



HAL
open science

Approximation par champ-moyen : le couplage à la Sznitman pour les nuls

Julien Chevallier

► **To cite this version:**

Julien Chevallier. Approximation par champ-moyen : le couplage à la Sznitman pour les nuls. 2017. hal-01433292v1

HAL Id: hal-01433292

<https://hal.science/hal-01433292v1>

Preprint submitted on 12 Jan 2017 (v1), last revised 31 Jan 2017 (v2)

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

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



Distributed under a Creative Commons Attribution 4.0 International License

APPROXIMATION PAR CHAMP-MOYEN : LE COUPLAGE À LA SZNITMAN POUR LES NULS

Julien Chevallier

12 janvier 2017

La motivation des approches par champ-moyen est de relier deux échelles de modélisation :

- une échelle **microscopique** pour laquelle on dispose d'un modèle détaillé décrivant la dynamique de nombreuses « particules »¹ en **interaction**,
- et une échelle **macroscopique** pour laquelle on dispose d'un modèle simplifié qui décrit la distribution statistique de ses particules.

Le but de ces notes est double.

1. Proposer une application, sur un exemple² très simple, de la technique de couplage proposée par Sznitman dans [7].
2. Mentionner les principales difficultés rencontrées lorsque le modèle que l'on considère est plus complexe.

Les pré-requis à une bonne lecture de ces notes sont des connaissances de base du **calcul stochastique**. La rédaction de ces notes a été réalisée de manière conjointe avec la préparation de l'exposé intitulé « Méthode d'approximation de champ-moyen : couplage à la Sznitman sur un exemple jouet. » et présenté le 27 Janvier 2017 dans le cadre de la 5ème journée **Young Statisticians and Probabilists**.

I Le modèle jouet

Soit n un entier qui désigne le nombre de particules dans le système. Les n particules sont décrites par leurs positions respectives X_t^1, \dots, X_t^n , pour tout temps t positif. Nous supposons dans toute la suite que les particules vérifient l'équation différentielle stochastique

$$\forall i = 1, \dots, n, \quad dX_t^i = - \left(X_t^i - \frac{1}{n} \sum_{j=1}^n X_t^j \right) dt + dB_t^i, \quad (1)$$

où les $(B_t^i)_{t \geq 0}$ sont des mouvements Browniens indépendants. Nous supposons que les valeurs initiales X_0^1, \dots, X_0^n sont indépendantes et identiquement distribuées (i.i.d.) de loi centrée μ_0 . De plus, les mouvements Browniens et les valeurs initiales sont indépendants.

En l'absence du bruit, le système (1) se réduit à une équation différentielle ordinaire : chaque particule est **attirée**, à vitesse exponentielle, **vers la moyenne de la position des particules**.

1. Le mot particule est un terme générique qui peut désigner : une particule physique, un neurone, un agent financier, etc. Cette terminologie vient du fait que, à l'origine, les limites de champ-moyen concernaient principalement des systèmes de particules physiques. Ces particules en interaction hamiltonienne sont décrites à la limite par une équation cinétique (Boltzmann, Vlasov). La littérature sur ce sujet est très dense [1, 4, 5, 6, 7].

2. Le modèle jouet considéré dans ces notes est une version simplifiée du *laboratory example* étudié dans la section I.1. de [7].

Proposition 1. *Le système d'EDS (1) est linéaire. En particulier, nous avons existence et unicité trajectorielle des solutions.*

Démonstration. En notant $\mathbf{X}_t = (X_t^1, \dots, X_t^n)$ et $\mathbf{B}_t = (B_t^1, \dots, B_t^n)$, le système (1) se réécrit

$$d\mathbf{X}_t = A\mathbf{X}_t + d\mathbf{B}_t,$$

avec $A = -Id + n^{-1}J$ où Id est la matrice identité et

$$J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

Pour finir, l'existence et l'unicité des solutions sont données par le théorème de Cauchy-Lipschitz stochastique. \square

II Le processus limite

La technique de couplage repose sur une **heuristique de type « loi des grands nombres »** et l'introduction d'un processus limite \bar{X} dont la dynamique est caractérisée par une version « limite » du système (1). En effet, on s'attend à ce que les particules soient asymptotiquement i.i.d. et donc, par la loi des grands nombres, on s'attend à ce que le terme de moyenne empirique $n^{-1} \sum_{j=1}^n X_t^j$ soit remplacé par un terme d'espérance : on peut par exemple penser à $\mathbb{E}[X_t^1]$, mais, encore une fois, on s'attend à ce que ce terme soit remplacé par l'espérance du processus limite, i.e. $\mathbb{E}[\bar{X}_t]$.

Dans la suite de ces notes, nous nous intéressons donc, en comparaison avec (1), à l'équation limite

$$d\bar{X}_t = -(\bar{X}_t - \mathbb{E}[\bar{X}_t]) dt + dB_t, \quad (2)$$

où $(B_t)_{t \geq 0}$ est un mouvement Brownien et \bar{X}_0 est distribué selon μ_0 . Le caractère bien posé de cette équation n'est pas trivial. En effet, le terme $\mathbb{E}[\bar{X}_t]$ induit une non-linéarité dite de *McKean-Vlasov* : **le processus \bar{X} dépend de sa propre loi**. Cependant, dans le cadre de notre modèle jouet, il existe une simplification majeure qui nous donne une preuve très simple de l'existence et de l'unicité des solutions de (2).

Proposition 2. *Sous l'hypothèse fondamentale que la distribution initiale μ_0 est centrée, l'équation (2) est équivalente à*

$$d\bar{X}_t = -\bar{X}_t + dB_t. \quad (3)$$

Cette dernière est une EDS linéaire et correspond à un cas très simple du processus de Ornstein-Uhlenbeck : nous avons existence et unicité trajectorielle des solutions. De plus, la solution de (3) est donnée par

$$\bar{X}_t = \bar{X}_0 e^{-t} + \int_0^t e^{-(t-s)} dB_s.$$

Démonstration. Raisonnons par équivalences. Le processus \bar{X} est solution de (2) si et seulement si (ssi),

$$\bar{X}_t = \bar{X}_0 - \int_0^t (\bar{X}_s - \mathbb{E}[\bar{X}_s]) ds + \int_0^t dB_s,$$

ssi (en utilisant le fait que μ_0 est centrée),

$$\begin{cases} \mathbb{E}[\bar{X}_t] = 0 - \int_0^t 0 ds + 0 = 0, \\ d\bar{X}_t = -\bar{X}_t + dB_t, \end{cases}$$

ssi le processus \bar{X} est solution de (3). \square

Remarque 3. Si on note $u(t, \cdot)$ la loi de \bar{X}_t , alors u est solution de

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} (xu). \quad (4)$$

Démonstration. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction de classe \mathcal{C}^2 . Par la formule d'Itô, on a

$$f(\bar{X}_t) = f(\bar{X}_0) + \int_0^t f'(\bar{X}_s) dB_s + \int_0^t \left(\frac{1}{2} f''(\bar{X}_s) - f'(\bar{X}_s) \bar{X}_s \right) ds.$$

En prenant l'espérance de l'équation ci-dessus, on obtient alors

$$\int_{\mathbb{R}} f(x) u(t, dx) = \int_{\mathbb{R}} f(x) u(0, dx) + \int_0^t \int_{\mathbb{R}} \left(\frac{1}{2} f''(x) - x f'(x) \right) u(t, dx) dt,$$

qui est la version faible de (4). □

III Le couplage

Soit $\Omega := (\mathbb{R} \times \mathcal{C}(\mathbb{R}_+, \mathbb{R}))^{\mathbb{N}}$ notre espace de probabilité muni de $(\mu_0 \otimes W)^{\otimes \mathbb{N}}$: les coordonnées réelles correspondent aux conditions initiales i.i.d. (nous les notons Y^i) et les coordonnées trajectorielles correspondent aux mouvements Browniens i.i.d. (nous les notons B_t^i) qui dirigent la dynamique. Il est possible de construire sur Ω les processus suivants :

1. les processus $(X_t^i)_{t \geq 0}$, $i = 1, \dots, n$, vérifiant

$$\begin{cases} dX_t^i = - \left(X_t^i - \frac{1}{n} \sum_{j=1}^n X_t^j \right) dt + dB_t^i, \\ X_0^i = Y^i, \end{cases}$$

2. et les processus $(\bar{X}_t^i)_{t \geq 0}$, $i \geq 1$, vérifiant

$$\begin{cases} d\bar{X}_t^i = -\bar{X}_t^i + dB_t^i, \\ \bar{X}_0^i = Y^i. \end{cases}$$

Autrement dit, **nous couplons** le système en interaction avec des copies i.i.d. du processus limite via leurs conditions initiales et leurs mouvements Browniens.

Théorème 4. Supposons que μ_0 admette un moment d'ordre 2 égal à $v_0 < +\infty$. Alors, pour tout $i = 1, \dots, n$ et tout horizon de temps $T > 0$, il existe une constante C telle que

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t^i - \bar{X}_t^i| \right] \leq \left(v_0 + \frac{1}{2} \right) T \exp(2T) n^{-1/2} \quad (5)$$

Démonstration. Par le couplage des conditions initiales et mouvements Browniens, il est clair que

$$X_t^i - \bar{X}_t^i = \int_0^t - \left(X_s^i - \bar{X}_s^i \right) + \left(\frac{1}{n} \sum_{j=1}^n X_s^j - \mathbb{E} [\bar{X}_s^1] \right) ds$$

et donc

$$|X_t^i - \bar{X}_t^i| \leq \int_0^t |X_s^i - \bar{X}_s^i| + \left| \frac{1}{n} \sum_{j=1}^n (X_s^j - \bar{X}_s^j) \right| + \left| \frac{1}{n} \sum_{j=1}^n \bar{X}_s^j - \mathbb{E} [\bar{X}_s^1] \right| ds.$$

Ainsi, en notant $\delta(t) = \mathbb{E} \left[\sup_{0 \leq s \leq t} \left| X_s^i - \bar{X}_s^i \right| \right]$, et en prenant l'espérance de l'équation ci-dessus, nous avons

$$\delta(t) \leq 2 \int_0^t \delta(s) ds + \int_0^t \mathbb{E} \left[\left| \frac{1}{n} \sum_{j=1}^n \bar{X}_s^j - \mathbb{E} [\bar{X}_s^1] \right| \right] ds. \quad (6)$$

Notons que $D_n(t) := \mathbb{E} [|n^{-1} \sum_{j=1}^n \bar{X}_t^j - \mathbb{E}[\bar{X}_t^1]|]$ est le moment d'ordre 1 d'une moyenne empirique de variables aléatoires centrées indépendantes. En utilisant l'inégalité de Cauchy-Schwarz, nous avons donc

$$D_n(t) \leq \frac{1}{\sqrt{n}} \text{Var}(\bar{X}_t^1). \quad (7)$$

Or,

$$\begin{aligned} \text{Var}(\bar{X}_t^1) &= \text{Var}(\bar{X}_0^1) + \mathbb{E} \left[\left(\int_0^t e^{-(t-s)} dB_s \right)^2 \right] \\ &= v_0 + \int_0^t e^{-2(t-s)} ds = v_0 + \frac{1}{2} [1 - e^{-2t}] \leq v_0 + \frac{1}{2}. \end{aligned} \quad (8)$$

Finalement, en regroupant les équations (6), (7) et (8), nous obtenons

$$\forall 0 \leq t \leq T, \quad \delta(t) \leq 2 \int_0^t \delta(s) ds + \left(v_0 + \frac{1}{2} \right) t n^{-1/2},$$

et le Lemme de Gronwall permet de conclure. \square

Un corollaire immédiat du théorème précédent est le résultat suivant.

Corollaire 5. *Pour tout $T > 0$ et $i \geq 1$, la trajectoire $(X_t^i)_{0 \leq t \leq T}$ converge en loi vers $(\bar{X}_t^i)_{0 \leq t \leq T}$. Ce résultat de convergence est également valide pour un temps $t \geq 0$ fixé, d'où*

$$\mathcal{L}(X_t^i) \xrightarrow{n \rightarrow +\infty} \mathcal{L}(\bar{X}_t^i) = u(t, \cdot).$$

Rappelons que $u(t, \cdot)$ est solution de (4) (voir Remarque 3).

Démonstration. Il suffit de voir que le contrôle (5) donne un contrôle de la distance de Wasserstein entre $\mathcal{L}((X_t^i)_{0 \leq t \leq T})$ et $\mathcal{L}((\bar{X}_t^i)_{0 \leq t \leq T})$ et que la distance de Wasserstein métrise la convergence en loi. \square

Notons que l'utilisation d'une théorie plus poussée (**propagation du chaos**, voir [7, Proposition 2.2], [4, Proposition 4.2] ou bien les notes de cours de DEA "Limite de champ moyen" de Cédric Villani) permet d'obtenir un résultat de convergence pour des mesures empiriques. De plus, cela donne un **lien fort** entre le système de particules (1) et la solution u de l'EDP (4).

Proposition 6. *En comparaison avec l'énoncé précédent, nous avons les convergences suivantes*

$$\frac{1}{n} \sum_{i=1}^n \delta_{(X_t^i)_{0 \leq t \leq T}} \xrightarrow{n \rightarrow +\infty} \mathcal{L}((\bar{X}_t^i)_{0 \leq t \leq T}),$$

et

$$\frac{1}{n} \sum_{i=1}^n \delta_{X_t^i} \xrightarrow{n \rightarrow +\infty} u(t, \cdot).$$

La première convergence est une convergence en probabilité pour des variables aléatoires à valeur dans $\mathcal{P}(\mathcal{C}([0, T], \mathbb{R}))$ et la seconde est une convergence en probabilités pour des variables aléatoires à valeur dans $\mathcal{P}(\mathbb{R})$.

IV Illustration par simulations

En premier lieu, nous illustrons le couplage introduit au début de la section III en superposant des trajectoires associées au système de n particules et leurs trajectoires limites correspondantes. De plus, nous faisons varier la taille du système n pour voir que l'approximation donnée par les processus limites s'améliore à mesure que n prend des valeurs plus grandes : la distance entre la courbe en couleur pleine et la courbe en pointillé correspondante diminue.

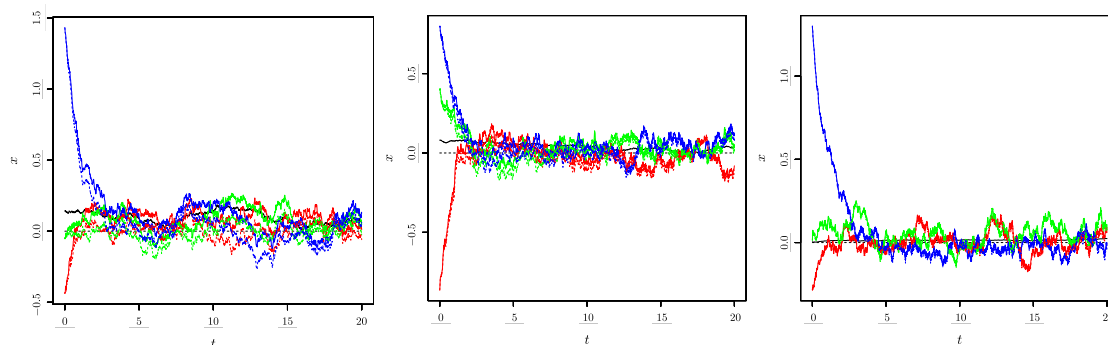


FIGURE 1 – Sont tracées sur chacun des graphes les trajectoires de X^1 , X^2 et X^3 (en couleur pleine), les trajectoires de \bar{X}^1 , \bar{X}^2 et \bar{X}^3 (en couleur pointillée), la trajectoire de la moyenne empirique $n^{-1} \sum_{i=1}^n X^i$ (en noir plein) et la fonction constante égale à 0 (en noir pointillé). De gauche à droite, la valeur de n est égale à 10, 100 et 1000. Notons que les conditions initiales $X_0^i = \bar{X}_0^i$ sont i.i.d. de loi normale centrée réduite.

Dans le graphe ci-dessous, nous proposons une illustration de la vitesse de convergence obtenue dans le Théorème 4. Chaque point du graphe correspond, pour une valeur de n donnée, à une estimation du terme d'espérance qui apparaît dans (5) via une méthode de Monte-Carlo avec 50 itérations (pour chaque point).

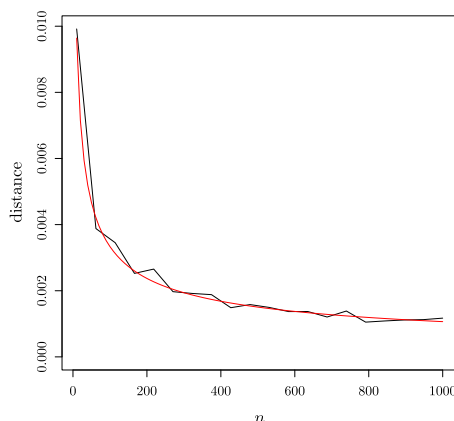


FIGURE 2 – En noir, le graphe de l'estimation (par Monte-Carlo) de l'espérance qui apparaît dans (5) ($T = 1$). En rouge, ajustement par une fonction de type $n \mapsto Cn^{-1/2}$, $C \in \mathbb{R}$.

V Pour aller plus loin

V.1 Interactions non-linéaires

Le terme d'interaction qui apparaît dans l'exemple jouet présenté dans ces notes est linéaire. Plus généralement, il peut être intéressant de regarder un terme d'interaction donné par un noyau **non linéaire** $b : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, et donc le système³

$$\forall i = 1, \dots, n, \quad dX_t^i = \frac{1}{n} \sum_{j=1}^n b(X_t^i, X_t^j) dt + dB_t^i. \quad (9)$$

En comparaison, le système (1) est obtenu avec le noyau linéaire $b(x_1, x_2) = x_2 - x_1$. La majeure différence dans l'utilisation de l'argument de couplage entre le modèle (1) et le modèle (9) réside dans la **plus grande complexité de l'équation limite**. En comparaison avec (2), nous obtenons ici

$$d\bar{X}_t = \int b(\bar{X}_t, y) u(t, dy) dt + dB_t, \quad (10)$$

où $u(t, \cdot)$ est la loi de \bar{X}_t . Dans ce cadre général, nous ne pouvons plus utiliser la simplification énoncée dans la Proposition 2. Il faut traiter la non-linéarité⁴ de l'équation limite (10) : **les trajectoires du processus \bar{X} dépendent de sa propre loi**. Sous l'hypothèse que le noyau b est **lipschitzien**, les principaux résultats énoncés dans la Proposition 2 et la Remarque 3 restent valides. Plus précisément, on a le résultat suivant.

Proposition 7 (voir Théorème 1.1 et Remarque 1.2 dans [7]). *Nous avons existence et unicité trajectorielle des solutions de (10). De plus, la loi des marginales temporelles du processus vérifie*

$$\frac{\partial u}{\partial t} = \frac{1}{2} \frac{\partial^2 u}{\partial x^2} - \frac{\partial}{\partial x} \left(\left[\int b(\cdot, y) u(t, dy) \right] u \right). \quad (11)$$

Une fois que l'existence du processus limite est garantie, il suffit de suivre (de manière quasiment identique) l'argument de couplage et les calculs menés dans la preuve du Théorème 4 (utiliser le fait que b est lipschitz).

V.2 Processus limite

Comme mentionnée dans la section précédente, la principale difficulté dans l'application de cet argument de couplage réside dans le caractère bien posé de l'équation limite (non-linéaire en général). Deux stratégies légèrement différentes sont possibles pour attaquer ce problème.

La première stratégie⁵ se place dans un cadre **probabiliste** et suit le schéma suivant :

1. Considérer une version linéarisée de (10) où u est remplacé par une mesure de probabilité quelconque m sur $\mathcal{C}([0, T], \mathbb{R})$.
2. Étudier la loi de la solution de cette équation linéarisée : elle dépend de m et nous la notons $\Phi(m)$.
3. Montrer que l'application Φ qui va, de l'espace des probabilités sur $\mathcal{C}([0, T], \mathbb{R})$, dans lui-même admet un unique point fixe : typiquement, on montre une propriété de contraction pour Φ .

3. Ce système est exactement le *laboratory example* étudié dans la section I.1. de [7] : on se restreint ici à la dimension $d = 1$.

4. Cette non-linéarité est qualifiée de type *McKean-Vlasov*.

5. Cette stratégie est utilisée dans [7].

La seconde stratégie repose sur une étude **analytique** de l'EDP que l'on espère obtenir – (11) par exemple – et suit le schéma suivant :

1. Montrer l'existence d'une solution v à l'EDP. Dans le fond, cela ressemble beaucoup au schéma précédent, mais transposé dans un cadre d'EDP.
2. Montrer l'existence et l'unicité des solutions pour l'équation (10) où u est remplacée par v : cette équation est devenue linéaire et donc plus simple à traiter. On note $u(t, \cdot)$ les marginales temporelles de la loi du processus solution.
3. Montrer que u est l'unique solution d'une certaine EDP, qui se trouve être également vérifiée par v . Par unicité, on a $u = v$, et donc le processus construit au point 2 est solution de (10).

V.3 Sortir du cadre des diffusions

Lorsque l'on sort du cadre des processus continus, la principale différence réside dans le fait que la distance la plus adaptée n'est pas forcément la norme L^1 qui apparaît dans (5).

Par exemple, dans le cadre des processus ponctuels (processus à saut pur), la distance en **variation totale** est plus adaptée. En comparaison avec (5), on préférera contrôler dans ce cas un terme de la forme

$$\mathbb{P} \left((X_t^i)_{0 \leq t \leq T} \neq (\bar{X}_t^i)_{0 \leq t \leq T} \right).$$

V.4 Ajouter une dimension spatiale

Sans nécessité de changements trop drastiques, l'argument de couplage à la Sznitman permet de donner des résultats d'approximation de champ-moyen avec une dimension spatiale dans le modèle.

Par exemple, supposons que les variables X_t^i représentent la valeur de l'action d'une entreprise au temps t . Chaque entreprise est localisée dans l'espace et on peut tenir compte du fait que la force d'interaction entre l'entreprise i et l'entreprise j dépend de leur localisation respective – par exemple par la distance (physique) qui les sépare – en plus de dépendre de leur cotation respective. Dans la suite, on note $y_i \in \mathcal{Y}$ la localisation de la i -ème particule. Ainsi, en comparaison avec (9), on pourrait considérer le système

$$\forall i = 1, \dots, n, \quad dX_t^i = \frac{1}{n} \sum_{j=1}^n K(y_i, y_j) b(X_t^i, X_t^j) dt + dB_t^i, \quad (12)$$

où $K : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$ est le noyau de dépendance spatiale. Dans ce cas, deux convergences sont à prendre en compte :

- la convergence des processus X^i vers des processus limites (dont la loi va dépendre de la localisation y_i),
- et la convergence de la mesure empirique sur \mathcal{Y} induite par les localisations, i.e. $n^{-1} \sum_{i=1}^n \delta_{y_i}$, vers une distribution limite.

Le lecteur intéressé pourra s'orienter vers les articles [2, 3] pour ce genre de considérations.

Bibliographie

- [1] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, volume 1955, pages 171–197, 1954.
- [2] E. Luçon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *The Annals of Applied Probability*, 24(5) :1946–1993, 2014.
- [3] E. Luçon and W. Stannat. Transition from gaussian to non-gaussian fluctuations for mean-field diffusions in spatial interaction. *arXiv preprint [arXiv:1502.00532](https://arxiv.org/abs/1502.00532)*, 2015.
- [4] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.
- [5] H. Spohn. Kinetic equations from hamiltonian dynamics : Markovian limits. *Reviews of Modern Physics*, 52(3) :569, 1980.
- [6] H. Spohn. *Large scale dynamics of interacting particles*. Springer Science & Business Media, 2012.
- [7] A.-S. Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.