_{\alpha,\infty}\left(1+||x||_{n_1}^{n_1}\right)$.

(vii) for all $g\in C_c^\alpha\left(\mathbb{R}^d\right)$ such that Supp $g\subseteq[-M,M]^d$, for each $1\leq i\leq d$, $\left|\left|\frac{\partial\bar{A}g}{\partial x_i}\right|\right|_\infty\leq C\,||g||_{\alpha,\infty}\left(1+M^{n_1}\right)$ and $\left|\left|\frac{\partial A^Ng}{\partial x_i}\right|\right|_\infty\leq C\,||g||_{\alpha,\infty}\left(1+M^{n_1}\right)$.

(viii) For all $g\in C_b^\alpha\left(\mathbb{R}^d\right)$ and $x,y\in\mathbb{R}^d$, $\left|\bar{A}g(x)-\bar{A}g(y)\right|\leq C\left(1+||x||_{n_1}^{n_1}+||y||_{n_1}^{n_1}\right)||x-y||_{n_2}^{n_2/2}$ and $\left|A^Ng(x)-A^Ng(y)\right|\leq C\left(1+||x||_{n_1}^{n_1}+||y||_{n_1}^{n_1}\right)||x-y||_{n_2}^{n_2/2}$.

(ix) we assume that $\lim\limits_{k\to\infty}\bar{A}g_k(x_k)=\lim\limits_{k\to\infty}A^Ng_k(x_k)=0$, for any bounded sequence of real numbers $(x_k)_k$, and for any sequence $(g_k)_k$ of $C_b^\alpha\left(\mathbb{R}^d\right)$ satisfying

(1) $\forall 0\leq i\leq\alpha-1, \forall 1\leq j\leq d, \forall x\in\mathbb{R}^d$, $\frac{\partial^ig_k}{\partial x_j^i}(x)\underset{k\to\infty}{\longrightarrow}0$

(2) $\sup\limits_k||g_k||_{\alpha,\infty}<\infty$

Then we have for each $g\in C_b^\alpha\left(\mathbb{R}^d\right), x\in\mathbb{R}^d$ and $t\in\mathbb{R}_+$

$$\left(\bar{P}_t-P_t^N\right)g(x)=\int_0^tP_{t-s}^N\left(\bar{A}-A^N\right)\bar{P}_sg(x)ds.$$

Remark A.0.4. *One can remark that under usual hypothesis, any diffusion process satisfies the conditions of Proposition A.0.3. For example, this is the case if the processes Y^N and $\bar{Y}$ satisfy a classical stochastic differential equation of the form:*

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t,$$

with W a Brownian motion and b and σ sublinear finite-dimensional valued functions. In addition, it is possible to add a jump term in the stochastic differential equation of Y^N. Whereas for $\bar{Y}$, the jump term could lead the conditions (iv) and (v) to be false.

Remark A.0.5. *Notice that the conditions of Proposition A.0.3 are not all symmetric with respect to the processes $\bar{Y}$ and Y^N. Indeed, the regularity hypothesis of the semigroup with respect to the initial condition only concerns $\bar{P}$ (see hypothesis (v) and (vi)). Moreover, hypothesis (iii) provides a stronger control on Y^N than what is needed for $\bar{Y}$.*

Proof. To begin with, let us emphasize the fact that hypothesis (i) implies that, for all $x \in \mathbb{R}^d, T \geq 0$,

$$\sup_{0\leq t\leq T} \mathbb{E}_x \left[||\bar{Y}_t||_{n_1}^{n_1} \right] \leq C_T \left(1 + ||x||_{n_1}^{n_1}\right) \text{ and } \sup_{0\leq t\leq T} \mathbb{E}_x^N \left[||Y_t^N||_{n_1}^{n_1} \right] \leq C_T \left(1 + ||x||_{n_1}^{n_1}\right) \tag{A.1}$$

since

$$\mathbb{E}_x \left[||\bar{Y}_t||_{n_1}^{n_1} \right] \leq \mathbb{E}_x \left[||\bar{Y}_t||_{n_1}^{2n_1} \right]^{1/2} \leq C_T \sqrt{1 + ||x||_{n_1}^{2n_1}} \leq C(1 + ||x||_{n_1}^{n_1}).$$

We fix $t \geq 0, g \in C_b^\alpha(\mathbb{R}^d), x \in \mathbb{R}^d$ in the proof. We note $u(s) = P_{t-s}^N \bar{P}_s g(x)$. Firstly we show that $s \mapsto \bar{P}_s g(x)$ and $s \mapsto P_s^N h(x)$ are differentiable for all $h \in C_b^\alpha(\mathbb{R}^d)$, by showing that $\bar{P}$ and P^N satisfy the hypothesis of Lemma A.0.2. The condition of the point (1) of the lemma is a straightforward consequence of (A.1) hypothesis (vi), and the conditions of the point (2) are satisfied for $M(x) = \sqrt{1 + ||x||_{n_1}^{2n_1}}$ using hypothesis (i), (iii) and $(viii)$.

Then, write $u = \Phi \circ \Psi$ with $\Phi : \mathbb{R}^2 \to \mathbb{R}$; $\Phi(v_1, v_2) = P_{v_1}^N \bar{P}_{v_2} g(x)$ and $\Psi : \mathbb{R} \to \mathbb{R}^2$; $\Psi(s) = (t - s, s)$. We have just proven that Φ is differentiable w.r.t. to v_1 and v_2 (for v_1 we have to use condition (iv) to guarantee that $\bar{P}_{v_2} g$ belongs to $C_b^\alpha(\mathbb{R}^d)$). As a consequence, and thanks to hypothesis (iv), u is differentiable and

$$\begin{aligned} u'(s) =& -\frac{d}{dr}\left(P_r^N \bar{P}_s g(x)\right)\Big|_{r=t-s} + \frac{d}{dr}\left(P_{t-s}^N \bar{P}_r g(x)\right)\Big|_{r=s} \\ =& -P_{t-s}^N A^N \bar{P}_s g(x) + P_{t-s}^N \bar{P}_s \bar{A} g(x) \\ =& P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x). \end{aligned}$$

The second equality comes from the fact that $\bar{P}$ satisfy the additional assumption of the point (1) of Lemma A.0.2 (see hypothesis (iv) and (vi)).

Now we show that u' is continuous. Indeed if it is the case, then we will have

$$u(t) - u(0) = \int_0^t u'(s)ds,$$

which is exactly the assertion. In order to prove the continuity of u', we consider a sequence $(s_k)_k$ that converges to some $s \in [0, t]$, and we write

$$\left|P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x) - P_{t-s_k}^N \left(\bar{A} - A^N\right) \bar{P}_{s_k} g(x)\right| \leq \left|\left(P_{t-s}^N - P_{t-s_k}^N\right) \left(\bar{A} - A^N\right) g_s(x)\right| \tag{A.2}$$

$$+ \left| P^N_{t-s_k} \left(\bar{A} - A^N \right) \left(\bar{P}_s - \bar{P}_{s_k} \right) g(x) \right|, \quad \text{(A.3)}$$

where $g_s = \bar{P}_s g \in C^\alpha_b(\mathbb{R}^d)$.

To show that the term (A.2) vanishes when k goes to infinity, we introduce, for all $M > 0$ the function $\varphi_M(g_s)(y) = g_s(y) \cdot \xi_M(y)$ where $\xi_M : \mathbb{R}^d \to [0,1]$ is C^∞, and $\forall ||y||_\infty \leq M, \xi_M(y) = 1$ and $\forall ||y||_\infty \geq M+1, \xi_M(y) = 0$. We note that the term (A.2) is bounded by

$$\left| \left(P^N_{t-s} - P^N_{t-s_k} \right) \left(\bar{A} - A^N \right) \varphi_M(g_s)(x) \right| + \left| \left(P^N_{t-s} - P^N_{t-s_k} \right) \left(\bar{A} - A^N \right) \left(g_s - \varphi_M(g_s) \right)(x) \right| =: A_1 + A_2.$$

If we consider the function $h_{M,s} = \left(\bar{A} - A^N \right) \varphi_M(g_s)$, using hypothesis (iii), (iv) and (vii), we have

$$\begin{aligned} A_1 \leq & \mathbb{E}^N_x \left[\left| h_{M,s}\left(Y^N_{t-s} \right) - h_{M,s}\left(Y^N_{t-s_k} \right) \right| \right] \\ \leq & \left(\sum_{i=1}^d \left|\left| \frac{\partial h_{M,s}}{\partial x_i} \right|\right|_\infty \right) \mathbb{E}^N_x \left[\left| Y^N_{t-s} - Y^N_{t-s_k} \right| \right] \\ \leq & C \left(1 + M^{n_1} \right) ||g||_{3,\infty} \varepsilon(|s - s_k|)^{1/n_2}. \end{aligned}$$

Choosing $M = M_k = \varepsilon(|s - s_k|)^{-1/(1+n_1 n_2)}$, it follows that $\lim\limits_{k\to\infty} A_1 = 0$. To see that the term A_2 vanishes, it is sufficient to notice that A_2 is bounded by

$$\mathbb{E}^N_x \left[\left| \left(\bar{A} - A^N \right) \left(g_s - \varphi_{M_k}(g_s) \right) \left(Y^N_{t-s} \right) \right| \right] + \mathbb{E}^N_x \left[\left| \left(\bar{A} - A^N \right) \left(g_s - \varphi_{M_k}(g_s) \right) \left(Y^N_{t-s_k} \right) \right| \right].$$

We know that the expressions in the expectations vanish almost surely (using hypothesis (ii) and (ix)), and then we can apply dominated convergence (using hypothesis (ii) and (vi)).

We just proved that the term (A.2) vanishes. To finish the proof, we need to show that the term (A.3) vanishes. We note that the term (A.3) is bounded by:

$$\mathbb{E}^N_x \left[\left| \bar{A} g_k \left(Y^N_{t-s_k} \right) \right| \right] + \mathbb{E}^N_x \left[\left| A^N g_k \left(Y^N_{t-s_k} \right) \right| \right],$$

where $g_k = \left(\bar{P}_s - \bar{P}_{s_k} \right) g \in C^\alpha_b(\mathbb{R}^d)$.

We have to show that the terms in the sum above vanish as k goes to infinity. Firstly we know that $\bar{A} g_k \left(Y^N_{t-s_k} \right)$ and $A^N g_k \left(Y^N_{t-s} \right)$ vanish almost surely when k goes to infinity (see hypothesis (ii), (iv), (v) and (ix)). Dominated convergence, using (A.1) and the hypothesis (ii), (iv) and (vi), then implies the result. □

Appendix B

Convergence of point processes

This appendix is based on Erny (2020).

This appendix is dedicated to prove a result interesting on its own: under some hypothesis, the convergence of point processes that admit stochastic intensities is implied by the convergence of their stochastic intensities. The result is stated formally in Theorem B.0.1 below. This result has been used to prove Theorem 2.4.1. Let us recall that it was also possible to prove Theorem 2.4.1 with Theorem IX.4.15 of Jacod and Shiryaev (2003). One advantage of Theorem B.0.1 is that it can be used in some frameworks where Theorem IX.4.15 of Jacod and Shiryaev (2003) cannot (e.g. in Example B.0.4), as it is explained below.

Theorem B.0.1. *Let $\bar{Y}^k$ and $Y^{N,k}$ $(N,k \in \mathbb{N}^*)$ be $D(\mathbb{R}_+,\mathbb{R}_+)$-valued random variables. Let $\left(\pi^k\right)_{k\in\mathbb{N}^*}$ and $\left(\bar{\pi}^k\right)_{k\in\mathbb{N}^*}$ be i.i.d. families of Poisson measures on $\mathbb{R}_+\times\mathbb{R}_+$ having Lebesgue intensity. Let $Z^{N,k}$ and $\bar{Z}^k$ be point processes defined as follows*

$$Z_t^{N,k} := \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\left\{z\leq Y_{s-}^{N,k}\right\}} d\pi^k(s,z), \quad \bar{Z}_t^k := \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\left\{z\leq \bar{Y}_{s-}^{k}\right\}} d\bar{\pi}^k(s,z), \quad k\geq 1.$$

Assume that, for every $n\geq 1$, $(Y^{N,1},\pi^1,...,Y^{N,n},\pi^n)$ converges in distribution to $(\bar{Y}^1,\bar{\pi}^1,...,\bar{Y}^n,\bar{\pi}^n)$ in $(D(\mathbb{R}_+,\mathbb{R})\times\mathcal{N})^n$, and that, for each $k\geq 1$, $\bar{Y}^k$ is independent of $\bar{\pi}^k$.

Then, for any $n\geq 1$, $\left(Z^{N,k}\right)_{1\leq k\leq n}$ converges to $\left(\bar{Z}^k\right)_{1\leq k\leq n}$ in distribution in $D(\mathbb{R}_+,\mathbb{R}^n)$. In particular, $\left(Z^{N,k}\right)_{k\geq 1}$ converges to $\left(\bar{Z}^k\right)_{k\geq 1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$ endowed with the product topology.

Remark B.0.2. *In the statement of Theorem B.0.1, we need to guarantee the following property: Poisson random measures are $\mathcal{N}-$valued random variables. This is a direct consequence of Theorem 2.6.III.(ii) of Daley and Vere-Jones (2003) and of the definition of Poisson measures (see Definition B.1.1).*

Remark B.0.3. *According to Lemma 4 of Brémaud and Massoulié (1996), a point process Z having stochastic intensity $(Y_{s-})_{s\geq 0}$ (where Y is a càdlàg process) can always be written in the form of Theorem B.0.1.*

Let us explain a bit more Theorem B.0.1 and how it can be applied. Let us note that, in Theorem B.0.1, the processes $Y^{N,k}$ are not assumed to be independent of the Poisson measures π^k.

Otherwise, the proof would be straightforward by conditioning by $Y^{N,k}$. Let us also note that the condition of independence of the limiting intensities $\bar{Y}^k$ from Poisson measures $\bar{\pi}^k$ is often satisfied and natural in many examples of application. It holds for example when the limiting intensities are deterministic, or when they are functionals of Brownian motions with respect to the same filtration as $\bar{\pi}^k$.

If the processes $Y^{N,k}$ are semimartingales, the result of Theorem B.0.1 can follow from Theorem IX.4.15 of Jacod and Shiryaev (2003), provided the convergence of the characteristics of the corresponding semimartingales holds. But we do not assume the semimartingale structure for the intensities $Y^{N,k}$ in Theorem B.0.1. This is the case in the model considered in Chapter 2.

When the limiting intensities are deterministic, Theorem B.0.1 can be compared to Theorem 1 of Brown (1978) that states that the convergence of point processes is implied by the pointwise convergence in distribution of their compensators (i.e. for each $t \geq 0$, the compensator at time t converges in distribution in $\mathbb{R}$). In Brown (1978), Theorem 1 holds when the compensator of the limit point process is a deterministic function, whereas in Theorem B.0.1, the limit point processes have stochastic intensities. Let us give an example of application in this framework.

Example B.0.4. *Consider any locally bounded $K : \mathbb{R}_+ \to \mathbb{R}$ and any Lipschitz continuous function $f : \mathbb{R} \to \mathbb{R}_+$. Define X^N as solution of*

$$X_t^N = \frac{1}{N} \sum_{j=1}^{N} \int_{[0,t]\times\mathbb{R}_+} K(t-s) \mathbb{1}_{\left\{z \leq f(X_{s-}^N)\right\}} d\pi^j(s,z),$$

where π^j $(j \geq 1)$ are independent Poisson measures on $\mathbb{R}_+^2$ with Lebesgue intensity, and $\bar{X}$ as the deterministic solution of

$$\bar{X}_t = \int_0^t K(t-s) f(\bar{X}_s) ds.$$

Let

$$Z_t^{N,i} = \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\left\{z \leq f(X_{s-}^N)\right\}} d\pi^i(s,z) \text{ and } \bar{Z}_t^i = \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\left\{z \leq f(\bar{X}_{s-})\right\}} d\pi^i(s,z).$$

Then, Theorem 8 of Delattre et al. (2016) states that, for $i \geq 1$, for all $T \geq 0$,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|Z_t^{N,i} - \bar{Z}_t^i\right|\right] \leq C_T N^{-1/2},$$

for some constant C_T. In other words, $Z^{N,i}$ converges to $\bar{Z}$ in a L^1-sense. As $\bar{X}$ is a deterministic function, Theorem 1 of Brown (1978), and a fortiori Theorem B.0.1, can be used to show the (weaker) convergence in distribution in $D(\mathbb{R}_+, \mathbb{R})$ of $Z^{N,i}$ to $\bar{Z}$, provided X^N converges to $\bar{X}$.

Theorem B.0.1 allows us to consider processes that are not semimartingales, such as Hawkes processes and Volterra processes. Since the stochastic intensities of Hawkes processes are not, in general, semimartingales, Theorem B.0.1 can be interesting to show the convergence of Hawkes processes, provided that one can show the convergence of their stochastic intensities. Let us give an example of application of Theorem B.0.1 in this case. The example is based on Examples 7.3 and 7.4 of Abi Jaber et al. (2019).

Example B.0.5. *Let us consider* $K(t) := t^\gamma$ *for some* $\gamma > 0$, $K^N(t) := K(t/N)$ *and some Poisson random measure* π *on* $\mathbb{R}_+^2$ *having Lebesgue intensity. Let* X^N *satisfies*

$$X_t^N = \int_{[0,t]\times\mathbb{R}_+} K^N(t-s)\mathbb{1}_{\{z\le |X_{s-}^N|\}} d\pi(s,z) - \int_0^t K^N(t-s)\left|X_s^N\right| ds.$$

Theorem 7.2 of Abi Jaber et al. (2019) implies that the sequence of processes $(\tilde{X}_t^N)_{t\ge 0} = (N^{-1}X_{Nt}^N)_{t\ge 0}$ *has converging subsequences (in distribution in the topology* L^2_{loc}*), and that every limit process* $(\bar{X}_t)_{t\ge 0}$ *satisfies*

$$\bar{X}_t = \int_0^t K(t-s)\sqrt{|\bar{X}_s|}dB_s, \tag{B.1}$$

for some standard Brownian motion B. *Besides, one can prove with standard arguments, the tightness of* $(\tilde{X}^N)_N$ *in Skorohod topology. Then, we can consider a subsequence of* $(\tilde{X}^N)_N$ *that converges in distribution in the topology of* L^2_{loc} *and in Skorohod topology. The limit for both topologies is necessarily the same on Skorohod space. Indeed, let* $\hat{x}$ *be the limit for* L^2_{loc} *topology and* $\check{x}$ *be the limit for Skorohod topology of a sequence of càdlàg functions* $(x^n)_n$. *The* L^2_{loc} *convergence implies that* x_t^n *converges to* $\hat{x}_t$ *for Lebesgue-a.e.* $t\ge 0$, *whence, the convergence for all continuity point* t *of* $\hat{x}$. *Besides, the convergence in Skorohod topology implies the convergence* x_t^n *to* $\check{x}_t$ *for every continuity point* t *of* $\check{x}$. *This implies that* $\hat{x}$ *and* $\check{x}$ *have the same continuity points, and that* $\check{x} = \hat{x}$.

This implies the convergence of (a subsequence of) $Y^N := |\tilde{X}^N|$ *to* $\bar{Y} := |\bar{X}|$ *in Skorohod topology. Moreover, the Brownian motion can be shown to be necessarily independent of the Poisson measure* π *(using Theorem II.6.3 of Ikeda and Watanabe (1989)). Then, Theorem B.0.1 implies the convergence in distribution in Skorohod topology of* $Z_t^N := \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\le Y_{s-}^N\}} d\pi(s,z)$ *to the point process* $\bar{Z}_t := \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\le \bar{Y}_{s-}\}} d\bar{\pi}(s,z)$, *where* $\bar{\pi}$ *is independent of* $\bar{Y}$. *To the best of our knowledge, there is no classical way to prove this convergence.*

Befor giving the proof of Theorem B.0.1, let us recall some useful results about Poisson measures.

B.1 Some basic properties of Poisson measures

Let us recall the usual definition of random measures and Poisson measures. We restrict this definition to the space $\mathbb{R}_+^2$ since we only need this space in the section, but Definition B.1.1 can be generalized to any measurable space.

Definition B.1.1. *A locally finite random measure on* $\mathbb{R}_+^2$ *is a* $\mathcal{M}-$*valued random variable, where* $\mathcal{M}$ *is endowed with the* $\sigma-$*algebra generated by the functions* $\pi \in \mathcal{N} \mapsto \pi(B)$ $(B \in \mathcal{B}(\mathbb{R}_+^2))$.

A Poisson measure on $\mathbb{R}_+^2$ *is a locally finite random measure* π *satisfying:*

- *for all* $B \in \mathcal{B}(\mathbb{R}_+^2)$, $\pi(B)$ *follows a Poisson distribution,*
- *for every* $n \in \mathbb{N}^*$, *for all disjoint sets* $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}_+^2)$, *the variables* $\pi(B_i)$ $(1 \le i \le n)$ *are independent.*

The function $\mu : B \in \mathcal{B}(\mathbb{R}_+^2) \mapsto \mathbb{E}\left[\pi(B)\right]$ *is a measure on* $\mathbb{R}_+^2$ *that is called the intensity of* π.

Remark B.1.2. *In Definition B.1.1, we can consider Poisson distribution with parameter infinity. A Poisson variable with parameter infinity is a random variable* X *satisfying* $X = +\infty$ *a.s.*

Let us begin with an elementary lemma.

Lemma B.1.3. *Let $D \in \mathcal{B}(\mathbb{R}_+^2)$, μ be a (deterministic) measure on $\mathbb{R}_+^2$, and π be a Poisson measure on $\mathbb{R}_+^2$ with intensity μ. If $\mu(D) = 0$, then, a.s. $\pi(D) = 0$.*

Proof. By definition, $\pi(D)$ is a Poisson variable with parameter $\mu(D) = 0$. □

Now, we state and prove another classical property of Poisson measures.

Lemma B.1.4. *Let π be a Poisson measure on $\mathbb{R}_+^2$ with Lebesgue intensity. Then,*

$$\mathbb{P}\left(\forall t \geq 0, \pi(\{t\} \times \mathbb{R}_+) \leq 1\right) = 1.$$

Proof. Let us write

$$\mathbb{P}\left(\exists t, \pi(\{t\} \times \mathbb{R}_+) \geq 2\right) \leq \mathbb{P}\left(\sum_{\substack{(t_1,x_1),(t_2,x_2)\in\pi \\ (t_1,x_1)\neq(t_2,x_2)}} \mathbb{1}_{\{t_1=t_2\}} \geq 1\right) \leq \mathbb{E}\left[\sum_{\substack{(t_1,x_1),(t_2,x_2)\in\pi \\ (t_1,x_1)\neq(t_2,x_2)}} \mathbb{1}_{\{t_1=t_2\}}\right].$$

Then, introducing the random measure $\pi^{(2)}$ (called the second factorial measure of π) defined as

$$\pi^{(2)} := \sum_{\substack{(t_1,x_1),(t_2,x_2)\in\pi \\ (t_1,x_1)\neq(t_2,x_2)}} \delta_{(t_1,x_1,t_2,x_2)},$$

we can write

$$\sum_{\substack{(t_1,x_1),(t_2,x_2)\in\pi \\ (t_1,x_1)\neq(t_2,x_2)}} \mathbb{1}_{\{t_1=t_2\}} = \int_{\mathbb{R}_+^4} \mathbb{1}_{\{t_1=t_2\}} d\pi^{(2)}(t_1,x_1,t_2,x_2).$$

According to the multivariate Mecke equation (see Theorem 4.4 of Last and Penrose (2017)) and Fubini-Tonelli's theorem, the expectation of the integral above is

$$\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \mathbb{1}_{\{t_1=t_2\}} dt_1 dt_2 dx_1 dx_2 = 0.$$

This allows to conclude that $\mathbb{P}\left(\exists t, \pi(\{t\} \times \mathbb{R}_+) \geq 2\right) = 0$. □

B.2 Proofs

In a first time, let us the following result: the vague convergence of locally finite point measures implies the convergence of their atoms.

Proposition B.2.1. *Let P^k ($k \in \mathbb{N}$) and P be locally finite simple point measures on $\mathbb{R}_+^2$ such that P^k converges vaguely to P. Let T, M be positive real numbers such that $P(\partial[0,T] \times [0,M]) = 0$. Denote $n_k := P^k([0,T] \times [0,M]), n := P([0,t] \times [0,M])$ and $(t_i^k, z_i^k)_{1\leq i\leq n_k}$ (resp. $(t_i, z_i)_{1\leq i\leq n}$) the atoms of $P^k_{|[0,T]\times[0,M]}$ (resp. $P_{|[0,T]\times[0,M]}$).*

Then, for k large enough, $n_k = n$, and there exists a sequence of permutation $(\sigma^k)_k$ of $[\![1,n]\!]$ such that, for all $1 \leq i \leq n$, $t^k_{\sigma^k(i)}$ (resp. $z^k_{\sigma^k(i)}$) converges to t_i (resp. z_i) as k goes to infinity.

Proof. To begin with, Proposition A2.6.II.(iv) of Daley and Vere-Jones (2003) implies that n_k converges to n as k goes to infinity. As n_k $(k \in \mathbb{N})$ and n are integers, this implies that $n_k = n$ for k large enough.

To show the convergence of t_i^k and z_i^k, $(1 \leq i \leq n)$ let us fix some $\varepsilon > 0$. Then, for each $1 \leq i \leq n$, consider an open ball B_i centered on (t_i, z_i) of radius smaller than ε (for the supremum norm) such that, for all $i \neq j, B_i \cap B_j = \varnothing$.

Thanks to Proposition A2.6.II.(iv) of Daley and Vere-Jones (2003), we know that $P^k(B_i)$ converges to $P(B_i) = 1$ $(1 \leq i \leq n)$. Since P^k are point measures, this implies that for all $1 \leq i \leq n$, $P^k(B_i) = 1$ for k large enough. As the sets B_i $(1 \leq i \leq n)$ are disjoint, there exists a permutation σ^k such that $(t^k_{\sigma^k(i)}, z^k_{\sigma^k(i)}) \in B_i$. Hence, for k large enough, for all $1 \leq i \leq n$, $|t^k_{\sigma^k(i)} - t_i| \leq \varepsilon$ and $|z^k_{\sigma^k(i)} - z_i| \leq \varepsilon$. □

Now, let us state and prove the main step to prove Theorem B.0.1.

Theorem B.2.2. *Let $\Phi : D(\mathbb{R}_+, \mathbb{R}_+)^m \times \mathcal{N}^m \to D(\mathbb{R}_+, \mathbb{R}^m)$ be defined as*

$$\Phi(x,\pi)_t := \left(\int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z \leq x^j_{s-}\}} d\pi^j(s,z) \right)_{1\leq j\leq m}.$$

Let $(x,\pi) \in D(\mathbb{R}_+, \mathbb{R}_+)^m \times \mathcal{N}^m$. A sufficient condition for Φ to be continuous at (x,π) is:

(a) for each $1 \leq j \leq m$, for every $t \geq 0$, $\pi^j(\{t\} \times \mathbb{R}_+) \leq 1$,

(b) for each $1 \leq j \leq m$, for every $t \geq 0$, if $\pi^j(\{t\} \times \mathbb{R}_+) = 1$, then, for all $i \neq j$, $\pi^i(\{t\} \times \mathbb{R}_+) = 0$,

(c) for each $1 \leq j \leq m$, for every $t \geq 0$ such that $\pi^j(\{t\} \times \mathbb{R}_+) = 1$, x^j is continuous at t,

(d) for each $1 \leq j \leq m$, $\pi^j \left(\{(t, x^j_{t-}) \ : \ t \geq 0\} \right) = 0$.

Before proving Theorem B.2.2, let us point out that, in general, Φ is not continuous at every point of $D(\mathbb{R}_+, \mathbb{R}_+)^m \times \mathcal{N}^m$. This is shown in Example B.2.3, where hypothesis (d) is not satisfied.

Example B.2.3. *Let us consider the point measure $\pi = \delta_{(1,1)}$ and the constant function $x : t \in \mathbb{R}_+ \mapsto 1$. In addition, we consider the functions x^n defined as in Figure B.1 below. Obviously, $||x - x^n||_\infty = 1/n$, but $\Phi(x,\pi)_t = \mathbb{1}_{\{t\geq 1\}}$ and $\Phi(x^n, \pi) = 0$. In other words, x^n converges uniformly to x, but $\Phi(x^n, \pi)$ does not converge to $\Phi(x,\pi)$ for non-trivial topologies.*

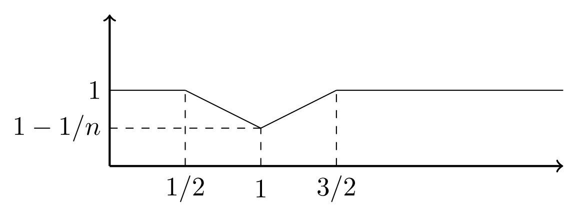

Figure B.1: Graph of x^n

Proof of Theorem B.2.2. Let $(x^k, \pi^k) = (x^{1,k}, \dots, x^{m,k}, \pi^{1,k}, \dots, \pi^{m,k})_k$ be a sequence of $D(\mathbb{R}_+, \mathbb{R})^m \times \mathcal{N}^m$ converging to $(x, \pi) = (x^1, \dots, x^m, \pi^1, \dots, \pi^m)$. Let $Z := \Phi(x, \pi)$ and $Z^k := \Phi(x^k, \pi^k)$.

Let us consider $T \geq 0$ such that for all $1 \leq j \leq m$, $\pi^j(\{T\} \times \mathbb{R}_+) = 0$ and for all $k \in \mathbb{N}^*, \pi^{j,k}(\{T\} \times \mathbb{R}_+) = 0$. In particular T is a point of continuity of Z and of each Z^k, and, as the set of all the atoms of the measures π^j and $\pi^{j,k}$ $(1 \leq j \leq m, k \geq 1)$ is countable, the set of points T satisfying the previous conditions is dense.

According to the proof of Theorem 16.2 of Billingsley (1999), in order to prove the convergence of Z^k to Z in $D(\mathbb{R}_+, \mathbb{R}^m)$, it is sufficient to prove this convergence in $D([0,T], \mathbb{R}^m)$ for the points T satisfying the conditions of the previous paragraph. Indeed, these points T are continuity points of Z and there exists an increasing sequence of such points T going to infinity. Then, by Lemma C.3.5 (whose hypothesis is satisfied thanks to hypothesis (a) and (b)), the convergence of Z^k to Z in $D([0,T], \mathbb{R}^m)$ will follow from the convergence of Z^k_t to Z_t for every point t satisfying the same conditions as T. Let us show the convergence of $Z^{j,k}_t$ to Z^j_t for every $1 \leq j \leq m$. In the rest of the proof, we work with fixed j, t, T.

To show this, fix some $M > \max(||x^j||_{\infty,[0,T]}, \sup_k ||x^{j,k}||_{\infty,[0,T]})$ (where we know that the supremum over k of $||x^{j,k}||_{\infty,[0,T]}$ is finite since $(x^{j,k})_k$ converges in Skorohod topology) such that $\pi^j(\mathbb{R}_+ \times [0,M]) = 0$, and write $\{(\tau_i, \zeta_i) : 1 \leq i \leq N\}$ the set of the atoms of $\pi^j_{|[0,t]\times[0,M]}$ and $\left\{\left(\tau^k_i, \zeta^k_i\right) : 1 \leq i \leq N_k\right\}$ that of $\pi^{j,k}_{|[0,t]\times[0,M]}$.

Then, Proposition B.2.1 implies that $N_k = N$ for all k (large enough), and that, for each $1 \leq i \leq N$, τ^k_i and ζ^k_i converge respectively to τ_i and ζ_i (we can assume that $\sigma^k = Id$ in the statement of Proposition B.2.1, possibly reordering the indexes of the atoms of every P^k).

Notice that

$$Z^{j,k}_t = \sum_{i=1}^{N} \mathbb{1}_{\left\{\zeta^k_i \leq x^{j,k}_{\tau^k_i -}\right\}} \mathbb{1}_{\{\tau^k_i \leq t\}}.$$

To end the proof, one has to note that $\mathbb{1}_{\{\zeta^k_i \leq x^{j,k}_{\tau^k_i -}\}}$ converges to $\mathbb{1}_{\{\zeta_i \leq x^j_{\tau_i -}\}}$, and that $\mathbb{1}_{\{\tau^k_i \leq t\}}$ converges to $\mathbb{1}_{\{\tau_i \leq t\}} = 1$. By hypothesis (d), $\zeta_i \neq x^j_{\tau_i -}$, whence there are two cases, either $\zeta_i < x^j_{\tau_i -}$ or $\zeta_i > x^j_{\tau_i -}$. In the first case, we consider $\varepsilon > 0$ such that $\zeta_i + \varepsilon < x^j_{\tau_i -}$. Then, noticing that hypothesis (c) guarantees that x^j is continuous at τ_i, Lemma C.3.4 and the convergence of τ^k_i and ζ^k_i respectively to τ_i and ζ_i imply that, for k large enough, $\zeta^k_i < \zeta_i + \varepsilon/3 < x^j_{\tau_i} - \varepsilon/3 < x^{j,k}_{\tau^k_i -}$, what implies the convergence of $\mathbb{1}_{\{\zeta^k_i \leq x^{j,k}_{\tau^k_i -}\}}$ to $\mathbb{1}_{\{\zeta_i \leq x^j_{\tau_i -}\}}$. The second case is handled in the same way, as well as the convergence of $\mathbb{1}_{\{\tau^k_i \leq t\}}$ to $\mathbb{1}_{\{\tau_i \leq t\}}$, recalling that $\pi^j(\{t\} \times \mathbb{R}_+) = 0$, and so $\tau_i < t$. □

Now, we can prove the main result of this section.

Proof of Theorem B.0.1. Step 1. Let us show that $(Z^{N,k})_{1 \leq k \leq n}$ converges to $(\bar{Z}^k)_{1 \leq k \leq n}$ as N goes to infinity in $D(\mathbb{R}_+, \mathbb{R}^n)$.

Since $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M})^n$ is a separable metric space (see Theorem 16.3 of Billingsley (1999) for $D(\mathbb{R}_+, \mathbb{R})$, and Theorem A2.6.III.(i) of Daley and Vere-Jones (2003) for $\mathcal{M}$), we can apply Skorohod representation theorem (see e.g. Theorem 6.7 of Billingsley (1999)) to show the almost sure convergence of a sequence $((\tilde{Y}^{N,1}, \tilde{\pi}^{N,1}), \dots, (\tilde{Y}^{N,n}, \tilde{\pi}^{N,n}))$ to $((\tilde{Y}^1, \tilde{\pi}^1), \dots, (\tilde{Y}^n, \tilde{\pi}^n))$ in $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{N})^n$ as N goes to infinity, where these variables have respectively the same distribution as $((Y^{N,1}, \pi^1), \dots, (Y^{N,n}, \pi^n))$ and $((\bar{Y}^1, \bar{\pi}^1), \dots, (\bar{Y}^n, \bar{\pi}^n))$.

Then Theorem B.2.2 implies the almost sure convergence of the multivariate point processes $(\widetilde{Z}^{N,k})_{1\leq k\leq n} := \Phi((\widetilde{Y}^{N,k},\widetilde{\pi}^{N,k})_{1\leq k\leq n})$ to $(\widetilde{Z}^{k})_{1\leq k\leq n} := \Phi((\widetilde{Y}^{k},\widetilde{\pi}^{k})_{1\leq k\leq n})$ in $D(\mathbb{R}_+,\mathbb{R}^n)$. Let us show that the hypothesis of Theorem B.2.2 are satisfied almost surely. Hypothesis (a) is a classical property of Poisson measures (see Lemma B.1.4). Hypothesis (b) is satisfied since the Poisson measures $\bar{\pi}^i$ $(i\geq 1)$ are independent, whence, considering $i\neq j$, denoting $A(\bar{\pi}^j)$ the set of points $t\geq 0$ such that $\bar{\pi}^j(\{t\}\times\mathbb{R}_+)=0$, we have that $\bigcup_{t\in A(\bar{\pi}^j)}\{t\}\times\mathbb{R}_+$ is a null set (since $A(\bar{\pi}^j)$ is finite or countable) independent of $\bar{\pi}^i$, and consequently

$$\mathbb{P}\left(\bar{\pi}^i(\bigcup_{t\in A(\bar{\pi}^j)}\{t\}\times\mathbb{R}_+)\neq 0\right)=\mathbb{E}\left[\mathbb{P}\left(\bar{\pi}^i(\bigcup_{t\in A(\bar{\pi}^j)}\{t\}\times\mathbb{R}_+)\neq 0|\bar{\pi}^j\right)\right]=0,$$

by Lemma B.1.3.

Hypothesis (c) and (d) are satisfied for a similar reason. For (c), one has to observe that

$$\mathbb{P}\left(\exists t>0,\pi^j(\{t\}\times\mathbb{R}_+)=1 \text{ and } \bar{Y}^j \text{ is not continuous at } t\right)=\mathbb{P}\left(\bar{\pi}^j(\bigcup_{t\in D(\bar{Y}^j)}\{t\}\times\mathbb{R}_+)\geq 1\right),$$

where $D(\bar{Y}^j)$ is the set of discontinuity points of $\bar{Y}^j$. As $D(\bar{Y}^j)$ is a.s. finite or countable (see e.g. the discussion after Lemma 1 of Section 12 of Billingsley (1999)), whence $\bigcup_{t\in D(\bar{Y}^j)}\{t\}\times\mathbb{R}_+$ is a null set independent of $\bar{\pi}^j$, and Lemma B.1.3 gives the result:

$$\mathbb{P}\left(\bar{\pi}^j(\bigcup_{t\in D(\bar{Y}^j)}\{t\}\times\mathbb{R}_+)\geq 1\right)=\mathbb{E}\left[\mathbb{P}\left(\bar{\pi}^j(\bigcup_{t\in D(\bar{Y}^j)}\{t\}\times\mathbb{R}_+)\geq 1|\bar{Y}^j\right)\right]=0.$$

And hypothesis (d) holds true, because the set $\{(t,\bar{Y}^j_{t-}):t\geq 0\}$ is also a null set independent of $\bar{\pi}^j$.

Then, the almost sure convergence of $(\widetilde{Z}^{N,k})_{1\leq k\leq n}$ to $(\widetilde{Z}^{k})_{1\leq k\leq n}$ implies the convergence in distribution of $(Z^{N,k})_{1\leq k\leq n}$ to $(Z^{k})_{1\leq k\leq n}$ in $D(\mathbb{R}_+,\mathbb{R}^n)$.

Step 2. We have shown that, for every $n\in\mathbb{N}^*$, $(Z^{N,k})_{1\leq k\leq n}$ converges to $(\bar{Z}^{k})_{1\leq k\leq n}$ in distribution in $D(\mathbb{R}_+,\mathbb{R}^n)$. This implies the convergence in the weaker topology $D(\mathbb{R}_+,\mathbb{R})^n$. Then, the convergence of $(Z^{N,k})_{k\geq 1}$ to $(\bar{Z}^k)_{k\geq 1}$ as N goes to infinity in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$ is classical (see e.g. Theorem 3.29 of Kallenberg (1997)). □

Appendix C

Standard results

C.1 Grönwall's lemma

The version of Grönwall's lemma we use in the thesis is a particular case of Grönwall's inequality (2019). We state it below.

Lemma C.1.1. *Let γ and u be non-negative measurable functions defined on $\mathbb{R}_+$, and let α be a non-negative constant. Assume that $u \in L^1_{loc}(dt)$, and that for all $t \geq 0$,*

$$u(t) \leq \gamma(t) + \alpha \int_0^t u(s)ds, \tag{C.1}$$

then for all $t \geq 0$, we have

$$u(t) \leq \gamma(t) + \alpha \int_0^t \gamma(s)e^{\alpha(t-s)}ds.$$

Moreover, if γ is nondecreasing then, for all $t \geq 0$, we have:

$$u(t) \leq \gamma(t)e^{\alpha t}.$$

An interesting point of Lemma C.1.1 is that it does not require any continuity hypothesis on u, contrarily to more common versions of Grönwall's lemma. We reproduce the proof of Grönwall's inequality (2019) for self-containedness.

Proof. We note μ the measure $\mu(dt) = \alpha dt$. Firstly we prove by induction on n that for all $n \in \mathbb{N}$

$$u(t) \leq \gamma(t) + \int_0^t \gamma(s) \sum_{k=0}^{n-1} \mu^{\otimes k}(A_k(s,t))\mu(ds) + R_n(t), \tag{C.2}$$

where $R_n(t) = \displaystyle\int_0^t u(s)\mu^{\otimes n}(A_n(s,t))\mu(ds)$ and $A_n(s,t) = \{(s_1, \dots, s_n) \in]s,t[^n :\ s < s_1 < \dots < s_n < t)\}$.

The case $n = 0$ is inequality (C.1). To show the induction step, we replace the assumed inequality in the expression of $R_n(t)$ and obtain

$$R_n(t) \leq \int_0^t \gamma(s)\mu^{\otimes n}(A_n(s,t))\mu(ds) + \widetilde{R}_n(t),$$

with $\widetilde{R}_n(t) = \int_0^t \left(\int_0^r u(s)\mu(ds)\right) \mu^{\otimes n}(A_n(r,t))\mu(dr)$.

Using Fubini-Tonelli's theorem, we have $\widetilde{R}_n(t) = R_{n+1}(t)$. As a consequence, equality (C.2) is proved for all $n \in \mathbb{N}$.

A straightforward induction gives

$$\mu^{\otimes n}(A_n(s,t)) = \frac{\alpha^n}{n!}(t-s)^n,$$

implying that, for all $n \in \mathbb{N}$,

$$u(t) \leq \gamma(t) + \int_0^t \gamma(s) \sum_{k=0}^{n-1} \frac{\alpha^k}{k!}(t-s)^k \mu(ds) + R_n(t). \tag{C.3}$$

As $R_n(t) = \frac{\alpha^n}{n!} \int_0^t u(s)(t-s)^n ds \leq \frac{\alpha^n}{n!} t^n \int_0^t u(s)ds$, we know that $R_n(t)$ vanishes when n goes to infinity, since u is locally integrable. Letting n go to infinity in equation (C.3), we obtain the assertion. □

C.2 Osgood's lemma

We have used many times in Chapter 6 a generalization of Grönwall's lemma, which is Osgood's lemma (see e.g. Lemma 3.4 of Bahouri et al. (2011)).

Lemma C.2.1. *Let ϱ be a measurable function from $[t_0, T]$ to $[0, b]$, γ a locally integrable function from $[t_0, T]$ to $\mathbb{R}_+$, and μ a continuous and non-decreasing function from $[0, b]$ to $\mathbb{R}_+$. Suppose that for all $t \in [t_0, T]$ and for some $c \in]0, b[$,*

$$\varrho(t) \leq c + \int_{t_0}^t \gamma(s)\mu(\varrho(s))ds.$$

Then, with $M(x) := \int_x^b \frac{ds}{\mu(s)}$,

$$-M(\varrho(t)) + M(c) \leq \int_{t_0}^t \gamma(s)ds.$$

C.3 Lemmas about Skorohod topology

In this section, we state and prove some results of Skorohod space that we use in the proofs in the other chapters. They rely strongly on Chapters 1 and 3 of Billingsley (1999).

Let us recall that if (E, d) is a Polish space and $T > 0$, $D([0, T], E)$ denotes the set of càdlàg $E-$valued functions defined on $[0, T]$, endowed with Skorohod topology. This space is Polish, and the convergence of a sequence $(x_n)_n$ to some x in this space is equivalent to: there exists a sequence of increasing continuous functions $(\lambda_n)_n$ such that $\lambda_n(0) = 0$, $\lambda_n(T) = T$ and both

$$||\lambda_n - Id||_{\infty,[0,T]} \text{ and } ||x_n - x \circ \lambda_n||_{\infty,[0,T]}$$

vanish as n goes to infinity. One can also consider the space $D(\mathbb{R}_+, E)$ with similar properties. This space is studied in detail in the section 16 of Billingsley (1999).

These two first lemmas are used in the proof of Proposition 6.1.8.

Lemma C.3.1. *Let $T > 0$, and $(x_n)_n$ be a sequence of càdlàg functions converging to some càdlàg function x in $D([0,T],\mathbb{R})$. In addition, if $(y_n)_n$ is a sequence of càdlàg functions that satisfies*

$$\sup_{0\le t\le T} |x_n(t) - y_n(t)| \underset{n\to\infty}{\longrightarrow} 0,$$

then, the sequence $(x_n, y_n)_n$ converges to (x,x) in $D([0,T],\mathbb{R}^2)$.

Proof. By hypothesis, there exists a sequence of increasing and continuous functions λ_n such that $\lambda_n(0) = 0$, $\lambda_n(T) = T$, and that both

$$\sup_{0\le t\le T} |x_n(t) - x(\lambda_n(t))| \text{ and } \sup_{0\le t\le T} |t - \lambda_n(t)|$$

vanish as n goes to infinity.

To prove the lemma, it is sufficient to prove that $(y_n)_n$ converges to x in the same sense as above with the same sequence of time-changes $(\lambda_n)_n$. It is a direct consequence of the fact that, for all $0 \le t \le T$,

$$|y_n(t) - x(\lambda_n(t))| \le |y_n(t) - x_n(t)| + |x_n(t) - x(\lambda_n(t))|.$$

□

Lemma C.3.2. *Let $T > 0$, and x and x_n ($n \in \mathbb{N}$) be càdlàg functions. Let λ_n ($n \in \mathbb{N}$) be continuous, increasing functions satisfying $\lambda_n(0) = 0$, $\lambda_n(T) = T$, and that both*

$$\sup_{0\le t\le T} |x_n(t) - x(\lambda_n(t))| \text{ and } \sup_{0\le t\le T} |t - \lambda_n(t)|$$

vanish as n goes to infinity. Then,

$$\sup_{0\le t\le T} \left| \int_0^t x_n(s)ds - \int_0^{\lambda_n(t)} x(s)ds \right| \underset{n\to\infty}{\longrightarrow} 0.$$

Proof. The proof relies on the following inequalities: for any $0 \le t \le T$

$$\begin{aligned}
&\left| \int_0^t x_n(s) - \int_0^{\lambda_n(t)} x(s)ds \right| \\
&\quad\le \int_0^t |x_n(s) - x(\lambda_n(s))|ds + \int_0^t |x(\lambda_n(s) - x(s)|ds + \left| \int_0^t x(s)ds - \int_0^{\lambda_n(t)} x(s)ds \right| \\
&\quad\le T \sup_{0\le s\le T} |x_n(s) - x(\lambda_n(s))| + \int_0^T |x(\lambda_n(s)) - x(s)|ds + \sup_{0\le s\le T} |x(s)| \cdot \sup_{0\le s\le T} |s - \lambda_n(s)|.
\end{aligned}$$

To conclude the proof, one has to notice that the three terms above do not depend on t and vanish as n goes to infinity. □

Remark C.3.3. *An interesting consequence of the previous lemma is that, if $(x_n, y_n)_n$ converges to (x, y) in $D([0,T],\mathbb{R}^2)$, then $(x_n, \int_0^\cdot y_n(s)ds)_n$ converges to $(x, \int_0^\cdot y(s)ds)$ in $D([0,T],\mathbb{R}^2)$. Note that it is important to have convergence in $D([0,T],\mathbb{R}^2)$ instead that in $D([0,T],\mathbb{R})^2$. The difference between these two topologies is that, the convergence in $D([0,T],\mathbb{R}^2)$ means that the two coordinates have to share the same sequence of time-changes, whereas for the convergence in $D([0,T],\mathbb{R})^2$ each of the coordinates has its own sequence.*

The two following lemmas are used to prove Theorem B.2.2

Lemma C.3.4. *Let* $(x_N)_N$ *be a sequence of* $D(\mathbb{R}_+, \mathbb{R})$ *converging to some* $x \in D(\mathbb{R}_+, \mathbb{R})$*, and* $(t_N)_N$ *be a sequence converging to some* $t > 0$*. If* x *is continuous at* t*, then* $x_N(t_N-) \to x(t)$*.*

Proof. Let $T > t$ such that x is continuous at T. By Theorem 16.2 of Billingsley (1999), x_N converges to x in $D([0,T], \mathbb{R})$. Consequently, there exists a sequence of continuous increasing bijective functions $\lambda_N : [0,T] \to [0,T]$ such that $||\lambda_N - Id||_{\infty,[0,T]}$ and $||x_N - x \circ \lambda_N||_{\infty,[0,T]}$ vanish as N goes to infinity. Then, as λ_N is continuous,

$$\begin{aligned} |x(t) - x_N(t_N-)| &\leq |x(t) - x(\lambda_N(t_N))| + |x(\lambda_N(t_N-)) - x_N(t_N-)| \\ &\leq |x(t) - x(\lambda_N(t_N))| + ||x \circ \lambda_N - x_N||_{\infty,[0,T]} \end{aligned}$$

vanishes as N goes to infinity since

$$|\lambda_N(t_N) - t| \leq |\lambda_N(t_N) - t_N| + |t_N - t| \leq ||\lambda_n - Id||_{\infty,[0,T]} + |t_N - t| \,.$$

□

Lemma C.3.5. *Let* $T > 0$*,* $k \in \mathbb{N}^*$*,* $n_i \in \mathbb{N}^*$*, and consider increasing sequences* $0 = t_{i,0} < t_{i,1} < \ldots < t_{i,n_i-1} < t_{i,n_i} = T$*,* $0 = t^N_{i,0} < t^N_{i,1} < \ldots < t^N_{i,n_i^N-1} < t^N_{i,n_i^N} = T$ *(*$1 \leq i \leq k$*). We define the functions* $g, g_N \in D([0,T], \mathbb{R}^k)$ *by*

$$\begin{cases} g(t) = \left(\sum_{j=0}^{n_i-1} \mathbb{1}_{[t_{i,j}, t_{i,j+1}[}(t) j\right)_{1\leq i\leq k} & \textit{for } t \in [0,T[, \\ g(T) = (n_i - 1)_{1\leq i\leq k} \,, \end{cases}$$

and

$$\begin{cases} g_N(t) = \left(\sum_{j=0}^{n_i^N-1} \mathbb{1}_{[t^N_{i,j}, t^N_{i,j+1}[}(t) j\right)_{1\leq i\leq k} & \textit{for } t \in [0,T[, \\ g_N(T) = \left(n_i^N - 1\right)_{1\leq i\leq k} . \end{cases}$$

We assume that there exists a dense subset $A \subseteq [0,T]$ *containing* T *and such that, for all* $t \in A, g_N(t)$ *converges to* $g(t)$*. Moreover, we assume that for all* $i_1 \neq i_2$*, for all* $j_1 \in [\![1, n_{i_1-1}]\!]$ *and* $j_2 \in [\![1, n_{i_2-1}]\!]$*,* $t_{i_1,j_1} \neq t_{i_2,j_2}$*. Then* g_N *converges to* g *in* $D\left([0,T], \mathbb{R}^k\right)$*.*

Proof. Since $g_N(T) = (n_i^N - 1)_{1\leq i\leq k}$ converges to $g(T) = (n_i - 1)_{1\leq i\leq k}$, we know that $n_i^N = n_i$ for all N (large enough) and all $1 \leq i \leq k$.

Now, we show that for each $1 \leq i \leq k, 1 \leq j \leq n_i - 1$, $t^N_{i,j}$ converges to $t_{i,j}$. As the sequence $(t^N_{i,j})_N$ is bounded, it is sufficient to show that $t_{i,j}$ is its only limit point. Let s be a limit of a subsequence $(t^{\varphi(N)}_{i,j})_N$.

We show that $s = t_{i,j}$. If $s > t_{i,j}$ there would exist some $r \in A \cap]t_{i,j}, s[$ satisfying that $g_{\varphi(N)}(r)$ converges to $g(r)$. This is not possible because, as $r < s = \lim_N t^{\varphi(N)}_{i,j}$, $g_{\varphi(N)}(r)_i \leq j - 1$ for N large enough, and as $r > t_{i,j}$, $g(r)_i \geq j$. For the same reason, it is not possible to have $s < t_{i,j}$. As a consequence, $t_{i,j}$ is the only limit point of the bounded sequence $\left(t^N_{i,j}\right)_N$. This implies the convergence of $t^N_{i,j}$ to $t_{i,j}$. In the rest of the proof, let us re-index the set $\{t_{i,j} : 1 \leq i \leq k, 1 \leq j \leq n_i - 1\}$ as $\{s_i : 1 \leq i \leq n\}$ where $n = \sum_{i=1}^k (n_i - 1)$, such that $s_1 < s_2 < \ldots < s_n$. And we consider the same indexes for the points $t^N_{i,j}$ $(1 \leq i \leq k, 1 \leq j \leq n_i - 1)$.

To prove the convergence of g_N to g in $D([0,T],\mathbb{R})$, we just have to define the sequence of functions $(\lambda_N)_N$ such that each λ_N is the function that is linear on each interval $[s_i^N, s_{i+1}^N]$ and that satisfies $\lambda_N(s_i^N) = s_i$. These functions verify $g_N = g \circ \lambda_N$ and

$$||\lambda_N - Id||_{\infty,[0,T]} = \max_{1\leq i\leq n} |s_i - s_i^N| = \max_{\substack{1\leq i\leq k \\ 1\leq j\leq n_i-1}} |t_{i,j} - t_{i,j}^N| \underset{n\to\infty}{\longrightarrow} 0.$$

This proves the result. □

C.4 Analytical lemmas

The first lemma of this section states that if a series of non-negative terms has a growth of order $n^{1-\varepsilon}$, then a subsequence of its general term vanishes. This property is used to prove Proposition 6.1.8.

Lemma C.4.1. *Let $(u_n)_{n\geq 0}$ be a sequence of non-negative real numbers, and $S_n = \sum_{k=0}^n u_k$ $(n \in \mathbb{N})$. If there exists $0 < \varepsilon < 1$, such that for all $n \in \mathbb{N}$, $S_n \leq Cn^{1-\varepsilon}$, then there exists a subsequence of $(u_n)_{n\geq 0}$ that converges to 0.*

Proof. We construct an increasing sequence of positive integers $(n_k)_{k\geq 1}$ such that for all $k \in \mathbb{N}^*$,

$$u_{n_k} \leq Cn_k^{-\varepsilon/2}. \tag{C.4}$$

Let $n_1 = 1$. By hypothesis, $u_1 = S_1 \leq C$. Let $k \in \mathbb{N}^*$, assume that the k first terms of the sequence $n_1, ..., n_k$ are defined such that (C.4) holds. Then we put $n_{k+1} := \min\{n > n_k : u_n \leq Cn^{-\varepsilon/2}\}$. We only have to prove that the above set is not empty. This is shown by contradiction. If the set was empty, this would imply that for all $n > n_k$, $u_n > Cn^{-\varepsilon/2}$. Consequently, for all $n > n_k$,

$$\begin{aligned} S_n = S_{n_k} + \sum_{i=n_k+1}^{n} u_i > S_{n_k} + C\sum_{i=n_k+1}^{n} i^{-\varepsilon/2} \geq S_{n_k} + C\int_{n_k+1}^{n+1} \frac{dt}{t^{\varepsilon/2}} \\ = S_{n_k} + \frac{C}{1-\varepsilon/2}(n+1)^{1-\varepsilon/2} - \frac{C}{1-\varepsilon/2}(n_k+1)^{1-\varepsilon/2}. \end{aligned}$$

Letting $n \to \infty$ in the above inequality we see that it is in contradiction with the hypothesis that for all $n \in \mathbb{N}^*$, $S_n \leq Cn^{1-\varepsilon}$. □

These two following lemmas were used to prove Theorem B.0.1 in a previous version. Although they are not used in the thesis, we still state and prove them because they are technical but not complicated, and even if they seem classical, we have not found their equivalent in the literature.

Lemma C.4.2. *For each $n \in \mathbb{N}$, let E_n be a topological space, and X_n be an E_n-valued random variable that is tight on E_n. Then $(X_n)_{n\in\mathbb{N}}$ is tight on $\prod_{n\in\mathbb{N}} E_n$ (endowed with the product topology).*

Proof. For all $\varepsilon > 0$ and for each $n \in \mathbb{N}$, let K_n^ε be a compact set of E_n such that

$$\mathbb{P}(X_n \notin K_n^\varepsilon) < \varepsilon.$$

Considering $L_\varepsilon = \prod_{n\in\mathbb{N}} K_n^{\varepsilon/2^n}$, which is a compact set of $\prod_{n\in\mathbb{N}} E_n$ (see Tykhonoff's theorem), we have

$$\mathbb{P}\left((X_n)_{n\in\mathbb{N}} \notin L_\varepsilon\right) = \mathbb{P}\left(\exists n \in \mathbb{N}, X_n \notin K_n^{\varepsilon/2^n}\right) \leq \sum_{n\in\mathbb{N}} \mathbb{P}\left(X_n \notin K_n^{\varepsilon/2^n}\right) < 2\varepsilon.$$

This proves the tightness of $(X_n)_n$ □

Lemma C.4.3. *For each $n \in \mathbb{N}$, let (E_n, d_n) be a separable metric space. Then $\prod_{n\in\mathbb{N}} E_n$ (endowed with the product topology) is a separable metric space.*

Proof. Firstly, we can assume that for each $n \in \mathbb{N}$, for all $x, y \in E_n, d_n(x,y) < 1$, since d_n is topology equivalent to the metric $\widetilde{d}_n(x,y) = \frac{d_n(x,y)}{1+d_n(x,y)}$. As a classical result, we know that the metric

$$d(x,y) = \sum_{n\in\mathbb{N}} \frac{d_n(x_n, y_n)}{2^n}$$

defines the product topology on $\prod_{n\in\mathbb{N}} E_n$.

For each $n \in \mathbb{N}$, let A_n be a dense countable subset of (E_n, d_n). We define the set

$$B = \bigcup_{n\in\mathbb{N}} \left(A_1 \times A_2 \times \ldots \times A_n \times \{0\} \times \{0\} \times \ldots\right),$$

which is a countable subset of $\prod_{n\in\mathbb{N}} E_n$.

Now we prove that B is dense. Let $x = (x_n)_{n\in\mathbb{N}} \in \prod_{n\in\mathbb{N}} E_n$ and $\varepsilon > 0$. We consider $n_\varepsilon \in \mathbb{N}$ such that $1/2^{n_\varepsilon} < \varepsilon$, and for each $0 \leq n \leq n_\varepsilon$, we choose some $y_n \in A_n$ such that $d_n(x_n, y_n) < \varepsilon$. For each $n > n_\varepsilon$, we define $y_n = 0$. We have

$$d(x,y) = \sum_{n=0}^{n_\varepsilon} \frac{d_n(x_n, y_n)}{2^n} + \sum_{n>n_\varepsilon} \frac{d_n(x_n, y_n)}{2^n} \leq 2\varepsilon + \frac{1}{2^{n_\varepsilon}} \leq 3\varepsilon.$$

□

Bibliography

Abi Jaber, E., Cuchiero, C., Larsson, M., and Pulido, S. (2019). A weak solution theory for stochastic Volterra equations of convolution type. *arXiv:1909.01166 [math]*.

Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. A. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.

Aldous, D. (1983). Exchangeability and related topics. In *Ecole d'Eté de Probabilités de Saint-Flour: XIII - 1983*, number 1117 in Lecture notes in mathematics. Springer-verlag edition.

Andreis, L., Dai Pra, P., and Fischer, M. (2018). McKean–Vlasov limit for interacting systems with simultaneous jumps. *Stochastic Analysis and Applications*, 36(6):960–995. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07362994.2018.1486202.

Bahouri, H., Chemin, J.-Y., and Danchin, R. (2011). *Fourier Analysis and Nonlinear Partial Differential Equations*, volume 343 of *Grundlehren der mathematischen Wissenschaften*. Springer.

Bauwens, L. and Hautsch, N. (2009). *Modelling financial high frequency data using point processes*. Springer Berlin Heidelberg.

Billingsley, P. (1999). *Convergence of Probability Measures*. Wiley Series In Probability And Statistics, second edition.

Brémaud, P. and Massoulié, L. (1996). Stability of Nonlinear Hawkes Processes. *The Annals of Probability*, 24(3):1563–1588.

Brown, T. (1978). A Martingale Approach to the Poisson Convergence of Simple Point Processes. *The Annals of Probability*, 6(4):615–628.

Carmona, R., Delarue, F., and Lacker, D. (2016). Mean field games with common noise. *The Annals of Probability*, 44(6):3740–3803.

Chaudru de Raynal, P. E. (2020). Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Processes and their Applications*, 130(1):79–107.

Chevallier, J. (2013). Détection de motifs de dépendance entre neurones. M2 internship report, Université de Nice Sophia-Antipolis.

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. Springer, second edition.

Davis, M. H. A. (1993). *Markov Models and Optimization*. Springer Science+Business Media Dordrecht, first edition.

De Masi, A., Galves, A., Löcherbach, E., and Presutti, E. (2015). Hydrodynamic Limit for Interacting Neurons. *Journal of Statistical Physics*, 158(4):866–902.

Delattre, S., Fournier, N., and Hoffmann, M. (2016). Hawkes processes on large networks. *The Annals of Applied Probability*, 26(1):216–261.

Ditlevsen, S. and Löcherbach, E. (2017). Multi-class Oscillating Systems of Interacting Neurons. *Stochastic Processes and their Applications*, 127(6):1840–1869.

Duarte, A., Löcherbach, E., and Ost, G. (2018). Stability, Convergence to Equilibrum and Simulation of Non-linear Hawkes Processes with Memory Kernels given by the Sum of Erlang Kernels. *ArXiv e-prints*, page arXiv:1610.03300v3.

El Karoui, N. and Méléard, S. (1990). Martingale measures and stochastic calculus. *Probability Theory and Related Fields*, 84(1):83–101.

Erny, X. (2020). A convergence criterion for systems of point processes from the convergence of their stochastic intensities. *arXiv:2003.11400 [math]*. arXiv: 2003.11400.

Erny, X. (2021). Well-posedness and propagation of chaos for McKean-Vlasov equations with jumps and locally Lipschitz coefficients. *arXiv:2102.06472 [math]*. arXiv: 2102.06472.

Erny, X., Löcherbach, E., and Loukianova, D. (2019). Mean field limits for interacting Hawkes processes in a diffusive regime. *arXiv:1904.06985 [math]*.

Erny, X., Löcherbach, E., and Loukianova, D. (2020). Conditional propagation of chaos for mean field systems of interacting neurons. *arXiv:1909.02925 [math]*. arXiv: 1909.02925.

Erny, X., Löcherbach, E., and Loukianova, D. (2021). White-noise driven conditional McKean-Vlasov limits for systems of particles with simultaneous and random jumps. *arXiv:2103.04100 [math]*. arXiv: 2103.04100.

Ethier, S. and Kurtz, T. (2005). *Markov Processes. Characterization and Convergence*. Wiley Series In Probability And Statistics.

Fischer, M. and Livieri, G. (2016). Continuous time mean-variance portfolio optimization through the mean field approach. *ESAIM: Probability and Statistics*, 20:30–44. Publisher: EDP Sciences.

Fournier, N. and Guillin, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3):707–738.

Fournier, N. and Löcherbach, E. (2016). On a toy model of interacting neurons. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, 52:1844–1876.

Fournier, N. and Meleard, S. (2002). A Stochastic Particle Numerical Method for 3D Boltzmann Equations without Cutoff. *Mathematics of Computation*, 71(238):583–604. Publisher: American Mathematical Society.

Fournier, N. and Mischler, S. (2016). Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Annals of Probability*, 44(1):589–627. Publisher: Institute of Mathematical Statistics.

Gärtner, J. (1988). On the McKean-Vlasov Limit for Interacting Diffusions. *Mathematische Nachrichten*, 137(1):197–248. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mana.19881370116.

Graham, C. (1992). McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications*, 40(1):69–82.

Graham, C. (2008). Chaoticity for Multiclass Systems and Exchangeability Within Classes. *Journal of Applied Probability*, 45(4):1196–1203. Publisher: Cambridge University Press.

Graham, C. and Méléard, S. (1997). Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Annals of Probability*, 25(1):115–132. Publisher: Institute of Mathematical Statistics.

Grönwall's inequality (2019). Grönwall's inequality — Wikipedia, The Free Encyclopedia.

Grün, S., Diedsmann, M., and Aertsen, A. M. (2010). *Analysis of Parallel Spike Trains*. Rotter, Springer series in computational neurosciences.

Guionnet, A. (1997). Averaged and quenched propagation of chaos for spin glass dynamics. *Probability Theory and Related Fields*, 109(2):183–215.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90.

Helmstetter, A. and Sornette, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research*, 107.

Hewlett, P. (2006). *Clustering of order arrivals, price impact and trade path optimisation*.

Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*. North-Holland Publishing Company, second edition.

Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*. Springer-Verlag BerlinHeidelberg NewYork, second edition.

Kallenberg, O. (1997). *Foundations of Modern Probability*. Probability and Its Applications. Springer-Verlag, New York.

Khasminskii, R. (2012). *Stochastic stability of differential equations*. Springer, second edition.

Kunita, H. (1986). *Lectures on Stochastic Flows and Applications for the Indian Institute Of Science Bangalore*. Tata Institute Of Fundamental Research Springer-Verlag.

Kunita, H. (1990). *Stochastic flows and stochastic differential equations*. Cambridge University Press.

Kurtz, T. (2014). Weak and strong solutions of general stochastic models. *Electronic Communications in Probability*, 19(58).

Lacker, D. (2018). On a strong form of propagation of chaos for McKean-Vlasov equations. *Electronic Communications in Probability*, 23. Publisher: The Institute of Mathematical Statistics and the Bernoulli Society.

Last, G. and Penrose, M. (2017). *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, Cambridge.

Lu, X. and Abergel, F. (2018). High dimensional Hawkes processes for limit order books Modelling, empirical analysis and numerical calibration. *Quantitative Finance*, pages 1–16.

Major, P. (1978). On the invariance principle for sums of independent identically distributed random variables. *Journal of Multivariate Analysis*, 8(4):487–517.

Méléard, S. (1998). Stochastic approximations of the solution of a full Boltzmann equation with small initial data. *ESAIM: Probability and Statistics*, 2:23–40. Publisher: EDP Sciences.

Meyn, S. P. and Tweedie, R. L. (1992). Stability of Markovian Processes I: Criteria for Discrete-Time Chains. *Advances in Applied Probability*, 24(3):542–574.

Meyn, S. P. and Tweedie, R. L. (1993). Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Applied Probability Trust*, 25:518–548.

Mishura, Y. S. and Veretennikov, A. Y. (2020). Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*. Publisher: American Mathematical Society.

Ogata, Y. (1999). Seismicity Analysis through Point-process Modeling: A Review. *Pure and applied geophysics*, 155:471–507.

Pillow, J. W., Wilson, M. A., and Brown, E. N. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454:995–999.

Reynaud-Bouret, P., Rivoirard, V., Grammont, F., and Tuleau-Malot, C. (2014). Goodness-of-Fit Tests and Nonparametric Adaptive Estimation for Spike Train Analysis. *Journal of Mathematical Neuroscience*, 4:3 – 330325.

Reynaud-Bouret, P. and Schbath, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *The Annals of Statistics*, 38(5):2781–2822.

Sznitman, A.-S. (1989). Topics in propagation of chaos. In *Ecole d'Eté de Probabilités de Saint-Flour: XIX - 1989*, number 1464 in Lecture notes in mathematics, pages 167–251. Springer, Berlin. OCLC: 23253880.

Tanaka, H. (1978). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 46(1):67–105.

Villani, C. (2008). *Optimal transport, old and new*. Springer.

Walsh, J. B. (1986). An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV - 1984*, Lecture Notes in Mathematics, pages 265–439. Springer, Berlin, Heidelberg.

Y. Kagan, Y. (2009). Statistical Distributions of Earthquake Numbers: Consequence of Branching Process. *Geophys. J. Int.*, 180.

Zhou, K., Zha, H., and Song, L. (2013). *Learning triggering kernels for multi-dimensional Hawkes processes.*

ÉCOLE DOCTORALE
de mathématiques
Hadamard (EDMH)

Titre : Limite de grande échelle de systèmes de particules en interaction avec sauts simultanés en régime diffusif

Mots clés : processus de Hawkes, propagation du chaos conditionnelle, équation McKean-Vlasov, système de particules, équation différentielle stochastique avec sauts

Résumé : Nous étudions des limites de grande échelle de systèmes de particules en interaction. Dans chaque modèle, nous considérons des systèmes à N particules en interaction champ moyen avec une force d'interaction aléatoire et centrée de l'ordre de $N^{-1/2}$, et démontrons la convergence de ces systèmes quand N tend vers l'infini. Dans le cadre plus classique où la force des interactions est N^{-1}, on constate une propagation du chaos. Dans notre cadre, nous montrons des propagations du chaos conditionnelles. Nous étudions trois types de modèles. D'abord, nous étudions des sytèmes de processus de Hawkes dont nous montrons la convergence des intensités stochastiques. Comme ces intensités sont, dans notre cadre, des processus de Markov, nous montrons leur convergence et obtenons une vitesse de convergence en exploitant leurs générateurs. Puis nous étudions une généralisation du premier type de modèle où les méthodes précédentes ne marchent plus. À la place, nous utilisons un nouveau type de problème martingale pour montrer que les mesures empiriques des systèmes de particules convergent vers la mesure directrice du système limite. Dans ces modèles, les équations de type McKean-Vlasov apparaissent naturellement à la limite, et lorsque les interactions sont spatiales, les EDS limites sont dirigées par des bruits blancs. Finalement, nous étudions des systèmes McKean-Vlasov avec une force d'interaction en N^{-1} et avec des coefficients localement lipschitziens. Nous montrons que ces équations sont bien posées au sens fort, et la propagation du chaos.

Title: Large scale limits for interacting particle systems with simultaneous jumps in diffusive regime

Keywords: Hawkes processes, conditional propagation of chaos, McKean-Vlasov equation, particle system, stochastic differential equation with jumps

Abstract: We study large scale limits for interacting particle systems. In each model, we consider mean field interacting $N-$particle systems with stochastic and centred interacting strength of order $N^{-1/2}$, and we prove the convergence of these systems as N goes to infinity. In the more classical framework where the interacting strength is N^{-1}, we observe a propagation of chaos. In our framework, we prove conditional propagations of chaos. We study three kinds of models. The first kind concerns systems of Hawkes processes. Their convergence is shown from the convergence of their stochastic intensities. As these intensities are, in our frame, Markov processes, we show their convergence and obtain a convergence speed using their generators. The second kind of model is a generalization of the first one where the previous methods cannot be applied. Instead we use a new type of martingale problem to show the convergence of the empirical measures of the particle systems to the directing measure of the limit system. In these models, McKean-Vlasov equations appear naturally at the limit, and when the interactions are spatial, the limit SDEs are directed by white noises. Finally, we study McKean-Vlasov systems with interacting strength in N^{-1} with coefficients that are locally Lipschitz. We prove the strong well-posedness of these equations and the propagation of chaos.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France